Computers and Electronics in Agriculture 162 (2019) 143–153

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

Original papers

Automated segmentation of soybean plants from 3D point cloud using machine learning

Jing Zhou[a], Xiuqing Fu[a,b], Shuiqin Zhou[a,c], Jianfeng Zhou[a,*], Heng Ye[d], Henry T. Nguyen[d]

[a] Division of Food Systems and Bioengineering, University of Missouri, Columbia, MO 65211, USA
[b] College of Engineering, Nanjing Agricultural University, Nanjing 210031, China
[c] Fair Friend Institute of Electromechanics, Hangzhou Vocational & Technical College, Hangzhou 310018, China
[d] Division of Plant Sciences, University of Missouri, Columbia, MO 65211, USA

ARTICLE INFO

Keywords:
Plant segmentation
Machine learning
Soybean breeding
High-throughput phenotyping
Greenhouse

ABSTRACT

Image-based plant phenotyping has become a promising method for high-throughput measurement of plant traits in breeding programs. Plant geometric features that are essential for understanding plant growth can be obtained from the point cloud built using three-dimensional (3D) reconstruction of plant imagery data. A key task in the data processing pipeline is the automated and accurate segmentation of individual plants. Machine learning is a promising approach due to its strong ability in the extraction of details from images and has been successfully applied in plant leaf segmentation from two-dimensional (2D) images. The aim of this paper was to evaluate the performance of three machine learning methods, i.e. *boosting*, Support Vector Machine (SVM) and *K-means* clustering, in the segmentation of non-overlapped and overlapped soybean plants at early growth stages using 3D point cloud. Images of 75 soybean plants at two growth stages in a greenhouse were collected using an image-based high-throughput phenotyping platform and were used to develop 3D point cloud using the Structure from Motion (SfM) method. Plant features including position (coordinate *x*, *y*, and *z*), and color (*Red*, *Green*, *Blue*, *hue*, *saturation* and *Triangular Greenness Index*) were used for background removal and the separation of non-overlapped plants. A Histogram of Oriented Gradient (*HOG*) descriptor was used for the separation of overlapped plants. The percentage of mismatched points between manual and automated segmentation was calculated and results showed that *K-means* clustering had the least mean error rates (0.36% and 0.20%) for the background removal and the non-overlapped plant separation. The least mean error rate for the separation of overlapped plants was 2.57% using SVM with labeled *HOG* descriptor. The developed image segmentation pipeline was evaluated in a case study where 69 plants at different growth stages were continuously monitored. Results showed that it took three minutes on average for completing all procedures in the pipeline and the extracted features (i.e. height and shooting area) were able to quantify the plant growth.

1. Introduction

Soybean is the second largest crop grown in the United States with 36.4 million ha planted, yielding US$40.0 billion in 2017 (USDA-ERS, 2018). Soybean production, however, can be greatly influenced by abiotic stresses imposed by environmental factors such as drought, flood, salt, and heavy metals (Deshmukh et al., 2014). Fast development of new soybean varieties with high yield potential and stress resistance is of a particular significance in providing sufficient food for the growing population under current adverse environmental conditions. However, conventional phenotyping methods (manual measurements and visual observations with naked eyes) are time-consuming and labor-intensive and have become the bottleneck for fast developing new crop varieties in breeding programs. Recently, image-based phenotyping technologies for fast measurement of plant traits from plant shoot (above ground part) have been used in stress detection, crop classification and yield prediction in the studies of genetics, breeding, and precision agriculture. Vollmann et al. (2011) showed that the chlorophyll content of soybean leaves measured using a chlorophyll meter was highly correlated ($r = -0.937$) to the green color in the RGB images of the leaves at the R3 (beginning pod development) growth stage. Purcell (2000) measured the proportion of ground area covered by soybean canopy from air-born digital images and found that the measurements of canopy-coverage were highly correlated to the light interception that is important for crop yield and crop growth. Rousseau et al. (2013) found that the maximum quantum yield of photosystem II

* Corresponding author at: 211 Ag Engineering Building, Columbia, MO 65211, USA.
E-mail address: zhoujianf@missouri.edu (J. Zhou).

https://doi.org/10.1016/j.compag.2019.04.014
Received 11 December 2018; Received in revised form 8 April 2019; Accepted 10 April 2019
Available online 13 April 2019
0168-1699/

photochemistry (F_v/F_m) obtained using chlorophyll fluorescence images might be used as an indicator of plant stress to distinguish disease-infected and healthy tissues.

In the past few years, high-throughput phenotyping approaches have been used for fast measurement of plant traits in controlled environments (Fiorani and Schurr, 2013). Compared to field conditions, controlled environments offer a uniform growing condition that reduces the environmental effects on plant breeding (Poorter et al., 2012). The academic and industrial sectors have taken efforts to develop high-throughput phenotyping systems for controlled environments. For example, the GROWSCREEN FLUORO developed by Jansen et al. (2009) was used to detect the stress tolerance in rosette plants based on traits of leaf growth and chlorophyll fluorescence. The Phenovator designed by Flood et al. (2016) is capable of screening more than 1000 Arabidopsis plants for the measurements of photosynthesis, growth and multispectral reflectance multiple times per day. The Scanalyzer HTS and 3D system by LemnaTec GmbH were adopted by many studies to analyze plant leaf chemical properties (Pandey et al., 2017) and measure the diurnal patterns of leaf hyponasty and leaf size (Dornbusch et al., 2012).

Most of the existing high-throughput phenotyping platforms for controlled environments were designed to take images of single plants using two different ways, i.e. “plant-to-sensor” or “stop-and-action”. Plant-to-sensor platforms transfer individual plants to a closed chamber with a standard background where plants are screened with multiple sensors (for example, Scanalyzer HTS and 3D system). However, the plant-to-sensor platforms require to be equipped with complicated conveyor systems and need a large investment in facilities, hardware and software. On the other hand, stop-and-action platforms carry sensors to individual plants and take measurements one by one (such as GROWSCREEN FLUORO and Phenovator). The stop-and-action platforms are mostly used for the applications of fluorescence imaging, as fluorescence cameras usually require the dark adaptation for plants (Jansen et al., 2009). However, stop-and-action platforms have relatively low efficiency in data acquisition, for example, the Phenovator (Flood et al., 2016) took around 1 h to screen an area of fewer than 5 m^2 with 1440 Arabidopsis plants. Besides, commercial stop-and-action platforms are high-cost and often require significant modifications to the structure of greenhouse facilities (Jianfeng Zhou et al., 2018).

A low-cost and efficient alternative is to take images of multiple plants at one action. An et al. (2016) developed a high-throughput imaging system that is able to measure plant traits such as leaf length and area of more than 1000 rosette plants using 108 cameras uniformly distributed on the top of plants. Jianfeng Zhou et al. (2018) developed a low-cost phenotyping platform to quantify soybean dynamic responses to salinity stress in a greenhouse by estimating leaf chlorophyll and plant height. Images of around 400 soybeans were continuously taken by moving a single camera along a preset path, and an orthomosaic including all soybean plants was developed by stitching the sequential images using the Structure from Motion (SfM) method. In these platforms, images of hundreds and even thousands of plants can be acquired in a short time period.

Plant geometric features are essential for understanding plant growth and plant responses to biotic and abiotic stresses. Plant height has been adopted to correlate with wheat biomass (Schirrmann et al., 2016) and corn yield (Geipel et al., 2014), estimate crop growth rate (Holman et al., 2016) and predict stress tolerance (Jianfeng Zhou et al., 2018). Some other plant geometric features including plant orientation, leaf angle, stem diameter, and plant height that were extracted from 3D models have shown high correlations with the measurements using conventional methods (Bao et al., 2019; Holman et al., 2016).

The SfM is one of the popular methods to build 3D models of objects, including plants, using 2D images with different view angles (Zhou et al., 2018). The 2D images can be collected using high-throughput platforms where sensors (cameras) move continuously along a certain pattern above plants growing in a fixed seedbed. The 3D point cloud generated by SfM contains position (coordinate x, y, and z) and color (red, green and blue) information that can be used to extract geometric features of plants following a systematic data processing pipeline (Jianfeng Zhou et al., 2018). One of the important tasks in the data processing pipeline is to segment individual plants from a 3D point cloud of a large number of plants. However, manual segmentation could be time-consuming and labor-intensive given a large number of plants and the high frequency of collection (i.e. plants are collected daily through their lifespan) in breeding programs. For example, Jianfeng Zhou et al. (2018) collected images of 405 plants for 5 days, leading to 2025 soybean plants in total to be manually segmented. Moreover, due to the lack of clearly discernible boundaries among overlapped plants, the segmentation is exceptionally difficult (Scharr et al., 2016). An automated and accurate segmentation method has become one of the bottlenecks that prevent image-based phenotyping from being widely adopted by breeding programs, although high-quality imaging sensors are now available at low cost (Walter et al., 2012).

Machine learning is a promising approach as its strong ability in the extraction of details from images and successful applications in the plant leaf segmentation from 2D images. Histogram thresholding (Pape and Klukas, 2014), unsupervised machine learning (Valliammal and Geethalakshmi, 2012) and supervised machine learning (Yin et al., 2014) have been extensively used in leaf segmentation. Those methods also have the potential in segmenting whole plants instead of leaves. Minervini et al. (2017) presented a system, *Phenotiki*, allowing automated plant segmentation and leaf count using machine learning methods. To the best of our knowledge, there is no report on plant segmentation using 3D point cloud in leafy plants like soybean. Therefore, the goal of this study was to develop a data processing pipeline for the segmentation of individual soybean plants at early growth stages from 3D point cloud generated from 2D images that were collected using an image-based high-throughput phenotyping platform in a greenhouse. The detailed objectives were: (a) to evaluate the accuracy of three machine learning methods (*boosting*, Support Vector Machine and *K-means* clustering) on the segmentation of soybean plants from 3D point cloud data, (b) to assess the performance of the Histogram of Oriented Gradients (*HOG*) descriptor on the segmentation of overlapped plants, and (3) to evaluate the performance of the developed plant segmentation pipeline using a case study.

2. Methods and materials

2.1. Experimental setup

A wooden tank was built as a plant test bed in a greenhouse at the University of Missouri-Columbia, USA. A group of 75 soybean plants were placed on five trays with 15 plants per tray (five rows by three columns) at 10 cm spacing between two adjacent plants in both between- and in-row directions. A black fiber sheet was used to cover the soil in plant pots, water solution and plant trays, which allows forming a strong color contrast to soybeans and makes it easy to remove background. The layout of the 75 plants is shown in Fig. 1. An image-based high-throughput phenotyping platform developed in the previous study (Jianfeng Zhou et al., 2018) was used to collect images for this study. The platform includes a digital camera (SX410, Canon U.S.A., Melville, NY) with a resolution of 20.0 Megapixel (5152 × 3864 pixels) that was mounted on a camera holder to make the camera facing the target objects at the nadir view. The camera was configured to take one image at a frequency of every 3 ± 0.2 s using an open source firmware Canon Hack Development Kit (CHDK, http://chdk.wikia.com). The plants were screened at 9 am on the 4 (non-overlapped stage) and 14 days (overlapped stage) after emergence (DAE). The plants were scanned with 80% side overlap, 90% forward overlap and a ground spatial resolution of 2.78 pixel·mm^{-1} was obtained.

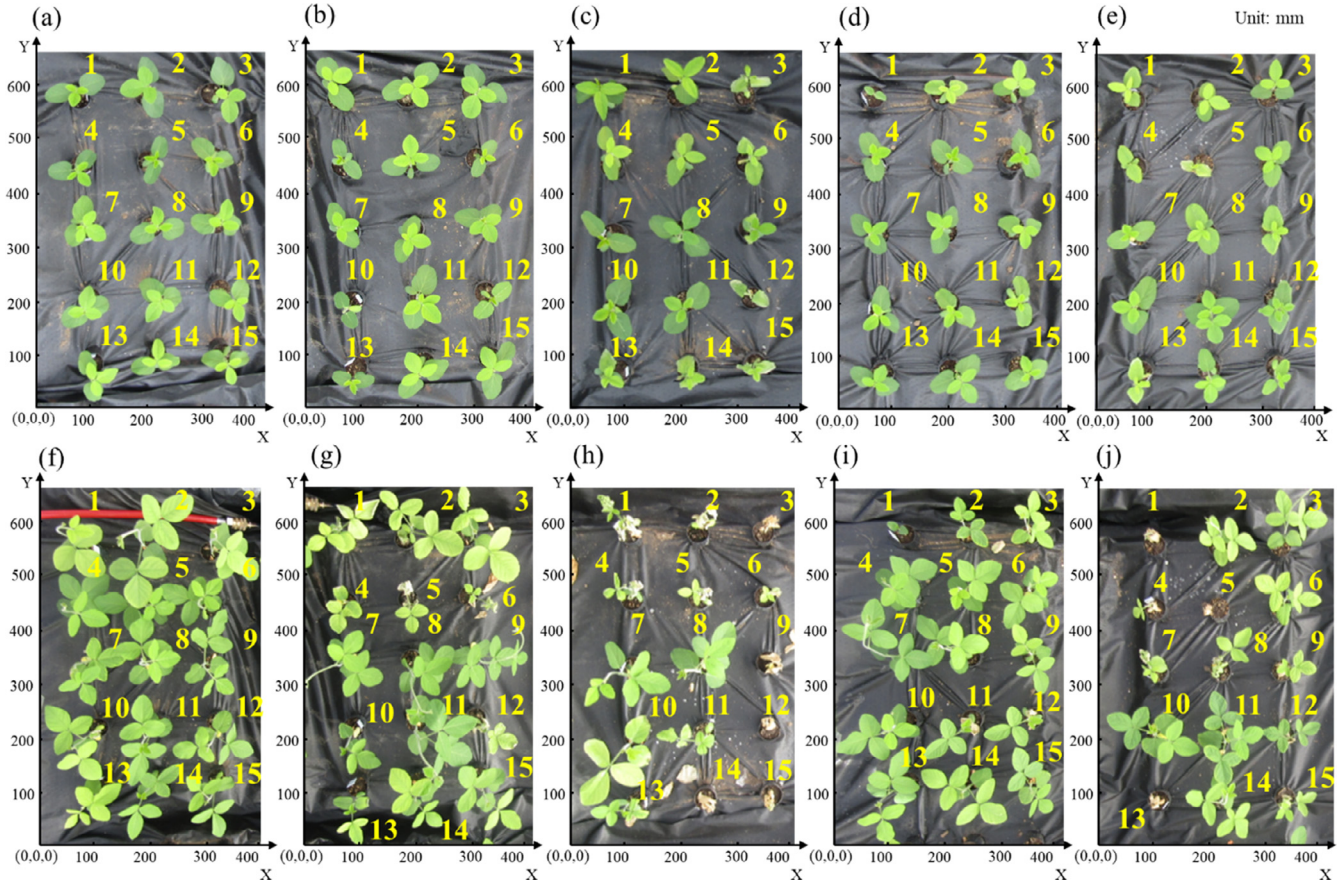

Fig. 1. The layout of the plant trays. (a) - (e) are No. 1 to 5 plant trays screened at the 4 DAE. (f) - (j) are No. 1 to 5 plant trays screened at the 14 DAE.

2.2. Dataset preparation

2.2.1. 3D point cloud development

3D point cloud data were built using the sequential images based on the SfM method, which is a low-cost photogrammetric method for 3D structure reconstruction from a series of multiple overlapping images (Westoby et al., 2012). It applies a highly redundant, iterative bundle adjustment procedure, based on a database of features automatically extracted from the set of multiple overlapping images to resolve the target's structure (Snavely et al., 2008). The method has been integrated by a number of image-processing software for developing 3D point cloud data, such as Agisoft PhotoScan Pro (v1.3.4, St. Petersburg, Russia) or Pix4D (v4.2.27, Lausanne, Switzerland), both of which can develop orthomosaic images and 3D point cloud data using user-uploaded and crowd-sourced images (Westoby et al., 2012). Recently, there has been a growing interest in using these tools in the field of agriculture, forestry, geoscience, archaeology, and architecture (Özyeşil et al., 2017).

In this study, sequential images were processed using Agisoft PhotoScan Pro to develop orthomosaic images and 3D point cloud data on a desktop PC (Dell Optiplex 5050). The PC was configured as Intel (R) Core i7-7700 CPU (8 cores), 16 GB RAM memory, a 512 GB solid state drive. The protocol of building 3D point cloud data involves three major stages: (1) importing sequential images and a geo-reference file, (2) aligning images and adding ground control points (GCPs), and (3) generating dense points. The parameters were set as "High" with Generic and Reference preselection for image alignment, "High" for reconstruction parameter and "Moderate" for filtering mode. The geo-reference file included location information of each image (camera), which was determined using the accumulated distance intervals from the origin in an established coordinate system (Jing Zhou et al., 2018). After the images were aligned, Agisoft software started searching similar features in the images to create 3D point cloud data that were used to calculate dimensions of target objects. The 3D point cloud data were exported to an M × 6 array (M is the number of points and six columns are three dimensions in three axes (*x*, *y*, and *z*), and three color channels (*Red* (*R*), *Green* (*G*) and *Blue* (*B*)) and saved as Comma Separated Value (*.csv*) files for further processing.

2.2.2. Manual segmentation of plants

Manual segmentation of plants was performed on the 3D point cloud data to give each point a true label that was used as the training set for supervised machine learning algorithms and ground truth data for the evaluation of segmentation accuracy using the machine learning methods. The 3D point cloud data were imported to CloudCompare (v2.9, www.cloudcompare.org), an open-source 3D point cloud processing software, to remove noises using the toolbox Statistical Outlier Removal (SOR) filter. The SOR filter computes the average distance of each point to its neighbors (considering *k* nearest neighbors for each) and rejects the points that are farther than a threshold (the average distance plus a number of times of the standard deviation) (CloudCompare, 2015). Background and 15 individual plants in each set were manually segmented and labeled from the cleaned point cloud, and then saved individually as *.csv* files, where all points were annotated to a specific label number (0 for background and 1–15 for plants, Fig. 1). The 16 files were combined into one file with each point having one label.

2.2.3. Features for segmentation

The position (coordinate *x*, *y*, and *z*), and color (*R*, *G*, and *B*) in RGB color space were directly obtained from the 3D point cloud data. The values of *hue* and *saturation* in HSV (*hue*, *saturation*, and *value*) color

space for each point were calculated as features for the segmentation. *hue* is a color resembling the pure color and is expressed as a number from 0 to 359, for example, red is from 0 to 59, green is from 120 to 179 and blue is from 240 to 299 (Smith, 1978). *saturation* is the variation of color depending on the lightness in the range of 0–100% (from the center of the black and white axis). The advantage of using HSV color space is that each of its attributes corresponds directly to the basic color concepts, which makes it conceptually simple in image processing (Shipman, 2012). The HSV was calculated using the *rgb2hsv* function in Image Processing Toolbox in MATLAB (ver. 2016b, The MathWorks, MA, USA). Meanwhile, the *Triangular Greenness Index* (*TGI*) (Hunt et al., 2011) that was suggested as an effective indicator of leaf chlorophyll content of plants was calculated using Eq. (1):

$$TGI = -0.5 \times [0.19(R - G) - 0.12(R - B)] \quad (1)$$

where *R*, *G*, and *B* are the pixel values of *Red*, *Green* and *Blue* channels in a 3D point cloud, respectively.

2.2.4. Calculation of HOG descriptors

The *HOG* is a feature descriptor commonly used in computer vision and image processing for object detection (Freeman and Roth, 1995). The *HOG* was introduced in this study to improve the accuracy in the segmentation of overlapped plants. The *HOG* descriptor was calculated using the *hog3D* function (Dupre et al., 2015) in MATLAB, which was designed based on the original 2D Histogram of Oriented Gradients (Dalal and Triggs, 2005). As the *hog3D* function requires its input to be an $m \times n \times p$ matrix defining voxels with values in the range of 0–1, the 3D point cloud data of each plant set were converted from *.csv* arrays to a logical matrix H with m, n and p equal to the range of coordinate x, y and z in millimeter, respectively, and the voxel $H(x, y, z)$ was defined as 1 if there was a point existing at the position in the 3D point cloud. The *HOG* descriptor of H was obtained with the parameters in Table 1.

With parameters in Table 1, the *hog3D* function separated the matrix H into k blocks using Eq. (2) and returned the position of each block and each cell in the block and the histograms of each cell (histogram in the *x-y* plane and histogram in the *x-z* plane). The magnitude and direction of the gradients of each cell in two planes were converted to two histograms, where the bins in each histogram represented the directions from 0 to 359 in the plane and the count of each bin represented the sum of magnitude (alpha) in the certain range of direction.

$$k = \text{floor}\left(\frac{\text{m}}{cell\_size \times step\_size}\right) \times \text{floor}\left(\frac{\text{n}}{cell\_size \times step\_size}\right) \times \text{floor}\left(\frac{\text{p}}{cell\_size \times step\_size}\right) \quad (2)$$

where cell_size and step_size were defined in Table 1.

The gradients of each cell were filtered using an alpha threshold to remove the values less than the threshold. The results of an example image filtered with different alpha thresholds are shown in Fig. 2 that indicates the higher the alpha threshold was, the more overlapping effects were removed. The alpha threshold was chosen to obtain the lowest error rate using 5-fold cross-validation.

The *HOG* descriptors of each overlapped plant set were first segmented using the *K-means* clustering (K = the number of plants in one classification) and labeled with the returned classes. The labeled *HOG* descriptors were used as the training sets to train the supervised learning algorithms for the segmentation of their original 3D point cloud data.

2.3. Machine learning methods

In this study, three machine learning methods were tested in the segmentation of individual plants, including *boosting*, Support Vector Machine (SVM) and *K-means* clustering. The *boosting* method is designed to improve the predicting accuracy based on the decision tree (James et al., 2013). The segmentation using '*boosting*' was performed using the ***adabag*** package (Alfaro et al., 2013) in the software *R* (Version 3.4.3, 2017, R Foundation). The '*importance*' function in this package was used to quantify the contributions of different features to the segmentation. The SVM method was successfully used in the tasks of two-class classification (Hsu and Lin, 2002), and was also effective in solving multi-class problems using the One-versus-One (OVO) method. The OVO method compares one of the K classes to the remaining K-1 classes each time (Schölkopf et al., 2002) and is verified more efficient than other algorithms regarding the processing time and accuracy (Hsu and Lin, 2002). The SVM method was performed using the '*svm*' function in *R* package for the segmentation (Meyer et al., 2015). The *K-means* clustering is a simple and widely used unsupervised machine learning method for partitioning a dataset into K clusters (James et al., 2013). The method makes no assumption between responses and observations. It interests in distinguishing similar relationships among observations, which is very close to the nature of the segmentation problem, i.e. discovering subset with similar features. The '*kmeans*' function in *R* was performed to conduct the segmentation.

The first step of segmentation was to remove the background (pixels that were not identified as plants) from the 3D point cloud data. Background removal was solved as a two-class classification problem by labeling background points as "0" (background) and non-zero points (plants) as the other class. The parameters of each machine learning method for background removal were set as the values listed in Table 2. In the *boosting* algorithm, the logical parameter *boos* that was used to call the bootstrap sampling procedure within the function was set to *FALSE*. The integer *mfinal* that is used to define the number of trees being computed in *boosting* algorithm was set as 50 considering of both the potential overfitting and processing efficiency (James et al., 2013). In SVM, besides applying linear boundary to the model, the 2nd polynomial boundary was also applied to address the potential non-linear relationship between color information of plant and background. The parameter *nstart* = 50 was determined in *K-means* clustering in this study, which is a typical value used in practice (James et al., 2013).

Plant separation was conducted using the data sets after removing the background that accounted for more than 80% of the points in one data set. The cleaned data set had fewer data points and was able to improve the data processing efficiency. The parameters for the plant separation were set as the values in the "Plant Separation" column in Table 2.

2.4. Performance assessment and data analysis

The performance of the background removal and plant separation

Table 1
Parameters used in *hog3D* function.

Parameters	Value	Description
cell_size	8	The matrix was divided into small connected regions called cells, and cell_size is the number of pixels in a cell
block_size	2	Groups of adjacent cells are considered as spatial regions called blocks, which are the basis for grouping and normalization of histograms
theta_histogram_bins	9	Number of bins in the x-y plane
phi_histogram_bins	18	Number of bins in the x-z plane
step_size	2	Number of cells overlapped by blocks

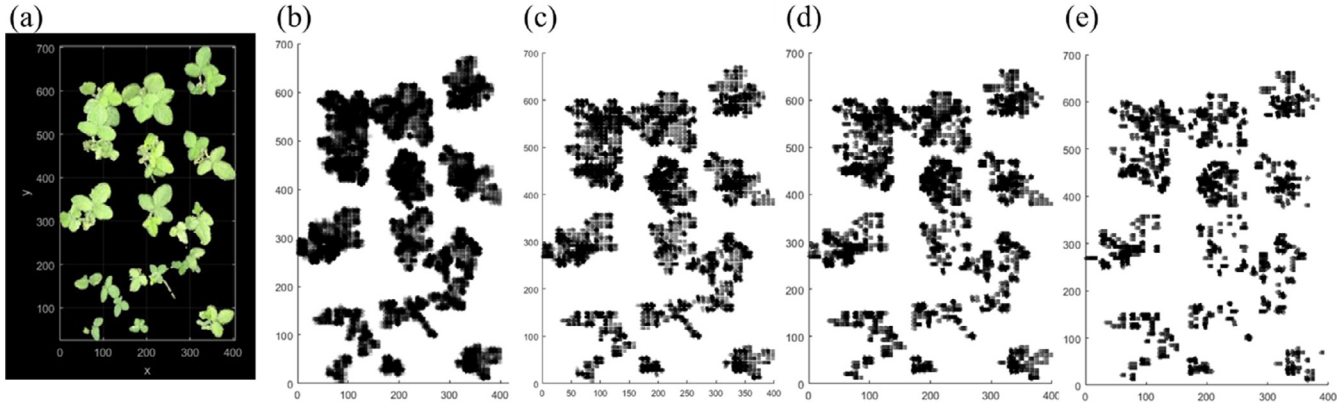

Fig. 2. The *HOG* descriptors with different alpha thresholds. (a) 3D point cloud of soybean plants at 14 DAE. (b) to (e) are gradient images with the alpha threshold = 0.1, 0.2, 0.3 and 0.4, respectively.

Table 2
Key parameters of three machine learning methods for background removal and plant separation.

Method	Parameters	Background Removal	Plant Separation
boosting	*coeflearn*	"Zhu" (SAMME)	"Freund" (AdaBoost.M1)
	boos	FALSE	FALSE
	mfinal	50	50
SVM	*boundary*	Linear The 2nd polynomial	Linear
	Class type	–	One versus one
K-means clustering	*nstart*	50	50

using different machine learning methods was assessed using error rates of the segmentation. The error rate is the percentage of the number of mismatches between the predicted labels (automated segmentation) and the true labels (manual segmentation) of points over the total number of points. All the machine learning methods for segmentation were performed using *R*. Packages and functions mentioned above can be referred from manuals on CRAN (The Comprehensive *R* Archive Network) website (cran.r-project.org). The variation in the error rates of each machine learning method was assessed using one-way ANOVA with default settings to compare the difference in means at a 0.05 level of significance in *R*.

2.5. Evaluation of the automated plant segmentation pipeline

To evaluate the performance of plant segmentation pipeline, the developed algorithms were used to segment plants at different growth stages in another experiment. Five soybean varieties (127, WM82, 87, 62, Lee) with 14 replicates for each variety were planted in salt solution to evaluate their performance under salt stress. Soybean seeds were sown in seedling tubes and filled with general purpose peat-based growing medium (PRO MIX, Premier Tech Horticulture, PA, USA). The soybeans were transferred to the wooden tank containing salt water solution (120 mM of NaCl) to induce salt stress after the emergence. The soybeans were placed on plant trays as 14 × 5 arrays with 15 cm space between two adjacent plants in both between- and in-row directions. The layout is shown in Fig. 3.

The screening strategy was the same as that introduced in Section 2.1. Plants were screened at five different days: 1, 6, 13, 17 and 23 (the last day) DAE. The flow chart in Fig. 4 shows the pipeline of the automated plant segmentation and the image feature extraction for all images collected through the plant's lifetime. After the 3D point cloud was built, the point cloud data were imported to MATLAB to remove noises and background using *kmeans* function based on *TGI*. The parameter *k* was set as 2 and the squared Euclidean distance between features was calculated during clustering. The function returned a vector labeling the class of each point ("1" stands for background and "2" stands for plants or vice versa). According to the principle of *K-means* clustering, *k* points from a dataset were randomly selected as initial clusters, and then the distance between each individual point and each cluster was calculated. The points with the shortest distance were assigned to the cluster. The procedure repeated among all points in the dataset (James et al., 2013). During the calculation, the labels representing two classes may switch to each other, and the labels with the higher average *TGI* were recognized as plants, and those with lower average *TGI* as background.

The cleaned data included only plants that were then segmented using *K-means* clustering as the *k* equals the number of plants (69 in this case). Since overlapped plants were not observed during the 23 days of the experiment, *K-means* clustering was always used for background removal and plant separation and there was no separation of overlapped plants included in this case study. Due to the randomness of the initial clusters, *K-means* clustering has poor repeatability in term of the results returned every time (James et al., 2013), resulting in switching labels within clusters as mentioned above. Therefore, a standard *K-means* clustering algorithm was modified by assigning the initial clusters with known points, instead of random ones. Before performing *K-means*, the initial clusters need to be selected manually according to the position of individual plants. The initial clusters do not require each point locating at the center of a plant exactly but any point in the plant. The positions of plants at the 1 DAE were manually selected one by one to record the initial clusters for separation and order the labels for the storage of plant features after the separation. This set of initial clusters at the 1 DAE can be applied to the separation of the dataset at any other days. The initial clusters were then loaded as the input of the modified *K-means* clustering and points with the shortest distance between the initial ones were assigned to these clusters. After separation, features of individual plants (i.e. plant height, color and shoot area) were extracted and saved for further analysis.

3. Results and discussion

3.1. Error rates in background removal

The error rates in background removal using three methods are shown in Fig. 5(a). The error rates varied from 0.36% (*K-means* clustering) to 33.26% (SVM with the 2nd polynomial kernel). ANOVA test indicates that there was no significant difference among the average error rates of 1.67%, 4.80%, and 6.32% obtained using methods of *K-means* clustering, SVM with the linear kernel and *boosting* respectively.

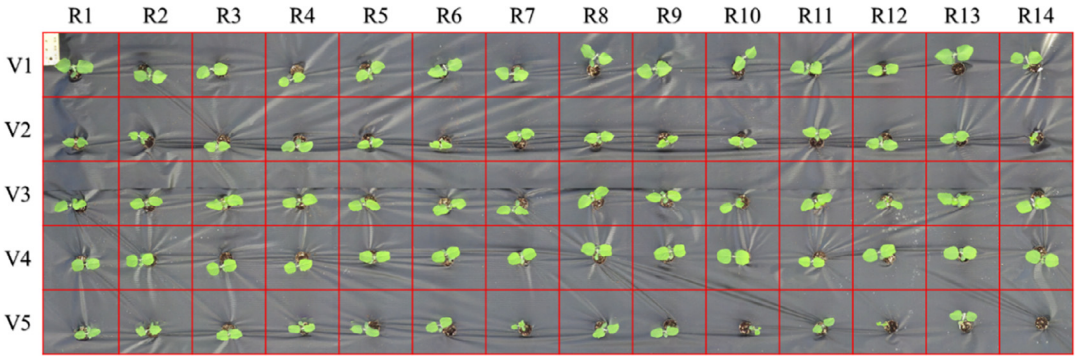

Fig. 3. The layout of plants in the salt-stress study. R1-R14 represents 14 replicates of one variety (R = replicate) and V1-V5 represents five varieties in this case (V = variety), which were 127, WM82, 87, 62, Lee. One plant in the right-bottom corner (V5-R14) died during the transfer.

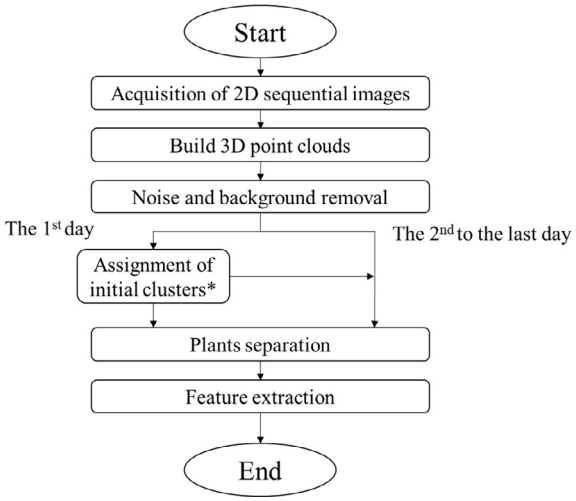

Fig. 4. The pipeline of the automated segmentation and feature extraction from individual plants. The asterisk marks the step requiring manual operations.

However, the error rate in the SVM algorithm with the 2nd polynomial kernel was significant higher (19.37%) than those from other algorithms (*p-value* = 0.015 at a 5% significance level), indicating that there was no the 2nd polynomial relationship between the features of plants and background.

K-means clustering not only had the smallest range of error rates but also took the least processing time in processing data (3–5 min for a scanning area of 0.2 m^2 compared to 30 min using *boosting* and more than one hour using SVM methods for the same area). In addition, there was only one step in segmentation by finding patterns among variables using *K-means* clustering, however, the segmentation procedure using other supervised methods depends on the trained models and may need additional plant sets, which is the most time-consuming step. Meanwhile, *K-means* clustering was able to identify the noises associated with soybean leaves by dismissing the effects of height (position coordinate *z*). The noises were caused by the errors in feature recognition when the 3D point cloud data were developed. A small proportion of the background points whose color was close to plant leaves had a higher potential to be classified as a part of leaves (3.1% for non-overlapped plants and 13% for overlapped plants on average). Due to

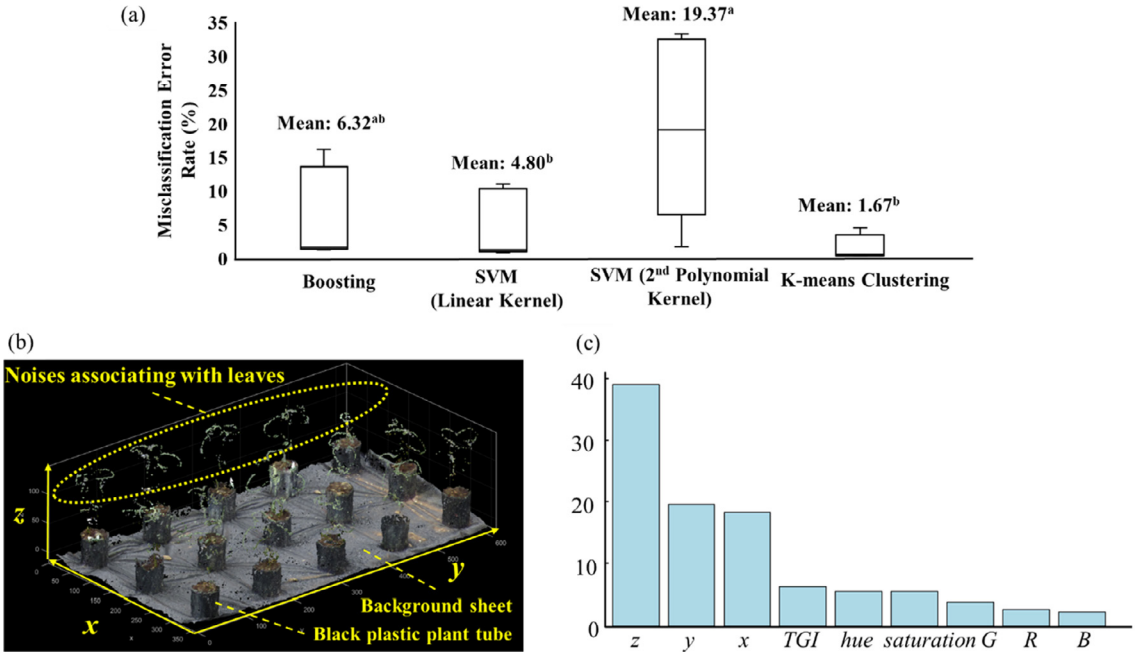

Fig. 5. Performance of background removal using three machine learning methods. (a) Boxplot of error rates in background removal. The different lower-case letters in superscripts (a, b and c) indicate the significant difference in means of misclassification error rate obtained using different methods at a 5% significance level. (b) Background removed using *K-means* clustering. There were rings of noise around leaves as shown in the dashed oval. (c) The relative importance of components contributing to background removal using *boosting* method.

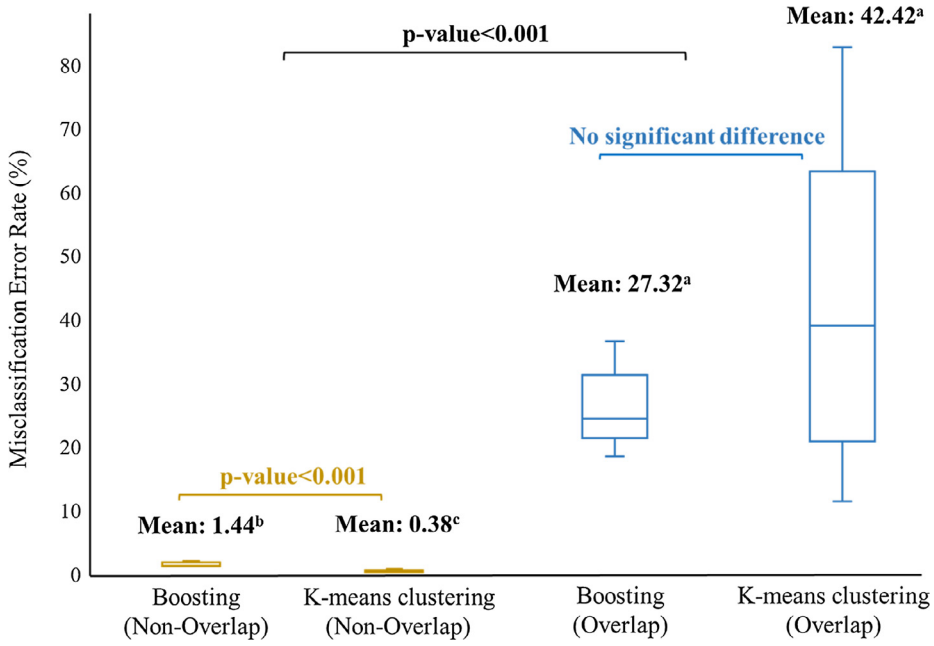

Fig. 6. ANOVA results of error rates in plants separation using different machine learning methods. The different lower-case letters in superscript (a, b and c) indicate the significant difference in means of misclassification error rates using different methods at a 5% significance level.

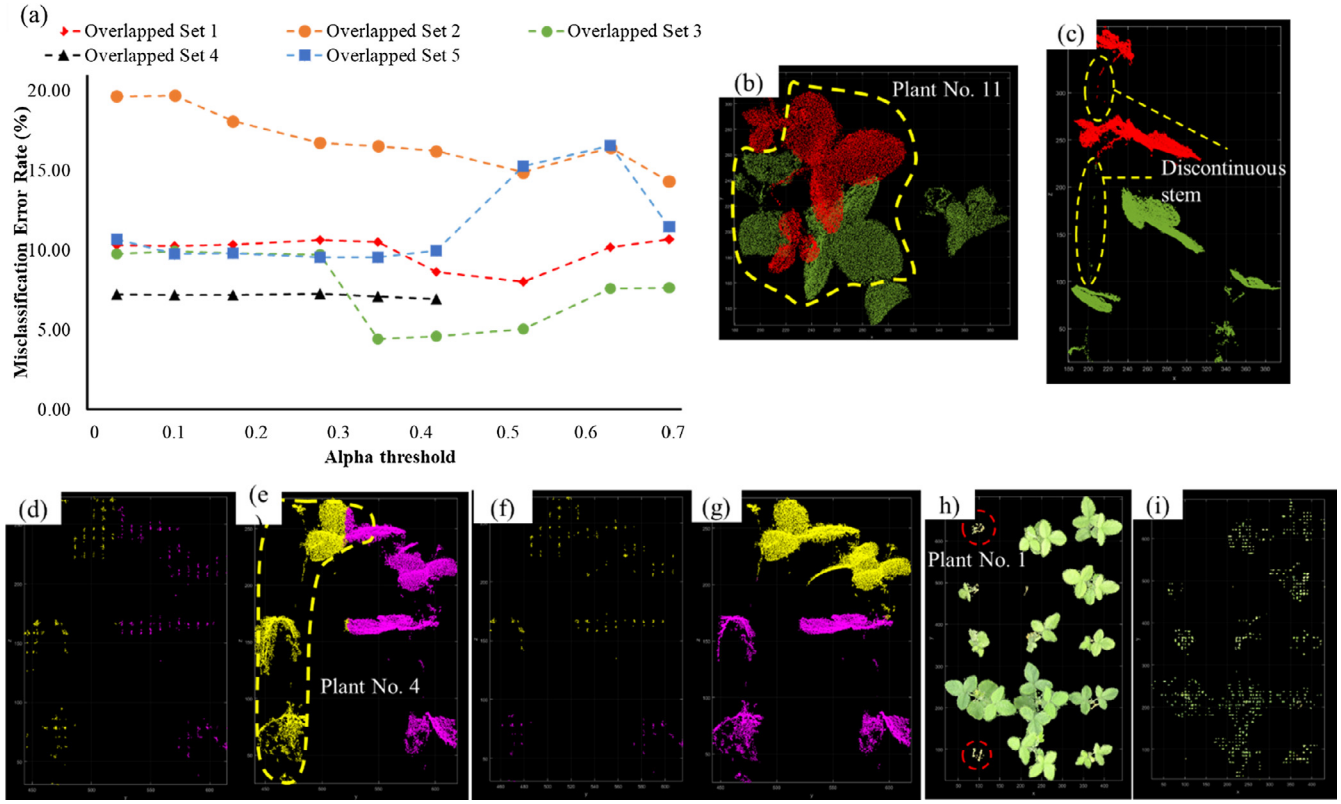

Fig. 7. The results of cross-validation for the alpha threshold of the *HOG* descriptors. (a) The variation in the error rates in different test sets of overlapped plant regarding different alpha thresholds. (b) Top view and (c) front view of the plant No. 11 in overlapped Set 2 shows that the long and thin leaf stem leaning on the neighboring plant, resulting in a discontinuity in the point cloud. (d) The *HOG* descriptor with alpha threshold = 0.5 and (e) the corresponding predicted result with alpha threshold = 0.5 of plant No. 4 in overlapped Set 5. (f) The *HOG* descriptor with alpha threshold = 0.4 and (g) the corresponding predicted result with alpha threshold = 0.4 of plant No. 4 in overlapped Set 5. Two colors in (e) and (g) stands for the *HOG* descriptors of two plants. (h) The top view of the 3D dense point cloud and (i) the *HOG* descriptors of plant No. 1 and 13 in overlapped Set 5.

the short spatial distance between these noises and leaves edge, they were not removed by the SOR filter, resulting in the outer rings around leaves (Fig. 5(b)). The '*importance*' of variables in *boosting* method (Fig. 5(c)) shows that the coordinate *z* was of the most important variable in background segmentation, leading to the noises associated with leaves segmented incorrectly when using other supervised

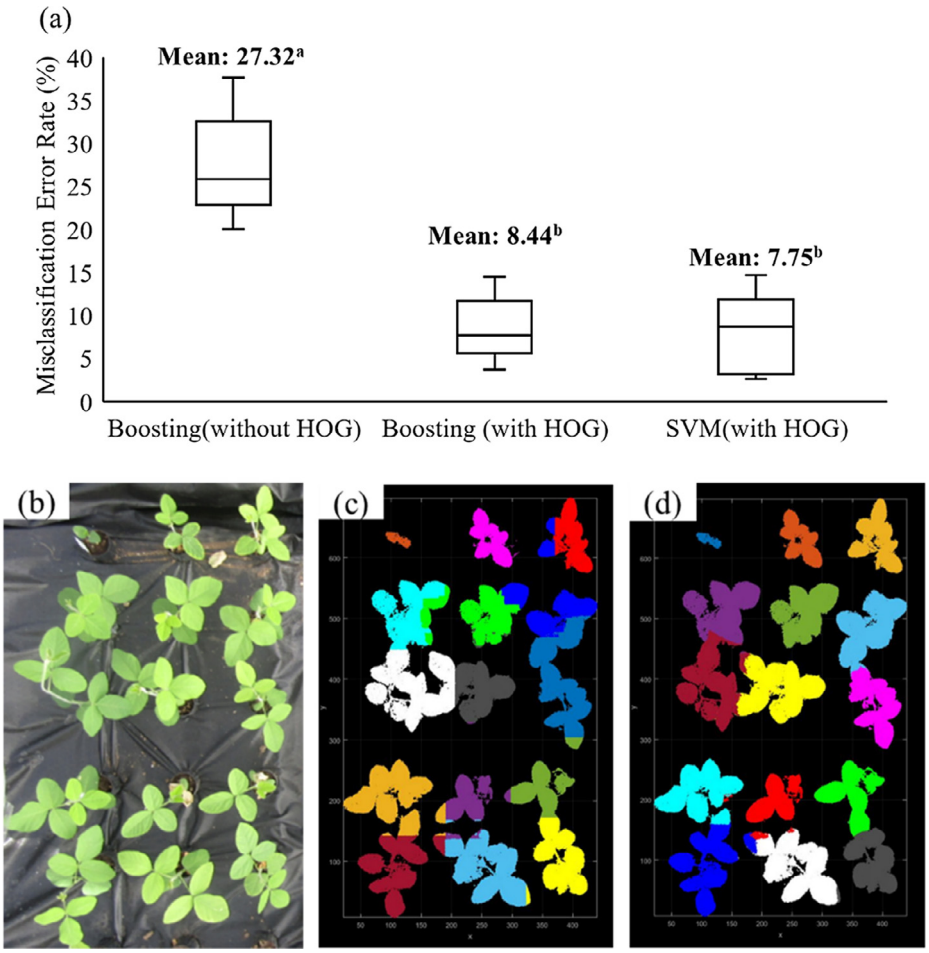

Fig. 8. The performance of plant separation for the overlapped plants using the *HOG* descriptor. (a) Comparison of the error rates with and without the *HOG* descriptor in the separation of overlapped plant sets with the alpha threshold = 0.4. The different lower-case letters in superscripts (a, b and c) indicate the significant difference in means of different methods at a 5% significance level. (b) RGB image of set 4. (c) Overlapped plants separated using *boosting* without the *HOG* descriptor. (d) Overlapped plants separated using *boosting* with the *HOG* descriptor.

machine learning and manual segmentation methods.

3.2. Error rates in plant separation without HOG descriptor

The error rates in individual plant separation using *boosting* method and *K-means* clustering for the non-overlapped and overlapped plants are shown in Fig. 6. The error rates varied from 0.20% to 1.96% for the non-overlapped plants and from 13.17% to 82.60% for the overlapped plants. ANOVA test indicates that *K-means* clustering had significantly lower error rates than *boosting* method in the non-overlapped plant set, however, there was no significant difference between the error rates in both methods in the overlapped plant set. In general, the error rates in the overlapped plant set were significantly higher than those in the non-overlapped plant set. The SVM was not performed on plant segmentation without *HOG* descriptor because it took over five hours in learning procedures and predicting results. Thus, for the non-overlapped plant segmentation, *K-means* clustering was considered as the best one among the three methods. The large variation of error rates for overlapped plants using *K-means* clustering was due to the extended leaves. When one or more leaves of a plant extended and touched or even covered their neighbors, they were easy to be classified into a wrong class. This is one of the drawbacks of the unsupervised machine learning methods, which make decisions based on whatever makes the overall variation the least. Therefore, efforts in this paper were made to reduce the effects of overlapped leaves.

3.3. Error rates in overlapped plant separation with HOG descriptor

3.3.1. Cross-validation for the alpha threshold of HOG descriptor

The results of cross-validation for tuning the alpha threshold of the *HOG* descriptor are shown in Fig. 7(a). The reason for the higher error rate in the overlapped plant set 2 than that in others is probably because of the thin and light green stems of tall plants in this set that caused fewer data points for the stems, which made it challenging to estimate plant dimensions accurately using 3D point cloud due to. Fig. 7(b) and (c) illustrated plant No. 11 in the set 2 whose stem was discontinuous in the 3D point cloud, and the top leaf lodged toward its neighbors, resulting in the main part of this leaf was misclassified. Fig. 7(d) to (g) shows another example of the overlapped plants in which the error rates increased sharply when the alpha threshold increased from 0.4 to 0.5. Similarly, the stem of No. 11 in the set 2 was filtered out when the alpha threshold was 0.5, resulting in the misclassification of *HOG* descriptors. The whole plant No. 1 and No. 13 in the overlapped plant set 5 were almost filtered out when the alpha threshold was larger than 0.4 (Fig. 7(h) and (i)), thus the *HOG* descriptors were failed to separate the plant set into 15 groups. Therefore, the alpha threshold = 0.4 was used in this study to separate the overlapped plants.

3.3.2. Error rates in the separation of the overlapped plant using the HOG descriptor

The error rates in the separation of the overlapped plant using the

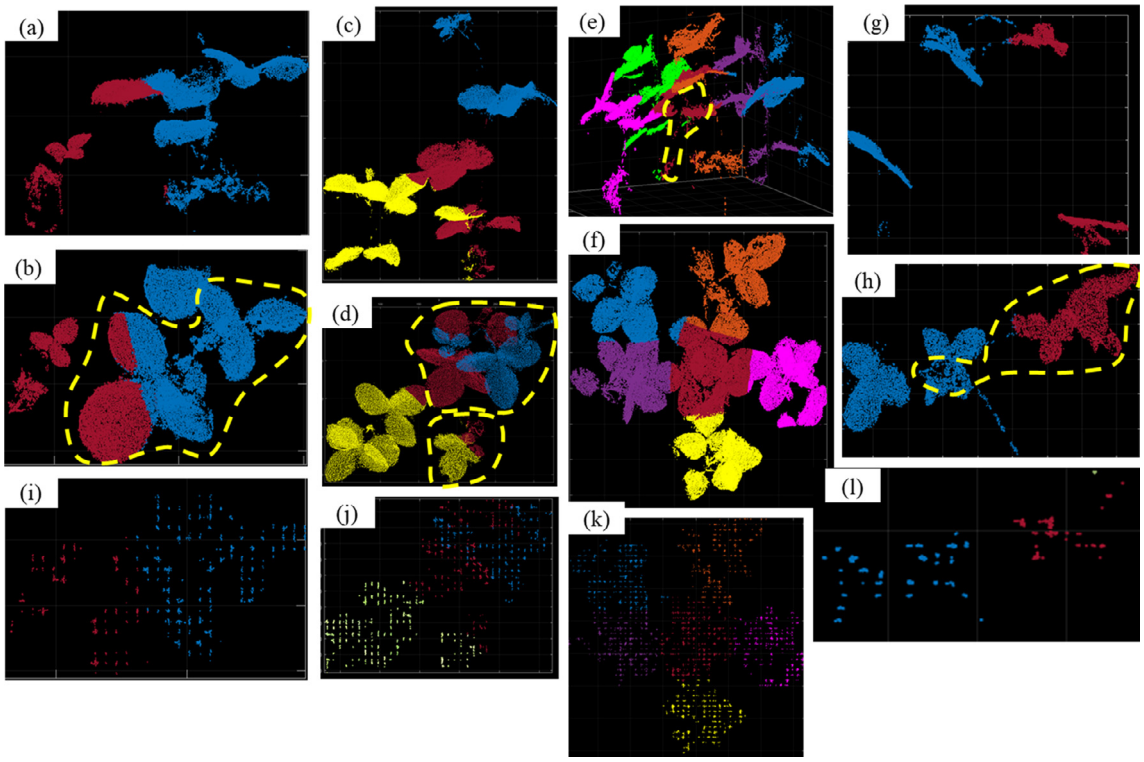

Fig. 9. Illustration of the overlapped plants with large error rates. (a) and (b) show the front view and top view of No. 3 in Set 2 with 34.4% error rate. (c) and (d) show the front view and top view of No. 11 in Set 2 with 56.5% and No. 12 with 74.87% error rate. (e) and (f) show the front view and top view of No. 5 in Set 1 with 49.5% error rate. (g) and (h) show the front view and top view of No. 12 in Set 5 with 58.3% error rate. (i) to (k) show the *HOG* descriptors of No. 3 in Set 2, No. 11 and No. 12 in Set 2, No. 5 in Set 1, and No. 12 in Set 5, respectively.

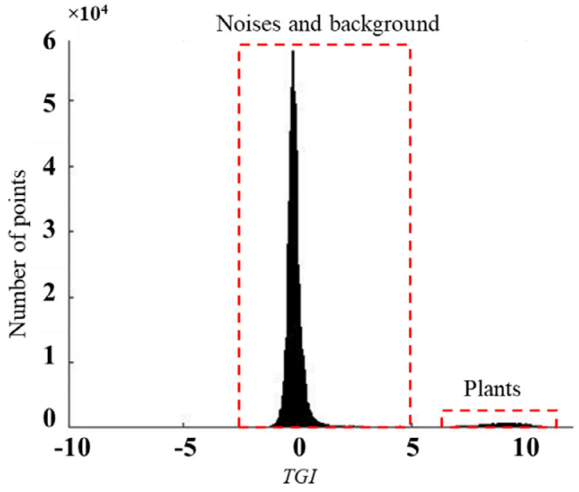

Fig. 10. Histogram of TGI for the 3D dense point cloud.

HOG descriptor as the training set (alpha threshold = 0.4) are shown in Fig. 8. Compared to the *boosting* method without a *HOG* descriptor, the error rates of the *boosting* and SVM methods with the *HOG* descriptor were significantly decreased (*p-value* < 0.001 at a 5% significance level). The mean error rates of the overlapped plant set (No.1 to 5) were from 2.57% to 14.65% using SVM and from 3.71% to 14.48% using the *boosting*, while the standard deviation (STD) of error rates of overlapped plant set varied from 5.13% to 23.16% using SVM and from 5.11% to 23.18% using *boosting*. The fluctuation of STD was due to various plant gestures in individual plant sets. Some plants were tall and lodging to the neighboring plants, while some plants were relatively small and short and did not touch their neighbors.

Fig. 8(b)-(d) shows the performance of the plant separation in the overlapped plant set 4 that was separated using the *boosting* method without (c) and with (d) the *HOG* descriptor. It can be seen that the error rate decreased significantly using the *HOG* descriptor due to the following reasons. First, plants that were not overlapped by their neighbors were separated correctly using the *HOG* descriptor, for example, plant No.3. Second, the mis-separated part of the overlapped plants (No. 6 and 8) were much smaller in Fig. 8(d) compared to those in Fig. 8(c), indicating that the *HOG* descriptor was able to remove overlapping effects to some extent even there were errors remaining in overlapped plants.

The overlapped plants with the error rate over 30% are shown in Fig. 9(a) - (b), and their *HOG* descriptors are shown in (e) - (f). In (a), an expanded leaf of plant No. 3 overlapped with its small-size neighbor without touching each other. As discussed in Section 3.3, tall plants, for example, No. 11 in (b), with discontinuous stem and lodging leaves always led to a high error rate. It is also demonstrated in (d) that a top leaf went far from the main part of this plant. In (c), the whole canopy of No. 5 plant was almost covered by its neighbors. From (e) - (f), the *HOG* descriptors of those plants had been in the wrong classes leading to further errors when predicting using supervised methods. Therefore, further improvement in the classification accuracy of overlapped plants is either to reduce the error rate of *HOG* descriptor using *K-means* clustering or to avoid the incomplete shape of leaves when predicting using supervised methods.

3.4. Case study

The histogram of *TGI* for the 3D point cloud built from 2D sequential images is shown in Fig. 10. The difference between background (noises and non-plant materials in the images) and the plants were successfully separated using the *K-means* clustering algorithm, indicating the effectiveness of the developed method.

The height and shoot area of plant V2-R5, V1-R6 and V1-R13

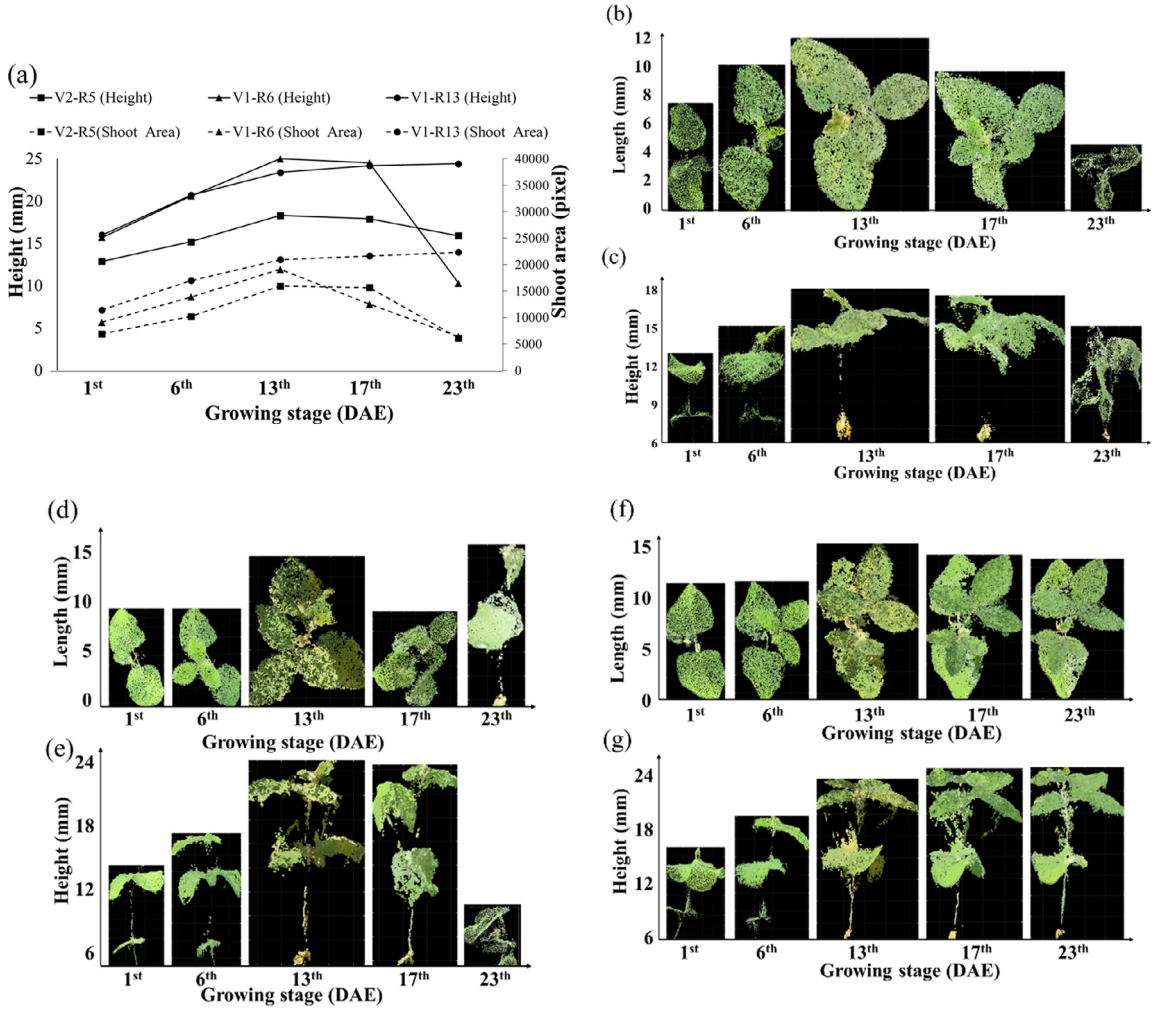

Fig. 11. Plant traits automatically acquired using the developed image processing pipeline and high-throughput phenotyping system. (a) Height and shoot area of plant V2-R5, V1-R6, and V1-R13 through their lifespan. (b) and (c) are the top and front view of the plant V2-R5. (d) and (e) are the top and front view of the plant V1-R6. (f) and (g) are the top and front view of the plant V1-R13.

extracted from the 3D point cloud through their lifespan are shown in Fig. 11. The curves in Fig. 11(a) was able to not only track the growth (height and shoot area) of individual plants but also quantify their difference that might be used to determine their level of resistance to salt stress. The segmentation method proposed here is automated except the adjustment step to ensure higher accuracy and takes three min on average to process 69 plants in this case, which makes it efficient and accurate to obtain useful information of plants in different genotypes using the high-throughput phenotyping platform.

The proposed methods for background removal and plant separation are potential in segmenting soybeans in field conditions. In existing studies, background of digital soybean images was removed by excluding pixels outside of a predefined color threshold. For example, Naik et al. (2017) combined saturation and hue thresholds to identify variances in soybeans and soil in 148 images. Similarly, Bai et al. (2018) observed that hue values of healthy and stressed (iron deficiency chlorosis) soybeans were from 40 to 170°, and those outside of this range were excluded. In these cases, the threshold for segmentation might be automatically achieved by applying *K-means* clustering on HSV images. For separating soybean plants from orthomosaic images with a large number of soybean lines, for example, in the case of Yu et al. (2016), *K-means* clustering based on position might be efficient for non-overlapped lines and SVM with the *HOG* descriptor has the potential in separating soybean lines with bulky canopies and connected to their neighbors.

4. Conclusions and future study

This study evaluated the performance of an automated segmentation method for soybean plants at early growth stages from 3D point cloud data built from 2D images collected in a greenhouse. Three machine learning methods, i.e. *boosting*, Support Vector Machine and *K-means* clustering were tested. Ten image features including position (coordinates *x*, *y* and *z*), colors (*Red*, *Green*, and *Blue* in RGB color space, *hue* and *saturation* in HSV color space and *TGI* greenness) and the *HOG* descriptor were used to segment background and individual plants. 2D sequential images of soybeans at two different growing stages (Non-overlapped and overlapped) were reconstructed to build 3D point cloud data using the Structure from Motion method. It was found that *K-means* clustering had the least mean error rates (0.36% and 0.20%) for background removal and non-overlapped plants separation. The least error rate in overlapped plants separation was decreased to 2.57% using SVM with the *HOG* descriptor. *K-means* clustering outperformed those methods in the aspect of processing efficiency and segmentation accuracy. In real experimental scenarios, efforts may be always made to avoid overlapped plants, but it happened sometimes due to the limitation of the experiment site or unforeseen overgrowth. Under this situation, the *HOG* descriptor is a potential method to segment overlapped plants. The developed method was validated using a case study, where 69 soybeans were planted under salt stress. Individual plants were automated separated and plant features (height and shoot area) of

each plant were successfully extracted using the proposed automatically segmentation pipeline. The pipeline took three minutes on average for each day's extraction of image features that were able to track the growth changes of all plants. The limitation of this study is that methods proposed in the study are developed and evaluated for soybeans at vegetative stages under salt stress, where the dimension of plants is relatively smaller than those at reproductive stages and under controlled treatments. However, the application of using *HOG* descriptor to reduce overlapping effects between plants could offer a reference for plants in the greenhouse with certain overlapping or for plant plots in the field with bulky canopy shapes at later stages or by weeds between plots. In future work, the classification accuracy of the *HOG* descriptor needs to be further improved, and the incomplete shape of leaf needs to be detected to get better segmentation performance.

Acknowledgment

We would like to thank colleagues in Molecular Genetics and Soybean Genomics Laboratory at the University of Missouri for their kindly help in plant material preparation and management. We also would like to thank colleagues Chin Nee Vong and Aijing Feng from Precision and Automated Agriculture Laboratory at the University of Missouri for their kind help in conducting experiments.

References

Alfaro, E., Gamez, M., Garcia, N., 2013. Adabag: An R package for classification with boosting and bagging. J. Stat. Softw. 54 (2), 1–35.

An, N., Palmer, C.M., Baker, R.L., Markelz, R.C., Ta, J., Covington, M.F., Maloof, J.N., Welch, S.M., Weinig, C., 2016. Plant high-throughput phenotyping using photogrammetry and imaging techniques to measure leaf length and rosette area. Comput. Electron. Agric. 127, 376–394.

Bai, G., Jenkins, S., Yuan, W., Graef, G.L., Ge, Y., 2018. Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning. frontiers. Plant Sci. 9 (1002). https://doi.org/10.3389/fpls.2018.01002.

Bao, Y., Tang, L., Srinivasan, S., Schnable, P.S., 2019. Field-based architectural traits characterisation of maize plant using time-of-flight 3D imaging. Biosyst. Eng. 178, 86–101. https://doi.org/10.1016/j.biosystemseng.2018.11.005.

CloudCompare, 2015. SOR filter. Retrieved from https://www.cloudcompare.org/doc/wiki/index.php?title=SOR_filter (10/6/2015).

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. Paper presented at the Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference.

Deshmukh, R., Sonah, H., Patil, G., Chen, W., Prince, S., Mutava, R., Vuong, T., Valliyodan, B., Nguyen, H.T., 2014. Integrating omic approaches for abiotic stress tolerance in soybean. Front. Plant Sci. 5, 244.

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. https://doi.org/10.1071/FP12018.

Dupre, R., Argyriou, V., Greenhill, D., Tzimiropoulos, G., 2015. A 3D scene analysis framework and descriptors for risk evaluation. Paper presented at the 3D Vision (3DV), 2015 International Conference.

Fiorani, F., Schurr, U., 2013. Future scenarios for plant phenotyping. Annu. Rev. Plant Biol. 64, 267–291.

Flood, P.J., Kruijer, W., Schnabel, S.K., van der Schoor, R., Jalink, H., Snel, J.F.H., Harbinson, J., Aarts, M.G.M., 2016. Phenomics for photosynthesis, growth and reflectance in Arabidopsis thaliana reveals circadian and long-term fluctuations in heritability. Plant Methods 12 (1), 14. https://doi.org/10.1186/s13007-016-0113-y.

Freeman, W.T., Roth, M., 1995. Orientation histograms for hand gesture recognition. Paper Presented at the International Workshop on Automatic Face and Gesture Recognition.

Geipel, J., Link, J., Claupein, W., 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. Remote Sens. 6 (11), 10335.

Holman, F.H., Riche, A.B., Michalski, A., Castle, M., Wooster, M.J., Hawkesford, M.J., 2016. High throughput field phenotyping of wheat plant height and growth rate in field plot trials using UAV based remote sensing. Remote Sens. 8 (12), 1031.

Hsu, C.-W., Lin, C.-J., 2002. A comparison of methods for multiclass support vector machines. IEEE Trans. Neural Networks 13 (2), 415–425.

Hunt, E.R., Daughtry, C., Eitel, J.U., Long, D.S., 2011. Remote sensing leaf chlorophyll content using a visible band index. Agron. J. 103 (4), 1090–1099.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning, vol. 112 Springer.

Jansen, M., Gilmer, F., Biskup, B., Nagel, K.A., Rascher, U., Fischbach, A., Briem, S., Dreissen, G., Tittmann, S., Braun, S., De Jaeger, I., Metzlaff, 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 (11), 902–914.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C., Lin, C., 2015. Misc functions of the department of statistics, probability theory group (formerly: E1071). Package e1071. TU Wien.

Minervini, M., Giuffrida, M.V., Perata, P., Tsaftaris, S.A., 2017. Phenotiki: an open software and hardware platform for affordable and easy image-based phenotyping of rosette-shaped plants. Plant J. 90 (1), 204–216. https://doi.org/10.1111/tpj.13472.

Naik, H.S., Zhang, J., Lofquist, A., Assefa, T., Sarkar, S., Ackerman, D., Singh, A., Singh, A.K., Ganapathysubramanian, B., 2017. A real-time phenotyping framework using machine learning for plant stress severity rating in soybean. Plant Methods 13 (1), 23. https://doi.org/10.1186/s13007-017-0173-7.

Özyeşil, O., Voroninski, V., Basri, R., Singer, A., 2017. A survey of structure from motion*. Acta Numerica 26, 305–364.

Pandey, P., Ge, Y., Stoerger, V., Schnable, J.C., 2017. High throughput in vivo analysis of plant leaf chemical properties using hyperspectral imaging. Front Plant Sci. 8, 1348. https://doi.org/10.3389/fpls.2017.01348.

Pape, J.-M., Klukas, C., 2014. 3-D histogram-based segmentation and leaf detection for rosette plants. Paper presented at the European Conference on Computer Vision.

Poorter, H., Fiorani, F., Stitt, M., Schurr, U., Finck, A., Gibon, Y., Usadel, B., Munns, R., Atkin, O.K., Tardieu, F., Pons, T.L., 2012. The art of growing plants for experimental purposes: a practical guide for the plant biologist. Funct. Plant Biol. 39 (11), 821–838. https://doi.org/10.1071/FP12028.

Purcell, L.C., 2000. Soybean canopy coverage and light interception measurements using digital imagery This paper is published with the approval of the director of the Arkansas Agric. Exp. Stn. (manuscript number 99–107). Crop. Sci. 40 (3), 834–837.

Rousseau, C., Belin, E., Bove, E., Rousseau, D., Fabre, F., Berruyer, R., Guillaumès, J., Manceau, C., Jacques, M.A., Boureau, T., 2013. High throughput quantitative phenotyping of plant resistance using chlorophyll fluorescence image analysis. Plant Methods 9 (1), 17.

Scharr, H., Minervini, M., French, A.P., Klukas, C., Kramer, D.M., Liu, X., Luengo, I., Pape, J.M., Polder, G., Vukadinovic, D., Yin, X., Tsaftaris, S.A., 2016. Leaf segmentation in plant phenotyping: a collation study. Mach. Vis. Appl. 27 (4), 585–606. https://doi.org/10.1007/s00138-015-0737-3.

Schirrmann, M., Hamdorf, A., Garz, A., Ustyuzhanin, A., Dammer, K.-H., 2016. Estimating wheat biomass by combining image clustering with crop height. Comput. Electron. Agric. 121, 374–384. https://doi.org/10.1016/j.compag.2016.01.007.

Schölkopf, B., Smola, A.J., Bach, F., 2002. Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond. MIT press.

Shipman, J.W., 2012, 10.16. 2012. Introduction to color theory. Retrieved from http://www.nmt.edu/tcc/help/pubs/colortheory/web/hsv.html.

Smith, A.R., 1978. Color gamut transform pairs. Comput. Graph (ACM) 12 (3), 12–19. https://doi.org/10.1145/965139.807361.

Snavely, N., Seitz, S.M., Szeliski, R., 2008. Modeling the world from internet photo collections. Int. J. Comput. Vision 80 (2), 189–210.

USDA-ERS, 2018. Oil Crops Outlook. Retrieved from http://www.ers.usda.gov.

Valliammal, N., Geethalakshmi, S., 2012. Plant leaf segmentation using non linear K means clustering. Int. J. Comput. Sci. Issues 9 (3), 212–218.

Vollmann, J., Walter, H., Sato, T., Schweiger, P., 2011. Digital image analysis and chlorophyll metering for phenotyping the effects of nodulation in soybean. Comput. Electron. Agric. 75 (1), 190–195. https://doi.org/10.1016/j.compag.2010.11.003.

Walter, A., Studer, B., Kolliker, R., 2012. Advanced phenotyping offers opportunities for improved breeding of forage and turf species. Ann. Bot. 110 (6), 1271–1279. https://doi.org/10.1093/aob/mcs026.

Westoby, M., Brasington, J., Glasser, N., Hambrey, M., Reynolds, J., 2012. 'Structure-from-Motion'photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology 179, 300–314.

Yin, X., Liu, X., Chen, J., Kramer, D.M., 2014. Multi-leaf alignment from fluorescence plant images. Paper presented at the Applications of Computer Vision (WACV), 2014 IEEE Winter Conference.

Yu, N., Li, L., Schmitz, N., Tian, L.F., Greenberg, J.A., Diers, B.W., 2016. Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. Remote Sens. Environ. 187, 91–101. https://doi.org/10.1016/j.rse.2016.10.005.

Zhou, J., Chen, H., Zhou, J., Fu, X., Ye, H., Nguyen, H.T., 2018. Development of an automated phenotyping platform for quantifying soybean dynamic responses to salinity stress in greenhouse environment. Comput. Electron. Agric. 151, 319–330. https://doi.org/10.1016/j.compag.2018.06.016.

Zhou, J., Fu, X., Schumacher, L., Zhou, J., 2018. Evaluating geometric measurement accuracy based on 3D reconstruction of automated imagery in a greenhouse. Sensors (Basel, Switzerland) 18 (7).