



## Research Paper

## Multi-view 3D reconstruction of seedling using 2D image contour



Qingguang Chen <sup>a,\*</sup>, Shentao Huang <sup>a</sup>, Shuang Liu <sup>a</sup>, Mingwei Zhong <sup>a</sup>, Guohao Zhang <sup>a</sup>, Liang Song <sup>a</sup>, Xinghao Zhang <sup>a</sup>, Jingcheng Zhang <sup>a</sup>, Kaihua Wu <sup>a</sup>, Ziran Ye <sup>b</sup>, Dedong Kong <sup>b</sup>

<sup>a</sup> School of Automation, Hangzhou Dianzi University, Hangzhou, 310018, China

<sup>b</sup> Institute of Digital Agriculture, Zhejiang Academy of Agricultural Sciences, Hangzhou, 310021, China

## ARTICLE INFO

## ABSTRACT

## Keywords:

RGB-D camera  
rotation axis optimisation  
Projected contour constraint  
Point cloud  
Phenotypic measurement

3D reconstruction of seedling can provide comprehensive and quantitative spatial structure information, offering an effective digital tool for breeding research. However, accurate and efficient reconstruction of seedling is still a challenging work due to limited performance of depth sensor for seedling with small-size stem and unavoidable error for multi-view point cloud registration. Therefore, in this paper, we propose an accurate multi-view 3D reconstruction method for seedling using 2D image contour to constrain 3D point cloud. The rotation axis is calibrated and optimised by minimising point-to-contour distance between 2D image contour and projected exterior points from 3D point cloud. Then, to remove outliers and noise, we introduce the seedling mask of 2D image to constrained and delete projected outlier points of 3D model from corresponding view. Furthermore, we propose a residual-guided method to recognise missing region for 3D model and complete 3D model of small-size stem. Finally, we can obtain an accurate 3D model of seedling. The reconstruction accuracy is evaluated by average distance between projected contour of 3D model and 2D image contour of all views (0.3185 mm). Then, the phenotypic parameters were calculated from 3D model and the results are close to manual measurements (Plant height:  $R^2 = 0.98$ , RMSE = 2.3 mm, rRMSE = 1.52%; Petioles inclination angle:  $R^2 = 0.99$ , RMSE = 0.73°, rRMSE = 1.41%; Leaf area:  $R^2 = 0.66$ , RMSE = 1.05 cm<sup>2</sup>, rRMSE = 7.63%; Leaf inclination angle:  $R^2 = 0.99$ , RMSE = 1.01°, rRMSE = 1.72%; Stem diameter:  $R^2 = 0.95$ , RMSE = 0.12 mm, rRMSE = 5.43%). Breeders can improve the selection of more resilient varieties and cultivars to different growing conditions starting from the dynamic analysis of their phenotype.

## Nomenclature

## (continued)

2D,3D	two-dimensional, three-dimensional
SV	Stereo Vision
SFM	Structure From Motion
NeRF	Neural Radiance Fields
LiDAR	Light Detection And Ranging
CT	Computed Tomography
RGB-D	depth image sensor
lm	Lumen
LED	Light Emitting Diode
PC	Personal Computer
KD-tree	K-Dimensional tree
ICP	Iterative Closest Points
<i>v</i>	view serial number
<i>V</i>	number of views
<i>RT</i> <sub>c</sub>	calibrated rotation and translation transformation matrices between view <i>v</i> camera pose and view 1 camera pose

(continued on next column)

<i>P</i> <sub>n</sub> <sup>v</sup>	the <i>n</i> -th corner points in view <i>v</i> image from different views are identified
<i>n</i>	the optimal fitting plane's normal unit vector
<i>A</i>	matrix of mean-centred points
<i>P̄</i>	mean value of corner points with same view
<i>P</i> <sub>rot</sub>	coordinates of projected point cloud
<i>P</i>	coordinates of point cloud
<i>k</i>	cross product between plane normal <i>n</i> and normal <i>n</i> <sub>0</sub> of the new x-y coords
<i>C</i> <sub>i</sub> = ( <i>x</i> <sub>i</sub> , <i>y</i> <sub>i</sub> )	coordinates of the <i>i</i> -th centre of the circle
<i>a</i> , <i>a'</i>	calibrated, optimised rotation axis direction vector
<i>b</i> , <i>b'</i>	calibrated, optimised rotation axis position
<i>k</i> , <i>k'</i>	calibrated, optimised the error correction coefficient between the motor rotation angle and the ratio of motor pulses
<i>θ</i> <sup>v</sup>	rotation angle from view <i>v</i> to view 1
<i>M</i>	the rotation and translation process

(continued on next page)

\* Corresponding author.

E-mail address: [optichen@hdu.edu.cn](mailto:optichen@hdu.edu.cn) (Q. Chen).

(continued)

$RT_o^v$	rotation and translation matrix calculated from original calibrated rotation axis
$\Pi$	perspective projection process
$cam$	the camera internal parameters
$P_{3d\_i}$	initial 3D model
$N$	number of spatial points in 3D model
$\mathbb{R}^{a \times b}$	real matrix of a columns and b rows
$P_{2d\_m}^v$	2D seedling mask point set
$P_{2d\_p}^v$	coordinates of 2D projected points
$P_{2d\_p}^v i = (x_i^v, y_i^v)$	the $i$ -th projected point
$dist$	the minimum distance calculation process between point and point set
$D_i^v$	the minimum distance between point and point set
$LeakyRelu(x)$	an activation function that output $x$ when $x \geq 0$ , and output $ax$ when $x < 0$ , and $a = 0.1$ is a constant
$\alpha$	distance threshold that determines whether a projected point is within the mask
$\operatorname{argminf}_{a,b}(a,b)$	minimum optimisation process to get the appropriate values of $a$ and $b$ such that objective function is minimised
$RT^v$	rotation and translation matrix calculated from optimised rotation axis
$P_{3d\_i}^v i = (x_i^v, y_i^v, z_i^v)$	$i$ -th point of seedling point clouds from view v
$P_{3d\_ni}^v = (x_i^v, y_i^v, z_i^w)$	$i$ -th point of 3D model with noise in the world coordinate system
$NoiseFlag_i$	indicates whether $P_{3d\_ni}^v$ is noise or not
$S_{(x,y)}^v$	the skeleton map
$P_{2d\_pm}^v(x,y)$	2D projection map
$S_{sub}^v i$	$i$ -th sub-skeleton points
$\epsilon$	the computation process of the residual sum of the skeleton map and 2D projection map
$N_{sub}^v i$	the number of $i$ -th sub-skeleton points
$R_i$	region residual between the $i$ -th sub-skeleton map and 2D projection map
$I_{min}, I_{max}$	minimum index and maximum index of sub-skeleton points
$S_{sub\_nci}$	$i$ -th sub-skeleton points set to be completed
$L$	cylinder central axis
$d$	cylinder diameter (mm)
$mn$	umber of matching points
$(x_{i,n}, y_{i,n}, z_{i,n})$	$(x_{j,n}, y_{j,n}, z_{j,n})$ the coordinates of the $n$ -th matching point in view $i$ and $j$
$E_c$	calibration error
$IoU$	Intersection over Union
$G$	ground truth
$SR$	segmented result
$TP$	target point cloud
$OP$	other point clouds
$H$	Hausdorff distance
$h_{avg}$	mean Hausdorff distance
$C_{\alpha\_shape}$	the process of extracting contours using alpha-shape
$P_{2d\_f}^v$	the view v final 3D model projection map
$P_{f\_c}^v$	2D contour points of the view v final 3D model projection map
$P_{m\_c}^v$	2D contour points of the view v seedling mask
$D_p$	projected point-contour distance (pixel)
$k_{pt}$	conversion coefficient from image pixels to ground truth scale
$N_v$	number of projected contour points in the view v
$D_c$	projected point-contour distance (mm)
$LA$	leaf area
$PH$	plant height
$SD$	stem diameter
$LIA$	leaf inclination angle
$PIA$	petiole inclination angle
$pi$	half perimeter of $i$ -th triangulated facet
$a_i, b_i, c_i$	triangle side length
$Y_{max}, Y_{min}$	the highest point of the 3D model, the lowest point of the 3D model
$l$	vector normal to the plane centred on the leaf centroid
$z$	zenith direction
$s$	direction vector along the petiole
$RMSE$	oot Mean Square Error
$R^2$	the coefficient of determination R-Square
$rRMSE$	relative Root Mean Square Error
$MAPE$	Mean Absolute Percentage Error

## 1. Introduction

Plant phenotype is a quantitative description of observable traits of plants caused by the interaction of its organism and environment during the growth process (Li et al., 2020a). It can be used for monitoring the occurrence of plant stress in crop cultivation such as disease, pests, nutrient stress at early stage. The accurate phenotypic parameters can provide effective guidance for crop breeding and growth management to improve breeding selection, crop productivity, reduce environmental impact and increase food security. (Li et al., 2020c; Sultan, 2003; Tardieu et al., 2017). Modern agriculture uses phenotypic data to implement closed control for digital, precise, standardised farm management (Li, Y. et al., 2021; Simbeye D.S. et al., 2023).

Morphological traits of plants including plant height, canopy leaf area index, leaf area, leaf count, and stem diameter contribute to many important functional features (Noshita K. et al., 2022). In the past, RGB images are widely utilised for analysing plant structure to reflect plant textures, disease resistance capabilities, and other attributes. However, RGB imaging have inherent limitations due to 2D imaging dimensionality reduction and plant structural complexity. For instance, the surface area of a curved leaf in a 2D image is significantly smaller than its true area (Gibbs et al., 2018).

3D reconstruction can provide spatial structure of plant for 3D morphological phenotypic data. It is crucial for capturing plant changes, quantifying growth across multiple spatial and temporal scales (Harandi et al., 2023; Jin et al., 2021; Sun & Wang, 2019). 3D reconstruction technique can be categorised as passive and active method. Passive method uses multi-view 2D images to reconstruct 3D model by feature point matching and parameters optimisation. Stereo Vision (SV) uses two or more cameras to simultaneously capture RGB images of the target from different views. Nguyen et al. (2016) develops a multi-view stereo vision system for 3D reconstruction and phenotypic analysis of plants to investigate phenotypic changes during plant growth. Bao et al. (2019) employs stereo vision for 3D reconstruction of field sorghum, demonstrating the applicability of stereo vision for field-based 3D plant phenotyping. Structure from Motion (SFM) uses a moved camera around the target to capture multi-view images with overlapped regions between adjacent images to reconstruct 3D model. Rossi et al. (2020) utilises SFM to reconstruct 3D plant models, and proposes an automatic segmentation algorithm to measure 3D phenotypic results of potted tomato (Rossi et al., 2022). However, passive imaging-based 3D reconstruction method is susceptible to environmental lighting conditions. SFM requires sequences of images to be captured, often resulting in data redundancy (Jay et al., 2015) and time-consuming method (Lu et al., 2020; Zhou et al., 2018). There is a trade-off between reconstruction accuracy and time. Some researchers also consider employing neural networks to reconstruct 3D models from 2D images, such as Neural Radiance Fields (NeRF) (Mildenhall et al., 2021). Unlike SFM, which estimates camera motion trajectories and object depths, NeRF trains a neural network to estimate the colour and density of each spatial point. This model can then be used to generate new views and achieve realistic rendering effects, thereby reconstructing lifelike 3D models. However, NeRF requires high image quality and computational performance, and its reconstruction accuracy needs further investigation. Currently, NeRF has not yet been widely adopted in the field of 3D plant phenotyping research.

Active 3D reconstruction methods use projected light including X-ray, pulsed laser, structured pattern to encode distance information. X-ray CT is used to reconstruct internal spatial structures of plants tissues and organs for studying plant phenotypic traits and the transport forms of internal plant substances (Piovesan et al., 2021; Li et al. (2020b) reconstructed the 3D phenotype of sorghum flowers using X-ray CT. The seed number, panicle depth, and panicle width measured from the 3D model were highly consistent with actual measurements, with R2 values of 0.98, 0.94, and 0.92, respectively. This study reveals the relationship between plant phenotype, genetic identity, and domestication history.

Light Detection and Ranging (LiDAR) uses time of flight (ToF) principal to detect range. It can't reconstruct the internal plant substances, but its cost is lower than X-ray CT and has commendable ranging precision. Forero et al. (2022) establishes a LiDAR-based 3D plant reconstruction platform to create a 3D maize phenotype database with a plant height average absolute error of 0.03 cm. Red Green Blue - Depth (RGB-D) camera uses binocular stereo imaging to reconstruct depth information by matching projected random speckle points. RGB-D camera can provide 2D colour image and 3D depth image simultaneously. Due to rich information and low cost, RGB-D camera has widespread application in the field of plant 3D reconstruction and phenotyping research. Lao et al. (2019) successfully reconstructs a 3D maize model using a consumer-level depth camera, achieving reconstruction times and accuracy surpassing those of the commercial software. The average measurement error for leaf length is 4%. This shows the feasibility of utilising consumer-level RGB-D cameras in the search of 3D plant phenotyping. Currently, utilising RGB-D cameras, researchers can reconstruct 3D models of plants using images from a single perspective or multiple perspectives. Ma et al. (2022) employs an array of RGB-D cameras fixed on an indoor iron frame to reconstruct 3D models of potted plants, measuring plant height and leaf area index value with R2 values more than 0.94. Qiu et al. (2021, 2022) mounts an RGB-D camera on a movable cart's underside for field-based maize 3D reconstruction, combining thermal imaging to imbue the 3D model with temperature information and analyse the impact of crop water stress on canopy-air temperature differences. R2 for plant heights manually measured and estimated are all greater than 0.95. Syed et al. (2019) develops an RGB-D integrated detection system to horizontally capture seedlings, detecting changes in seedling height and stem diameter. The R2 values for plant height and stem diameter are greater than 0.99 and 0.35, respectively. These researchers achieve partial 3D model reconstruction using single-view depth and colour images, typically obtaining phenotype information determinable from a single viewpoint, such as plant height and stem diameter. However, more researchers aim to reconstruct complete 3D phenotype models to capture comprehensive phenotype information. Teng et al. (2021) proposes a neighbourhood extreme filtering method and an improved iterative closest point (ICP) algorithm with adaptive registration threshold, achieving 3D reconstruction of rapeseed's full growth cycle with a single RGB-D camera. Liu, Y et al. (2023) mounts an RGB-D camera on each side of the plant to rapidly reconstruct 3D peanut model using plant self-rotation to capture multi-view images. Then, they obtained measurements of plant height and bounding box volume from the 3D model, with R2 values of 0.99 and 0.96, respectively.

Although extensive investigations of 3D reconstruction using multi-view imaging, there are still challenges for object with small size due to limited spatial resolution and measurement accuracy of depth camera. For stems of the seedling with diameter less than 3 mm, depth camera measurements might significantly deviate from true positions, causing substantial distortion in the reconstructed 3D model. This is due to the scarcity of speckles from the infrared projector's speckle pattern or laser beam on small stems, resulting in increased measurement errors or even loss of depth information in those areas. This makes incomplete stem structures to be a common issue in 3D reconstruction of seedlings. To tackle this issue, several researchers had investigated methods for completing point clouds. Jiang et al. (2022) extract L1-medial skeletons from the plant point cloud to determine the position of the missing stem. Then, they use Bezier curve fitting based on the skeleton points at both ends of the missing stem to complete point cloud. This method directly repairs the point cloud without requiring additional input, and exhibited high generality. However, stem point clouds generate by fitting lack authenticity based solely on the data at the two ends of the missing stem. Lu et al. (2020) have explored using the relationship between SFM reconstruction results and image sequence quality to enhance data quality, thereby proposing the use of image sequences captured from optimal angles to reconstruct complete plant

point clouds. This approach has led to improved reconstruction accuracy. However, it is exclusively applicable to SFM-based 3D reconstruction. Multi-view 3D reconstruction also urgently needs to address the point cloud stitching problem. Typically, using a rotation axis to stitch point clouds can result in significant errors. Wang et al. (2020) improve the stitching accuracy of multi-view point clouds by employing registration algorithms such as ICP. However, registration methods often consume a significant amount of time (Sun & Wang, 2019). Additionally, this incremental registration method accumulates stitching errors as the number of viewpoints increases, thereby reducing the stitching accuracy.

Due to the limited performance of depth sensors on seedlings with small-size stems and unavoidable errors in multi-view point cloud registration, accurately and efficiently reconstructing seedlings remains challenging. This paper presents an accurate multi-view 3D reconstruction method for seedlings using 2D image contour to constrain 3D point cloud. This method projects the point cloud onto colour images from different views using the camera pose and the camera intrinsic parameters. The relationship between the plant's contour in the colour image and the projected points constrains the 3D model, removing points outside the contour for denoising and filling missing points inside the contour for point cloud completion. 3D model provides a non-destructive and precise measurement method for breeders to track seedling growth in real time, study seedling adaptability to various stresses, and select more resilient varieties. This paper reports the development of a novel method for 3D reconstruction of seedlings from different species and estimation of their phenotypic features.

## 2. Materials and methods

### 2.1. Experiment system

As shown in Fig. 1, a low-cost multi-view 3D reconstruction experimental system for accurate reconstruction of seedlings is designed and constructed. This system consists of a 3D reconstruction platform and 3D reconstruction algorithms. The platform includes a circular rail, sliding platform, servo motor and controller, LED light panel, and Intel RealSense L515 LiDAR camera. The camera and servo motor are mounted on the sliding platform, which could move on the circular rail controlled by a Portable Computer (PC). During experiment, the camera captures multi-view images of seedlings at the centre of the circular rail with a resolution of 1920 × 1080 pixel. Considering that ambient light will cause outlier noise and measurement error of depth value, 2064 lumen controllable LED light panels are installed along the four sides of the platform to provide uniform lighting environment. The light source is turned on for 2D image capturing, but turned off for depth image collection by a light controller. In our experiment, imaging chamber is set as 78cm × 78 cm × 47 cm. According to the camera's field of view parameters (70° × 55°) and the working distance D to the object, the actual image size can be calculated as 35cm × 26 cm using Eq. (1). Therefore, the object size is limited to be within 30cm × 20 cm.

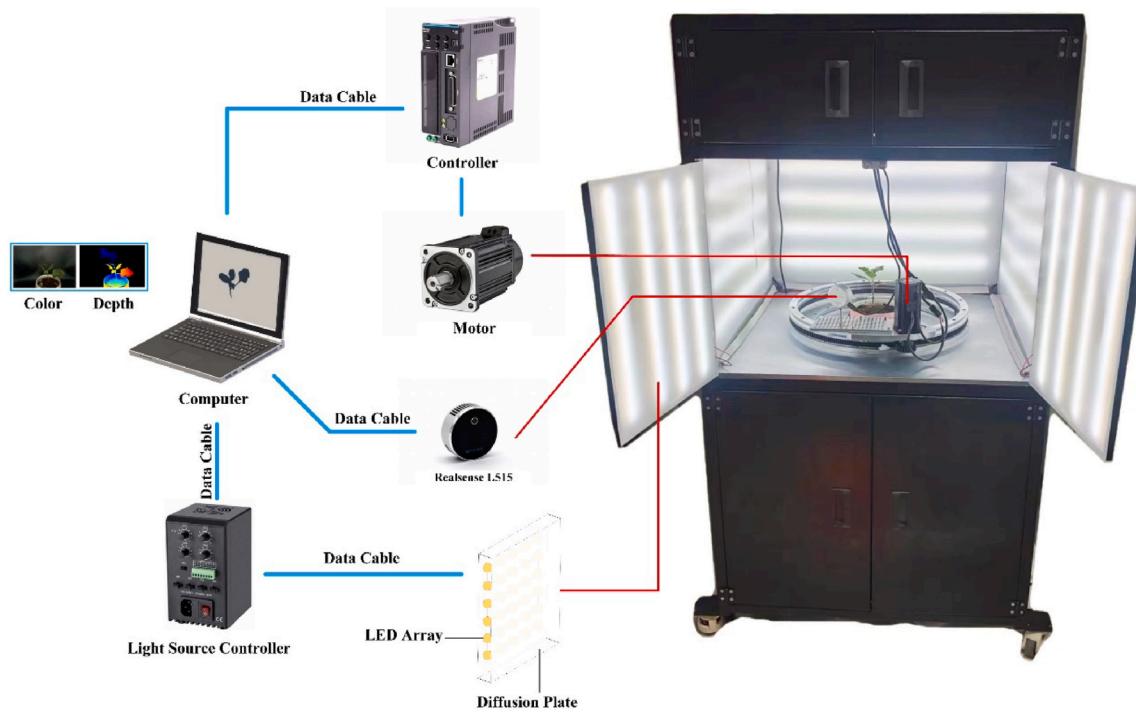
The maximum height and width of the plant that can be captured ( $p_H$  and  $p_W$ , respectively) can be calculated according to:

$$(p_H, p_W) = \left( 2 * \tan \frac{\theta_H}{2} * D, 2 * \tan \frac{\theta_W}{2} * D \right) \quad (1)$$

where  $\theta_H$  and  $\theta_W$  denote the camera's vertical and horizontal field of view angles, respectively, and D is the distance to the object.

Fig. 2 illustrates algorithm workflow for multi-view 3D reconstruction, which consisted of rotation axis calibration, data acquisition, rotation axis optimisation, point cloud denoising, point cloud completion. The specific description is as follows:

Rotation axis calibration and optimisation algorithm improves calibration accuracy by minimising point-to-contour distance between 2D image contour and projected exterior points from 3D point cloud. It



**Fig. 1.** Schematic representation of the proposed 3D reconstruction experimental setup.

plays a fundamental role to fuse multi-view point cloud to reconstruct accurate 3D models of seedling. U<sup>2</sup>-Net model is used for image segmentation to obtain the contour of seedling for constraining 3D point cloud. The 3D reconstruction process comprises point cloud stitching, point cloud denoising, and a point cloud completion. Point cloud stitching merges multi-viewpoint point clouds into a single 3D model. Point cloud denoising algorithm achieves noise reduction by removing points outside the contour constraints. The point cloud completion algorithm identifies the small-size stem regions in the 3D model that require repair through residual-guided. It fills these regions with cylindrical point clouds and then employs the point cloud denoising algorithm to remove redundant points, ultimately obtaining a complete 3D model.

## 2.2. Rotation axis calibration and optimisation

### 2.2.1. Rotation axis calibration

In our experiment, multi-view 2D images and 3D point clouds are captured along circular rail. Therefore, the rotation axis calibration is necessary for point cloud fusing. Fig. 3 illustrates the rotation axis calibration method. The camera is mounted on the sliding platform of the circular rail, while the calibration board is positioned at the centre of the circular rail. The camera rotates counterclockwise along the circular rail and captures consecutive calibration board images from view  $v$ . Then, the captured images are subjected to camera calibration using MATLAB software, resulting in the acquisition of rotation and translation transformation matrices  $RT_c^v$  between different camera poses. Subsequently, the corner points  $P_n^v$  in the images from different views are identified, with subscripts  $n$  denoting the corner point index within the same view and superscripts  $v$  indicating the view number. The corner points with the same subscript are subjected to plane fitting, and the optimal fitting plane's normal unit vector  $\mathbf{n}$  is obtained by applying singular value decomposition (SVD) to  $A$ . The optimisation objective can be formulated as:

$$\underset{\|\mathbf{n}\|=1}{\operatorname{argmin}} \|\mathbf{A}\mathbf{n}\|^2 \quad (2)$$

$$A = (P_n^1 - \bar{P}, P_n^2 - \bar{P}, \dots, P_n^v - \bar{P})^T \quad (3)$$

where  $A$  is matrix of mean-centred points,  $\bar{P}$  is mean value of corner points with same view.

Subsequently, the point cloud is projected onto  $x$ - $y$  coords using the Rodriguez rotation formula (Rodrigues, 1840) (Eq. (4)).

$$P_{rot} = P + (1 - \cos(\theta))\mathbf{k} \times (\mathbf{k} \times P) + (\mathbf{k} \times P)\sin(\theta) \quad (4)$$

where  $\mathbf{k}$  is cross product between plane normal  $\mathbf{n}$  and normal  $\mathbf{n}_0$  of the new  $x$ - $y$  coords.  $\theta$  is rotation angle from  $\mathbf{n}$  to  $\mathbf{n}_0$ .  $P$  is coordinates of point cloud.  $P_{rot}$  is coordinates of projected point cloud. In Fig. 4 (a), the solid rectangle represents the plane fitted to the spatial points, while the dashed rectangle represents the  $x$ - $y$  plane. The spatial points on the fitted plane are rotated around the rotation axis  $\mathbf{k}$  and projected onto the  $x$ - $y$  plane. Thus, the spatial circle fitting problem is reduced to a planar circle fitting problem.

A least squares circle fitting is then performed on the projected points in the 2D plane to determine the coordinates of the circle centre. Then, these circle centre coordinates are transformed back into 3D space using the Rodriguez rotation formula (where  $\mathbf{k}$  is cross product between normal  $\mathbf{n}_0$  of the new  $x$ - $y$  coords and plane normal  $\mathbf{n}$ .  $\theta$  is rotation angle from  $\mathbf{n}_0$  to  $\mathbf{n}$ ). In Fig. 4 (b), the solid rectangle represents the  $x$ - $y$  plane, while the dashed rectangle represents the plane fitted to the spatial points. The calculated centre of the circle on the  $x$ - $y$  plane is then projected back onto the fitted plane through rotation around the axis  $\mathbf{k}$ , resulting in the coordinates of the three-dimensional centre.

Finally, a least-squares linear fitting of all the fitted circle centres is conducted to obtain the direction and position of the rotation axis in the camera coordinate system. The function can be formulated as:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i z_i & \sum_{i=1}^n x_i \\ \sum_{i=1}^n y_i z_i & \sum_{i=1}^n y_i \end{bmatrix} \begin{bmatrix} \sum_{i=1}^n z_i^2 & \sum_{i=1}^n z_i \\ \sum_{i=1}^n z_i & n \end{bmatrix}^{-1} \quad (5)$$

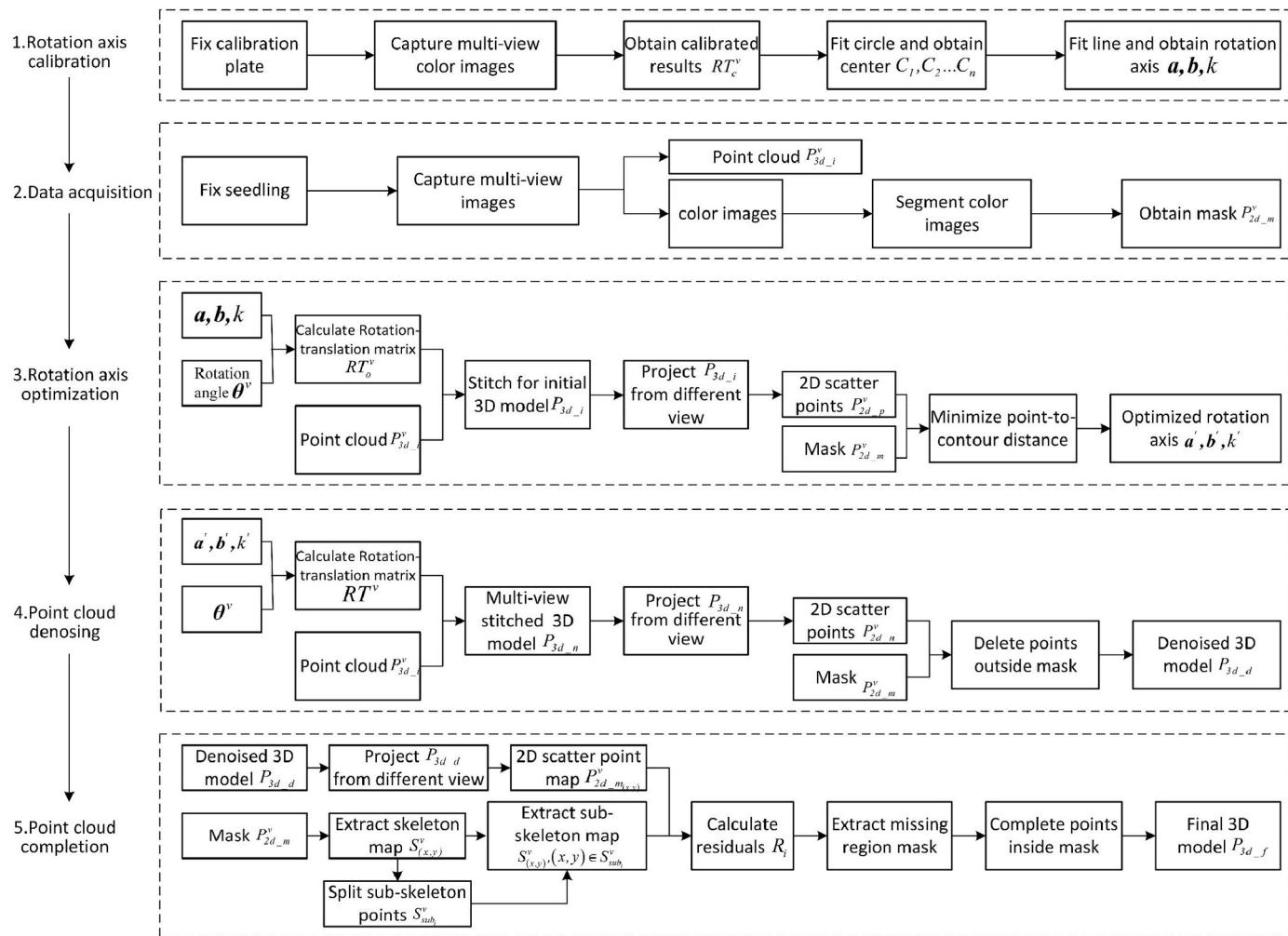


Fig. 2. Algorithm workflow for our proposed 3D reconstruction of seedlings.

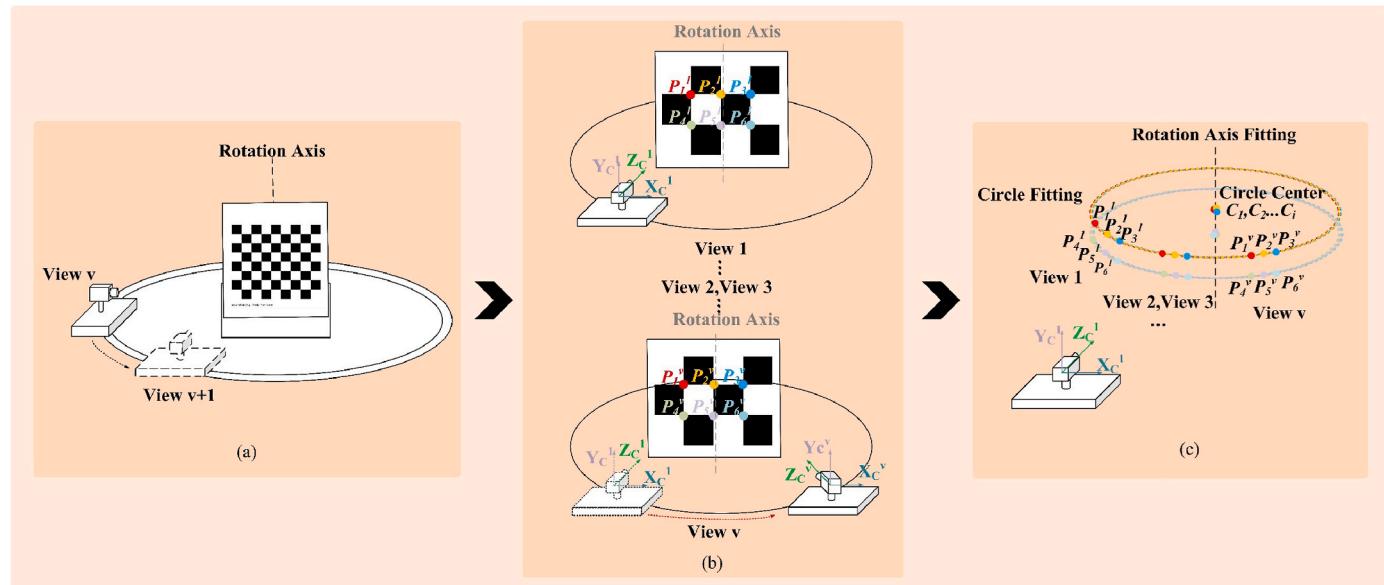
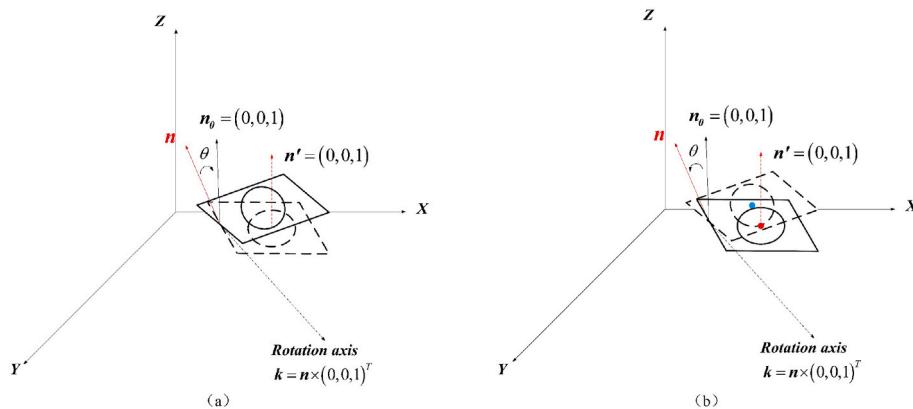


Fig. 3. Schematic representation of rotation axis calibration (a) Sequence image acquisition, (b) Corner points extraction of different views, (c) Rotation axis fitting.



**Fig. 4.** Using the Rodrigues rotation formula (a) Project the point cloud onto the x-y plane, (b) Convert the projection on the x-y plane back to 3D space.

where  $(x_i, y_i, z_i)$  is coordinates of the  $i$ -th centre of the circle  $C_i$ .  $a$ ,  $b$ ,  $c$  and  $d$  are the coefficients of the linear equation that can be expressed as  $\frac{x-b}{a} = \frac{y-d}{c} = \frac{z}{1}$ .

### 2.2.2. Rotation axis calibration optimisation

Accurate rotation axis calibration can shorten registration time and improve multi-view fusion performance. In practical implementation, rotation axis calibration is influenced by various factors including camera calibration accuracy, camera performance, lighting conditions, placement of the calibration board, and the number of calibration images. There is obvious offset among the stitched multi-view point cloud using calibrated results. Although the ICP registration algorithm boasts higher stitching accuracy, it is extremely time-consuming. To improve calibration performance and stitching efficiency, we propose a rotation axis calibration optimisation algorithm based on projected contour consistency constraint, which minimises the distance between the projected points outside the mask contour from different perspectives and the mask contour to optimise the rotation axis parameters. The optimisation objective can be formulated as:

$$\arg \min_{a,b,k} \sum_{v=1}^V \sum_{i=1}^N \text{LeakyRelu}(D_i^v - \alpha) \quad (6)$$

where  $D_i^v$  is the projection Euclidean distance.  $\text{LeakyRelu}(x)$  is an activation function that output  $x$  when  $x \geq 0$ , and output  $ax$  when  $x < 0$ , and  $a = 0.1$ .

The projection distance  $D_i^v$  is defined as follows:

$$D_i^v = \text{dist}\left(\left(\Pi_{RT_o^v, \text{cam}}(P_{3d-i})\right)_i, P_{2d-m}^v\right) \quad (7)$$

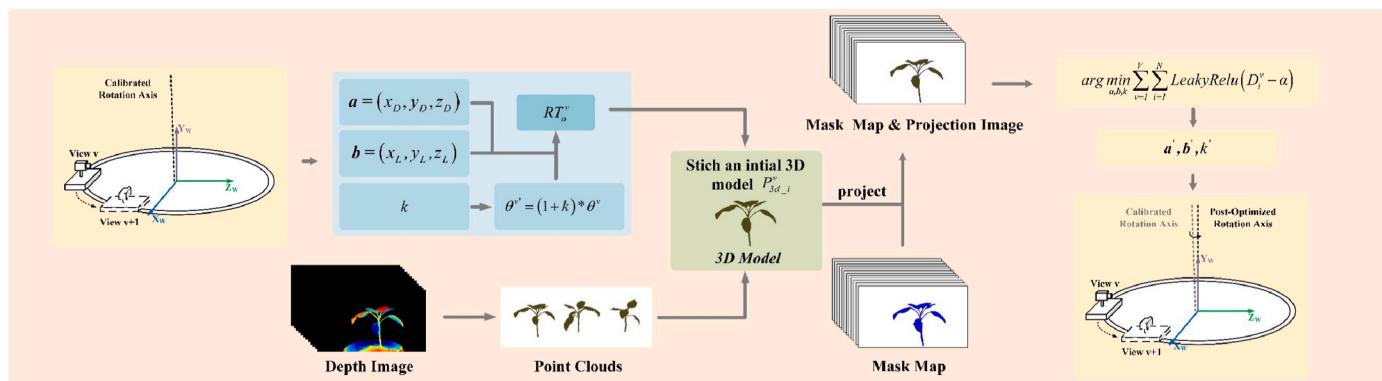
where  $\Pi$  represents perspective projection process using the rotation-translation matrix  $RT_o^v$  and the camera internal parameters  $\text{cam}$ , and  $\text{dist}$  represents the distance between the  $i$ -th projection point of the initial 3D model  $P_{3d-i} \in \mathbb{R}^{3 \times N}$  onto the mask of the view  $v$  and the nearest point within the 2D seedling mask point set  $P_{2d-m}^v$ .

The rotation and translation matrix  $RT_o^v$  is defined as follows:

$$RT_o^v = M(a, b, (1+k)\theta^v) \quad (8)$$

where  $M$  represents the rotation and translation process using rotation axis direction vector  $a$ , position  $b$  and rotation angle  $\theta^v$  from view  $v$  to view 1.  $k$  represents the error correction coefficient between the motor rotation angle and the ratio of motor pulses.

The workflow of the optimisation algorithm is illustrated in Fig. 5. Initially, the optimisation model acquires the depth image from the view  $v$  of the reconstructed object. Subsequently, considering the camera coordinate system of the view 1 as the world coordinate system, the optimisation model calculates the rotation-translation matrix  $RT_o^v$  from the camera coordinate system of view  $v$  to the world coordinate system. This computation is based on the parameters  $a$  and  $b$  of the rotation axis, the error correction coefficient  $k$  and the rotation angle  $\theta^v$  (Eq. (8)). The point clouds generated from depth images are transformed using  $RT_o^v$  from their camera coordinate systems to the world coordinate system, resulting in the 3D model  $P_{3d-i}$ . Subsequently,  $P_{3d-i}$  is projected onto the  $v$ -th 2D seedling mask to obtain the 2D projected points  $P_{2d-p}^v \in \mathbb{R}^{2 \times N}$ . Next, a KD-tree is established for the  $v$ -th seedling mask points  $P_{2d-m}^v$ . Within



**Fig. 5.** Schematic representation of rotation axis optimisation. Blue box: The rotation and transformation matrix Calculating that converts the camera coordinate system of the view  $v$  to the world coordinate system  $RT_o^v$ . Green box: Projection of the point cloud model from the world coordinate system to the view  $v$  image coordinate system. Yellow box: Optimisation through the minimisation of the distance between the 2D image contour and the projected exterior points from the 3D point cloud. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

the KD-tree, a search is conducted to find the nearest seedling mask pixel point to each  $P_{2d\_p}^v = (x_i^v, y_i^v)$  (where  $P_{2d\_p}^v$  represents the coordinates of the  $i$ -th projected point in  $P_{2d\_p}$ ). Then,  $D_i^v$  between these two points is computed (Eq. (7)).  $\text{LeakyRelu}(x)$  is employed to intensify the penalisation for points outside the contour, while points within the contour are rewarded slightly. Finally, the adaptive moment estimation (Adam) optimisation algorithm is employed to solve this optimisation problem (Eq. (6)), resulting in the optimised parameters  $a'$ ,  $b'$ , and  $k'$ . As the value of optimisation objective decreased, it indicates a higher level of overlap between  $P_{2d\_p}^v$  and  $P_{2d\_m}^v$ , reflecting the increased accuracy of the relative pose relationship  $RT^v$  and enhanced precision in rotation axis calibration. This method improves the accuracy of rotation axis calibration and avoids the need for ICP registration, thereby improving reconstruction efficiency.

### 2.2.3. Evaluation of rotational axis calibration

Based on the optimised rotation axis, the corner points of the calibration board from different views are stitched together. Points corresponding to the same physical corner location but in different camera coordinate systems are considered as matching points. Firstly, we generate the initial view corner point cloud as target point cloud. Then, using both the calibrated and optimised rotation axis, we stitch the point clouds from other views onto the target point cloud. Evaluation of the rotation axis calibration is performed by calculating the mean stitching error of all matched points, denoted as  $E_c$  (Eq. (9)), which corresponds to the distances between the blue or green points and the red points.

$$E_c = \frac{1}{V} \sum_{i=1}^V \frac{1}{V-i} \sum_{j=i}^V \frac{1}{m} \sum_{n=1}^m \sqrt{(x_{i,n} - x_{j,n})^2 + (y_{i,n} - y_{j,n})^2 + (z_{i,n} - z_{j,n})^2} \quad (9)$$

where  $V$  is the number of views,  $m$  is the number of matching points, and  $(x_{i,n}, y_{i,n}, z_{i,n})$ ,  $(x_{j,n}, y_{j,n}, z_{j,n})$  represent the coordinates of the  $n$ -th matching point in view  $i$  and  $j$ , respectively.

### 2.3. $U^2$ -Net-based segmentation

To obtain contour mask for optimising contour distance from different view, we introduce  $U^2$ -Net deep learning model to segment seedling images (Qin et al., 2020).  $U^2$ -Net has a nested two-level U-shaped architecture with greater depth and offer an enriched multi-scale feature, while maintaining low memory overhead and computational costs. Notably, the  $U^2$ -Net architecture is built using residual U-blocks and does not rely on any pretrained backbones adapted from image classification, rendering it both straightforward and flexible.

A dataset is created by collecting 120 images using the 3D reconstruction system described in Section 2.1 and manually annotating them using the Labelme software. Expanding the dataset to 600 images by performing data augmentation operations such as rotation, cropping, and brightness adjustments on the dataset images. Then, the  $U^2$ -Net model is trained with 420 images in total, while the 180 images are used for testing the model and observing the pixel-wise results. Finally, the trained segmentation model is utilised to provide seedling mask images for the subsequent methodologies proposed in this paper.

To demonstrate the effectiveness of our proposed  $U^2$ -Net-based segmentation method, Intersection over Union (IoU), Recall and Precision are used to evaluate the performance of our segmentation (Eq (10), (11) and (12)).

$$\text{IoU} = \frac{|G \cap SR|}{|G \cup SR|} \quad (10)$$

where  $G$  is the ground truth, and  $SR$  is the segmented result based on  $U^2$ -Net (they are a two-dimensional matrix, with 1 indicating seedlings).

$$\text{Recall} = \frac{TP}{(TP + FN)} \quad (11)$$

where  $TP$  represents the model predicting a positive for a positive, and  $FN$  represents the model predicting a negative for a positive.

$$\text{Precision} = \frac{TP}{(TP + FP)} \quad (12)$$

where  $FP$  represents the model predicting a positive for a negative.

### 2.4. 3D reconstruction of seedling

To reconstruct accurate 3D model of seedling using captured multi-view images and point cloud, there are three steps: multi-view stitching, point cloud denoising, and point cloud completion for fine stem.

#### 2.4.1. Multi-view stitching

The 3D reconstruction platform automatically captures colour images and depth images from  $V$  views. Point clouds are generated from the depth images. The camera coordinate system of the first view is adopted as the world coordinate system. Utilising the optimised rotation axis,  $RT^v$  is computed to transform the camera coordinate system of view  $v$  to the world coordinate system.  $P'_{3d\_i}$  from view  $v$  is then transformed to the world coordinate system. By traversing all views, the point clouds are ultimately stitched together to form a 3D model  $P_{3d\_n}$  with noise in the world coordinate system (Eq. (13)).

$$\begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \\ 1 \end{bmatrix} = RT^v \begin{bmatrix} x_i^v \\ y_i^v \\ z_i^v \\ 1 \end{bmatrix} \quad (13)$$

where  $P'_{3d\_i} = (x_i^v, y_i^v, z_i^v)$  is  $i$ -th point of seedling point clouds, and  $P_{3d\_n} = (x_i^w, y_i^w, z_i^w)$  is  $i$ -th point of 3D model.

Hausdorff distance is used to evaluate the accuracy of multi-view point cloud stitching. Initially, the first view is selected as the target point cloud. The point clouds from other views are then stitched to the target point cloud using different methods. The Hausdorff distance  $H$  between the target point cloud  $TP$  and each of the other point clouds  $OP$  is calculated to generate a heatmap that evaluates the stitching accuracy (Equation (14)). The mean Hausdorff distance  $h_{avg}$  is also calculated (Equation (15)).

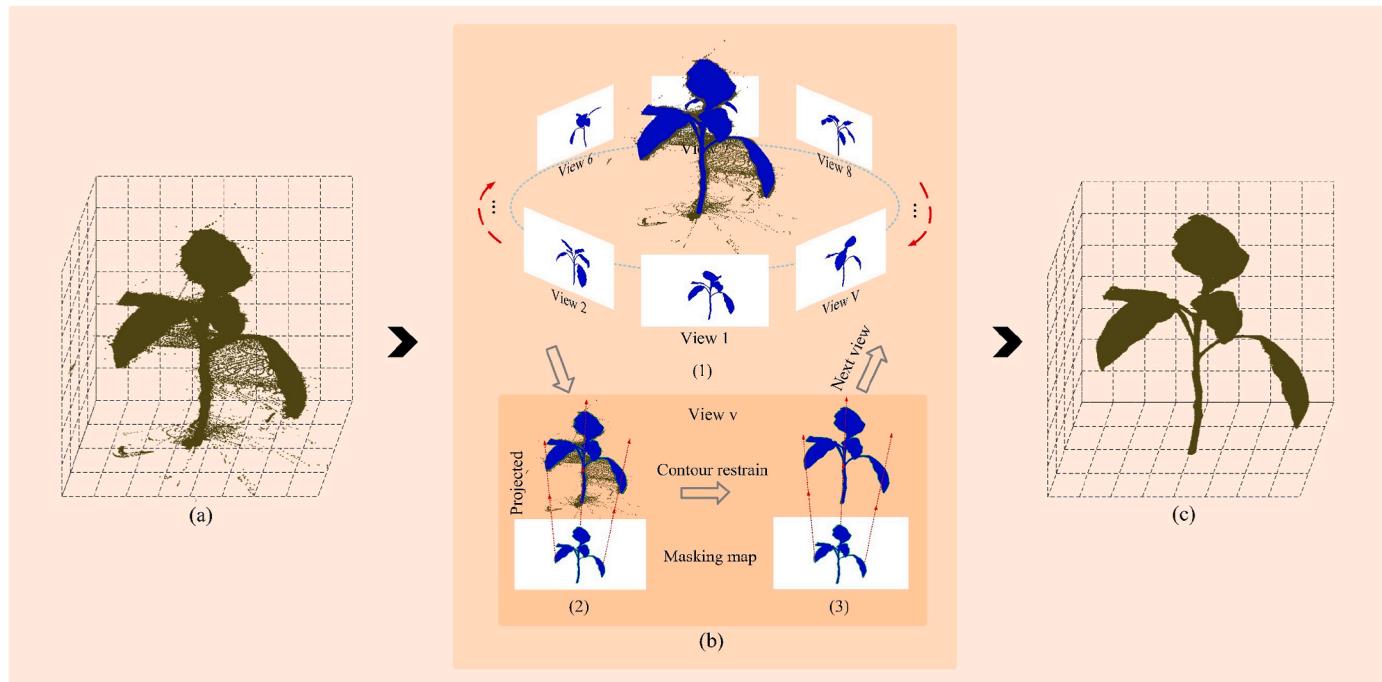
$$H(TP, OP) = \max \left\{ \max_{a \in TP} \left\{ \min_{b \in OP} \|a - b\| \right\}, \max_{b \in OP} \left\{ \min_{a \in TP} \|b - a\| \right\} \right\} \quad (14)$$

where  $a, b$  are spatial points in  $TP$  and  $OP$ , respectively.

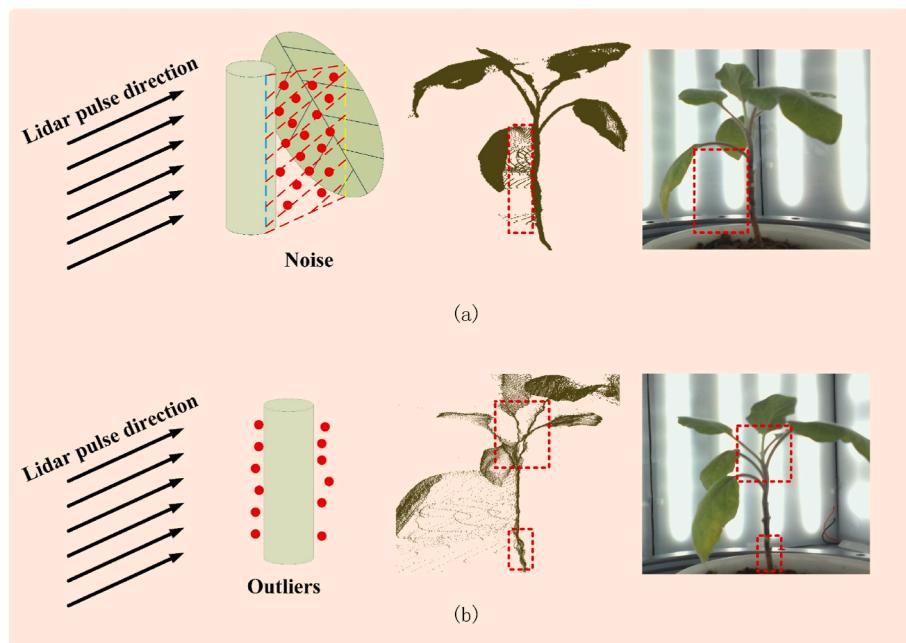
$$h_{avg}(TP, OP) = \text{avg}_{a \in TP} \left\{ \min_{b \in OP} \|a - b\| \right\} \quad (15)$$

#### 2.4.2. Point cloud denoising

RGB-D camera usually introduces a multitude of noise and outliers (Wang, J. et al., 2020). Fig. 6 presents the distribution of noise and outliers encountered during the 3D reconstruction process using depth cameras. Noise (Fig. 7 (a)) is typically observed along the edges of leaves or stems. This phenomenon arises from the inherent divergence angle in the laser emitted by time-of-flight (ToF) depth cameras. When a laser beam is emitted, it concurrently strikes two surfaces, causing the radar to erroneously determine that the measurement target lies between these two surfaces. Then, outliers are commonly found near the object's surface, leading to irregularities in the surface of 3D model (Fig. 7 (b)). These outliers result from inherent measurement errors in the depth



**Fig. 6.** Schematic representation of point cloud filtering algorithm based on projected contour consistency constraint (a) Point cloud to be denoised, (b) (1) Applying constraint-based denoising through traversal of multi-view mask maps for the point cloud, (2) (3) Illustration of constraint-based denoising on point cloud using single view mask map, (c) Denoised 3D point cloud.



**Fig. 7.** Schematic diagram of noise sources (a) The blue dashed line represents the edges of foreground objects, while the yellow dashed line represents the projection of these edges onto the surfaces of background objects. Due to the inherent divergence angle of the laser beam, a significant amount of noise is observed on the planes defined by the blue and yellow dashed lines, (b) The red outliers, stemming from measurement errors inherent to the laser radar system, are distributed beyond the surface of the measured object and are located very close to the surface. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

camera, leading to noticeable distortions in the 3D model of the stem.

Traditional denoising methods, such as statistical and spatial filtering, require iterative experimentation to fine-tune parameter thresholds for enhanced denoising quality. However, these approaches often suffer from inefficiency. To address this challenge, this study proposes a point cloud denoising algorithm based on projected contour

consistency constraint. This approach effectively denoises the registered seedling point cloud without necessitating parameter adjustments, thereby significantly enhancing filtering efficiency.

Firstly, 3D model  $P_{3d\_n}$  is projected onto the mask of view  $v$  as  $P_{2d\_n}^v$  using the known rotation-translation matrix  $RT^v$  and the camera intrinsic parameter matrix  $cam$ . Subsequently, based on whether  $i$ -th

projected point  $P_{2d\_n}^v i$  belongs to the mask  $P_{2d\_m}^v$  of seedling, the presence of noise points in the 3D model is determined according to Eq. (16). If  $Noise\_flag_i = 1$ , then the point  $P_{3d\_ni}$  of 3D model associated with  $P_{2d\_p}^v i$  is classified as a noise point and is permanently removed. Iterate through all views, and ultimately output a denoised 3D model  $P_{3d\_d}$ .

$$Noise\_flag_i = \begin{cases} 1, P_{2d\_n}^v i \notin P_{2d\_m}^v \\ 0, P_{2d\_n}^v i \in P_{2d\_m}^v \end{cases} \quad (16)$$

#### 2.4.3. Point cloud completion

For small size stem with diameter less than 3 mm, measurement errors and denoising process will lead to missing region. To reconstruct complete 3D model of seedlings, we propose a novel small-size stem point cloud completion algorithm based on projected contour consistency constraint (Fig. 8). It consists of two steps: residual-guided missing region recognition, projected contour-constraint point cloud completion.

Region residual is defined as the sum of difference between 2D projection map  $P_{2d\_pm}^v(x, y)$  of 3D model and the skeleton map  $S_{(x,y)}^v$  from mask from all views. They are 2D matrices. And for each element of the matrix, if the pixel at the corresponding element coordinate belongs to the point cloud projection or the seedling mask, it is set to 1.

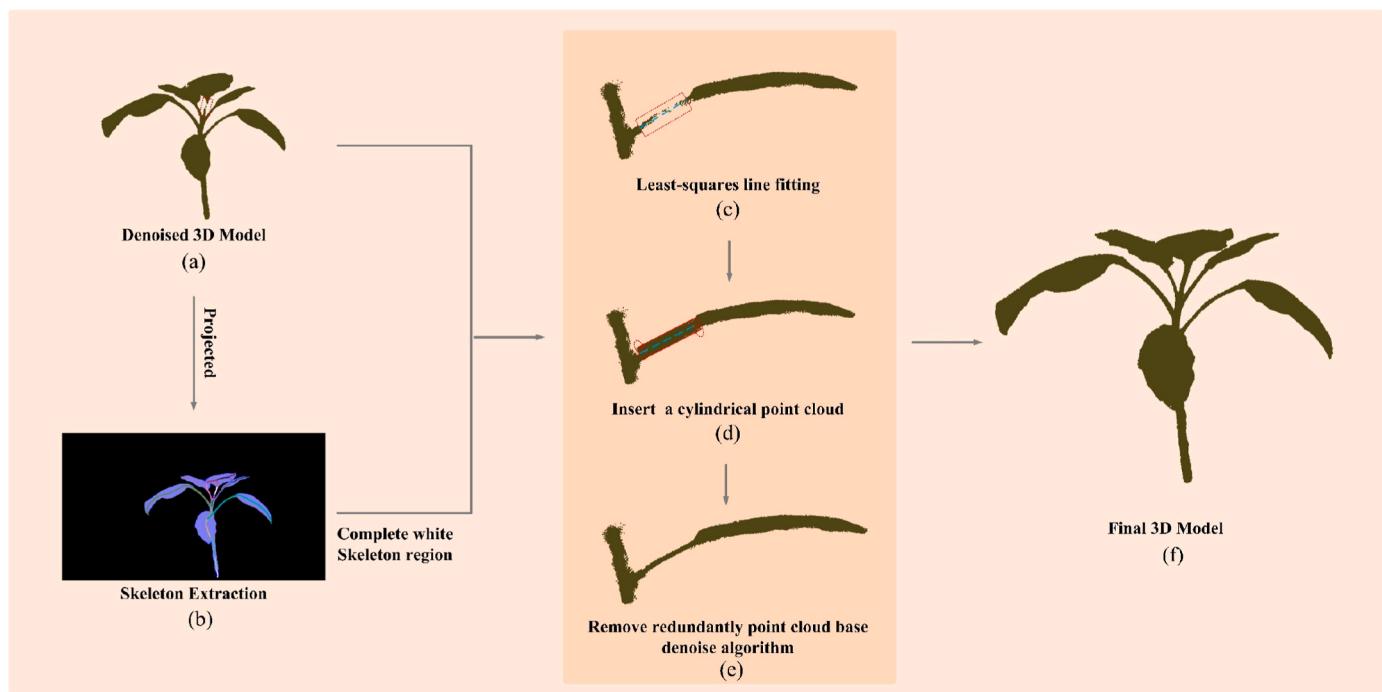
$$R_i = \frac{\sum_{(x,y) \in S_{sub}^v} (S_{(x,y)}^v - P_{2d\_pm}^v(x, y))}{N_{sub}^v i} \quad (16)$$

where  $R_i$  represents sub-skeleton region residual,  $\epsilon$  represents the computation process of the residual sum of  $S_{(x,y)}^v$  and  $P_{2d\_pm}^v(x, y)$  at coordinates included in  $i$ -th sub-skeleton points  $S_{sub}^v i$ .  $N_{sub}^v i$  represents the number of  $S_{sub}^v i$ .

Missing region recognition is performed by computing the residual between the binarized seedling skeleton map and the binarized projection map of 3D model to identify regions for completion. At first, the skimage python library's median skeleton extraction function is utilised

to extract the skeleton list  $S_{(x,y)}^v$  from the mask map of the seedling region. While iterating through all points in  $S_{(x,y)}^v$ , if 8-neighbourhood of a point contains more than three skeleton points, its neighboring points are identified as skeleton intersection points and are removed. Then skeleton points are subdivided into multiple sub-skeletons points  $S_{sub}^v i$ , enhancing the accuracy of region localisation for completion. The sub-skeleton points are sorted in descending order based on abscissa, followed by sorting based on ordinate if abscissa is equal. This sorting process assigns a sequential index  $I$  to each point in  $S_{sub}^v i$ . For the points of sub-skeletons, we calculate the residuals between  $S_{(x,y)}^v$  and  $P_{2d\_pm}^v(x, y)$ . And recording the minimum index  $I_{min}$  and maximum index  $I_{max}$ . In our paper, it is proposed that when  $R_i$  exceeds 0.1 (Equation (16)), the sub-skeleton is identified as a missing region. This threshold can exclude leaves and stems with almost no missing parts from being considered for repair, thereby enhancing the reconstruction efficiency. Then sub-skeleton points between index  $I_{min} - \beta$  and  $I_{max} + \beta$  are selected as the skeleton points to be completed  $S_{sub\_nci}$  (Fig. 8 (b)). To ensure sufficient coverage of 3D model projection points within the missing region, it is imperative to set an optimal  $\beta$  value. A smaller  $\beta$  value results in a scarcity of projected points, while a larger  $\beta$  value encompasses points from neighboring areas, impacting the spatial localisation accuracy of the incomplete section. Through meticulous parameter tuning,  $\beta = 5$  is determined to yield the most effective restoration outcome (see the supplementary material for details). These points will be utilised to localise the position of the 3D missing region to be completed in the world coordinate system. Then, along the skeleton  $S_{sub\_nci}$ , the perpendicular line to the line connecting  $i$ -th skeleton point and the next skeletal point is calculated. Starting from  $i$ -th skeleton point, a bidirectional search along the perpendicular line is performed to locate the boundary of the seedling mask, and the search distance  $d$  is recorded. This process yields the seedling mask that needs to be completed.

In the second step, the point cloud within the seedling mask is extracted, and a least-squares line fitting is performed to obtain line



**Fig. 8.** Schematic representation of small-size stem point cloud completion algorithm based on projected contour consistency constraint (a) The original 3D model of the seedling for repair, (b) Extraction of the median skeleton from the seedling mask, (c) Least squares linear fitting of the point cloud within the area to be repaired, (d) Filling cylindrical point cloud based on the fitted line from (c), (e) Segmentation of the filled point cloud based on our filtering algorithm, (f) Completed 3D model of the seedling.

segment  $L$  (Fig. 8 (c)), which matches the length of the corresponding sub-skeleton. A cylindrical point cloud is inserted, with  $L$  serving as cylinder central axis and 3 times the search distance  $d$  as the cylinder diameter (Fig. 8 (d)) (see the supplementary material for details). The completed cylindrical point cloud is then subject to a point cloud denoising algorithm (Section 2.4.2). This algorithm segments the supplementary point cloud and removes redundantly added points as noise (Fig. 8 (e)). Iterate through all views and repeat the aforementioned process to eventually obtain the completed 3D model. (Fig. 8 (f)).

This proposed approach provides an effective solution for completing missing point clouds of fine plant stems by leveraging multi-view 2D contour constraints, ensuring a more accurate and comprehensive reconstruction of plant seedlings.

#### 2.4.4. Evaluation of 3D model reconstruction

To demonstrate the robustness of the proposed approach, 3D point cloud models were reconstructed for seven different species of seedlings, including the three varieties of eggplant seedlings, two varieties of chili pepper seedlings, okra seedlings, and cucumber seedlings. The three varieties of eggplants are represented by four plants each, while the other four varieties have two plants each, totalling twenty plants. Each species exhibits unique characteristics: eggplant seedlings have dispersed spatial distribution of stems and leaves, with large and flat leaves in diverse orientations; chili pepper seedlings have relatively dense spatial distribution of stems and leaves, forming complex canopies among seedlings; okra seedlings feature regular leaf distribution, often with three leaves growing on a short branch of the main stem; cucumber seedlings have sparse leaves, but the leaf shapes are intricate and exhibit pronounced curling.

In this paper, the reconstruction accuracy is quantified by demonstrating the consistency of projected contours to assess the similarity between the reconstruction results and the real plants. At first, we employ the alpha-shape method to extract the 2D contour points  $P_{f\_c}^v$  of the view  $v$  final 3D model projection map  $P_{2d,f}^v$  as well as the 2D contour points  $P_{m\_c}^v$  of the view  $v$  seedling mask  $P_{2d,m}^v$  (Edelsbrunner et al., 2010) (Eq. (18) and (19)).

$$P_{f_c}^v = C_{\alpha\text{-shape}}(P_{2d_f}^v) \quad (18)$$

$$P_{m_c}^v = C_{\alpha\text{-shape}}(P_{2d_m}^v) \quad (19)$$

where  $C_{\alpha\text{-shape}}$  represents the process of extracting contours using alpha-shape.

Then, we use the quantifiable metric  $D_p$  and  $D_c$  to assess the reconstruction precision of the 3D model (Eq. (20) and (21)), where function  $dist$  is defined as Eq. (7).

$$D_p = \frac{1}{V} \sum_{v=1}^V \frac{\sum_{i=1}^{N_v} dist(P_{f_c}^v i, P_{m_c}^v)}{N_v} \quad (20)$$

$$D_c = k_{pt} D_p \quad (21)$$

where  $k_{pt}$  is the conversion coefficient from image pixels to ground truth scale, and  $N_v$  is the number of projected contour points in the view  $v$ .

#### 2.5. 3D model phenotypic measurement

In this paper, five phenotypic traits are measured: leaf area ( $LA$ ), plant height ( $PH$ ), stem diameter ( $SD$ ), leaf inclination angle ( $LIA$ ), and petiole inclination angle ( $PIA$ ). The measurement accuracy of these phenotypic traits is evaluated using quantitative indicators. The simulated values are calculated from the 3D models as follows (Li, Y. et al., 2018; Rossi et al., 2022):

$LA$  is the total sum of the area of triangulated facets obtained from

the 3D model of leaves (Eq. (22)).

$$LA = \sum_{i=1}^n \sqrt{\frac{p_i}{2} \left( \frac{p_i}{2} - a_i \right) \left( \frac{p_i}{2} - b_i \right) \left( \frac{p_i}{2} - c_i \right)} \quad (22)$$

where  $p_i$  is half perimeter of  $i$ -th triangulated facet,  $a_i$ ,  $b_i$ ,  $c_i$  is triangle side length,  $n$  is total of triangulated facets.

$PH$  is calculated as the difference between the highest and lowest points of the 3D model (Eq. (23)).

$$PH = Y_{max} - Y_{min} \quad (23)$$

where  $Y_{max}$  is the highest point of the 3D model, while  $Y_{min}$  is the lowest point of the 3D model.

$SD$  is calculated as the width of the cross-section, which is obtained from the main stem at a height of 2 cm above the soil and from the side branch 2 cm away from the branching point. And it is assumed that the section of the stem is a perfect circle.

$LIA$  is calculated as the angle between the vector normal  $\mathbf{l}$  to the plane centred on the leaf centroid and the zenith direction  $\mathbf{z}$  (in this study  $\mathbf{z} = (\mathbf{0}, \mathbf{1}, \mathbf{0})$ ) (Eq. (24)).

$$LIA = \cos^{-1} \left( \frac{\mathbf{l} \bullet \mathbf{z}}{|\mathbf{l}| |\mathbf{z}|} \right) \quad (24)$$

$PIA$  is calculated as the angle between the direction vector  $\mathbf{s}$  along the petiole and the zenith direction  $\mathbf{z}$  (Eq. (25)).

$$PIA = \cos^{-1} \left( \frac{\mathbf{s} \bullet \mathbf{z}}{|\mathbf{s}| |\mathbf{z}|} \right) \quad (25)$$

Manual phenotypic measurements are conducted to obtain ground truth values for plant phenotypic parameter comparison. A grid coordinate scale is affixed to a vertical wall, and the plant is positioned in front of the scale. The plant is then horizontally projected onto the coordinate scale, and measurements are taken to acquire the true values of  $PH$ . Adjusting the plant's position so that the directional vectors of the stem and leaves to be measured are parallel to the wall, the directional vectors of the stem and leaves are then projected onto the coordinate scale. Taking the positive y-axis direction of the scale as the zenith direction, the angle between the two vectors is measured to obtain the true values of  $PIA$  and  $LIA$ . Each measurement is repeated five times, and the average value is calculated. Subsequently, the seedling is destructively sampled. Leaves are flattened and placed onto the grid coordinate scale. Using a camera positioned orthogonally to the grid surface, images are captured to measure  $LA$ . The specific manual measurement methods for  $PH$ ,  $PIA$ ,  $LIA$  and  $LA$  are shown in Fig. 9. Finally, using a vernier calliper (the accuracy is 0.02 mm) to measure the width of the stems at a height of 2 cm above the soil and at a point 2 cm away from the branching point, and the average of the five measurements is taken as  $SD$ .

The accuracy of the model measurements is evaluated using the coefficient of determination R-Square( $R^2$ ), Root Mean Square Error (RMSE), relative Root Mean Square Error (rRMSE), Mean Absolute Percentage Error (MAPE) (Eq. (26), (27), (28) and (29)). The model accuracy is considered excellent when rRMSE  $\leq 10\%$  (Despotovic et al., 2016).

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (26)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (27)$$

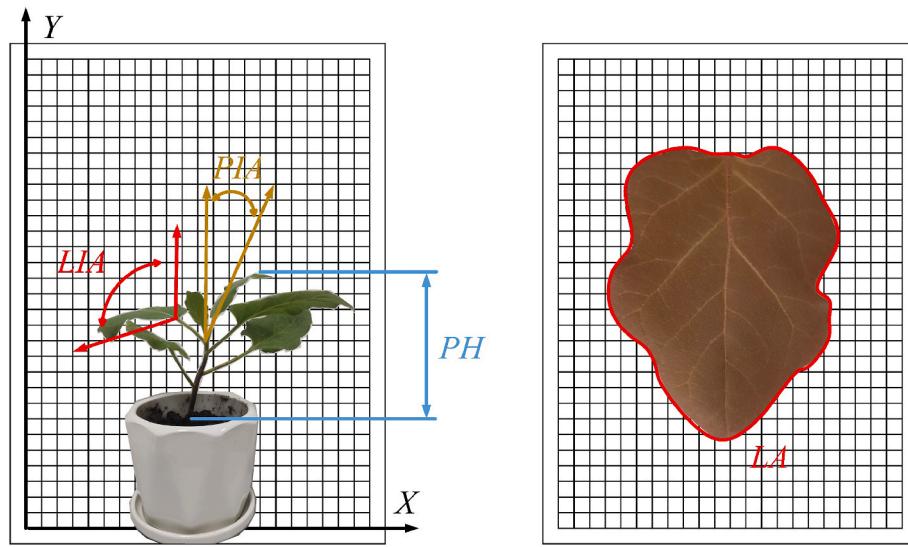


Fig. 9. Schematic representation of manual measurement.

$$rRMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} * 100\% \quad (28)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (29)$$

where  $\hat{y}_i$  is  $i$ -th simulated value,  $y_i$  is  $i$ -th ground truth value,  $\bar{y}$  is the average of the ground truth values.

### 3. Results

#### 3.1. Rotation axis calibration and optimisation

A stitching experiment was conducted using the optimised rotation axis and the calibrated rotation axis separately (Fig. 10). The red points represent the target point cloud, the green points represent the stitching results using the optimised rotation axis, and the blue points represent

the stitching results using the calibrated rotation axis. The point clouds of calibration board corners from 12 viewpoints are stitched, and the average stitching error  $E_c$  is employed to assess accuracy. The stitching errors before and after optimisation are 0.53 cm and 0.12 cm, respectively, indicating a reduction of 77.36% in stitching error.

#### 3.2. Semantic segmentation of seedling images

The 2D image segmentation results can be found in Fig. 11. The mean IoU is 0.94 for the test set (90 images). And the average recall and precision values are 0.98 and 0.95, respectively.

#### 3.3. 3D reconstruction results and evaluation

Fig. 12 illustrates the reconstruction process of the seedling, ranging from sparse to dense canopies, for seven different varieties. In Fig. 12 (a), real colour images of the reconstructed objects are presented, including eggplant seedlings (Fig. 12 a2, a5, and a6), chili pepper seedlings (Fig. 12 a3, and a7), okra seedlings (Fig. 12 a4), and cucumber

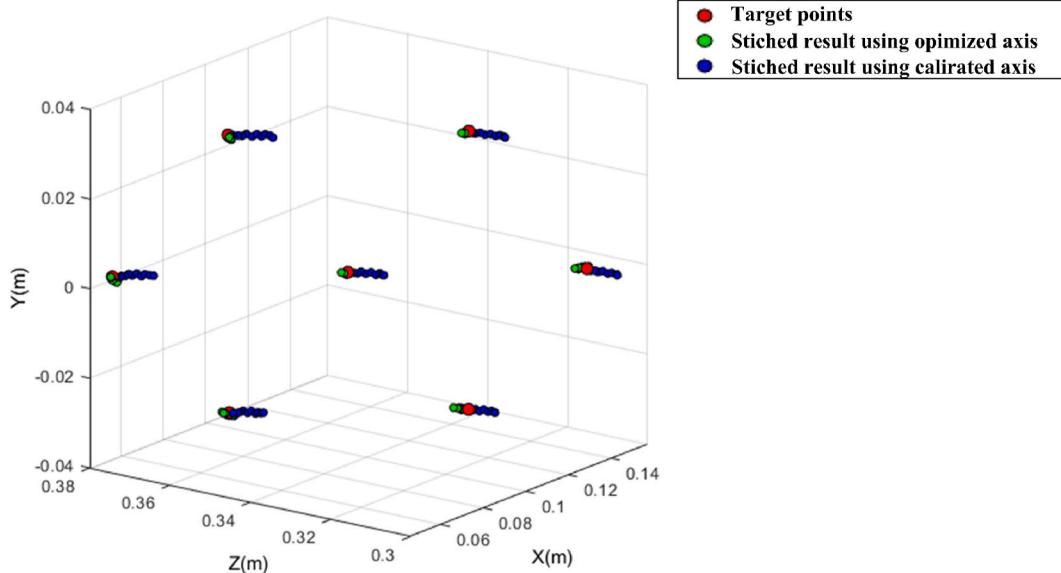
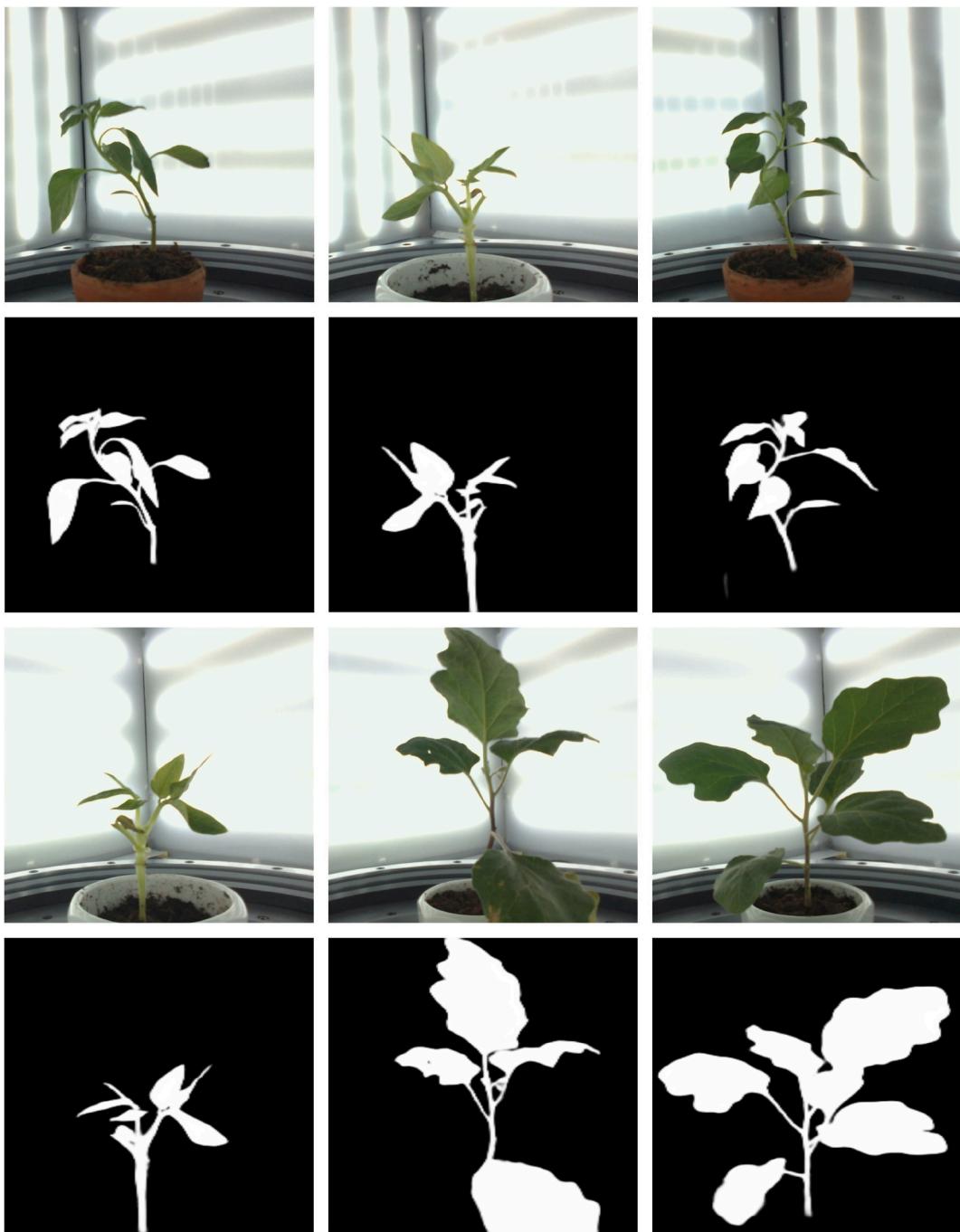


Fig. 10. Stitching multi-view calibration board corner point based on rotation axis.



**Fig. 11.** Captured RGB images and segmented mask results. The plants are chili peppers, okra, and eggplants that have grown for 30–45 days, respectively.

seedlings (Fig. 12 a1). Fig. 12 (b) displays the results of multi-view point cloud fusion based solely on rotation axes. At this stage, a substantial amount of noise points is present, resulting in unclear contours of the point cloud model. In Fig. 12 (c), the 3D model of the seedlings exhibits distinct stem and leaf contours after undergoing filtering algorithms. However, a missing stem point in the dashed box region is noticeable. This absence arises due to a significant depth measurement deviation caused by insufficient laser or auxiliary distance measurement patterns projected onto small-size stems by the RGB-D camera. Such discrepancies have been identified as outliers and removed during the filtering process. To render the 3D model of the small-size stem complete, a completion algorithm is applied. Fig. 12 (d) demonstrates the complete 3D model.

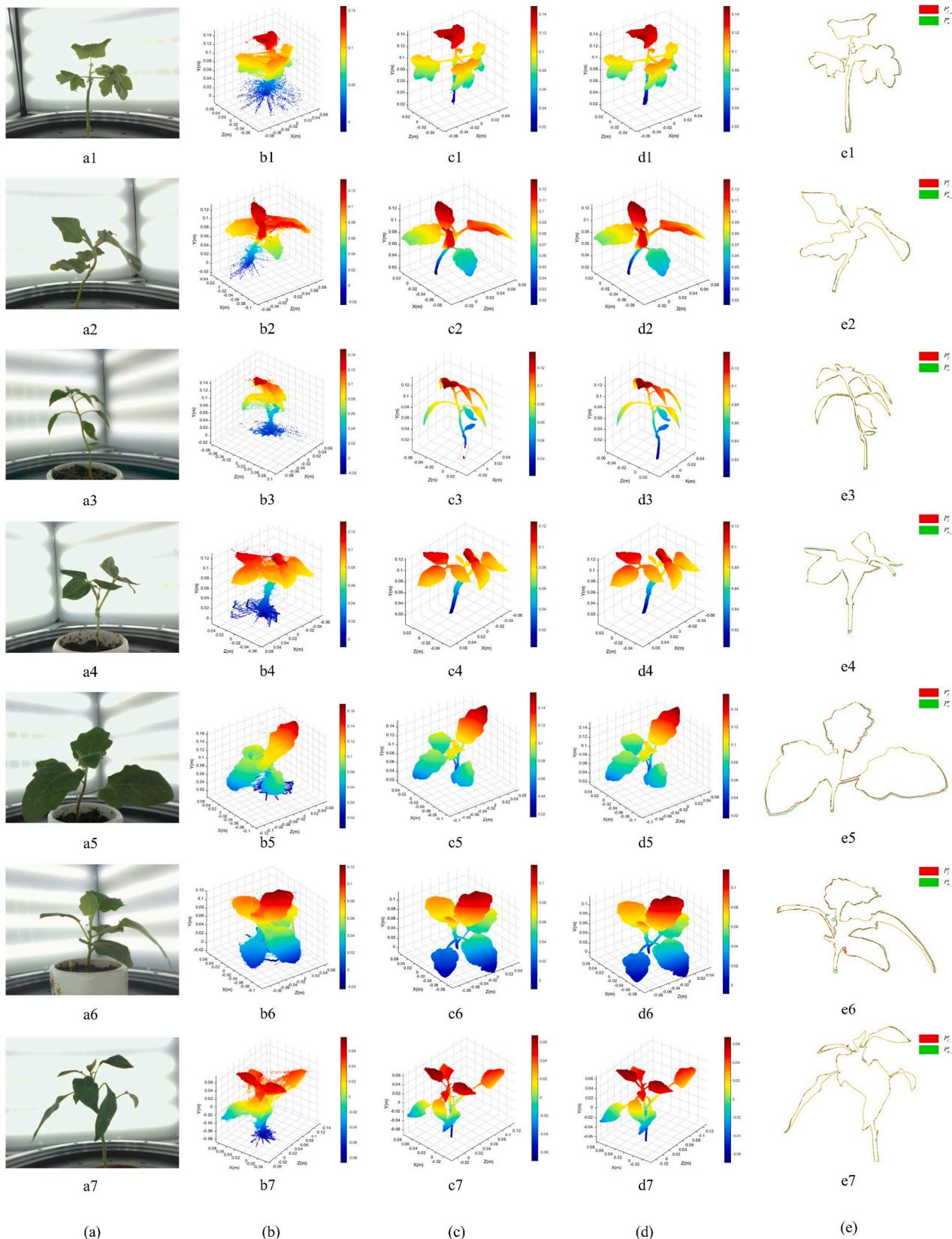
In this study, point cloud stitching is carried out based on the

optimised rotation axis. We compare the stitching results of our method with the results of stitching using the calibrated rotation axis and ICP (Fig. 13 (b) and (c)).

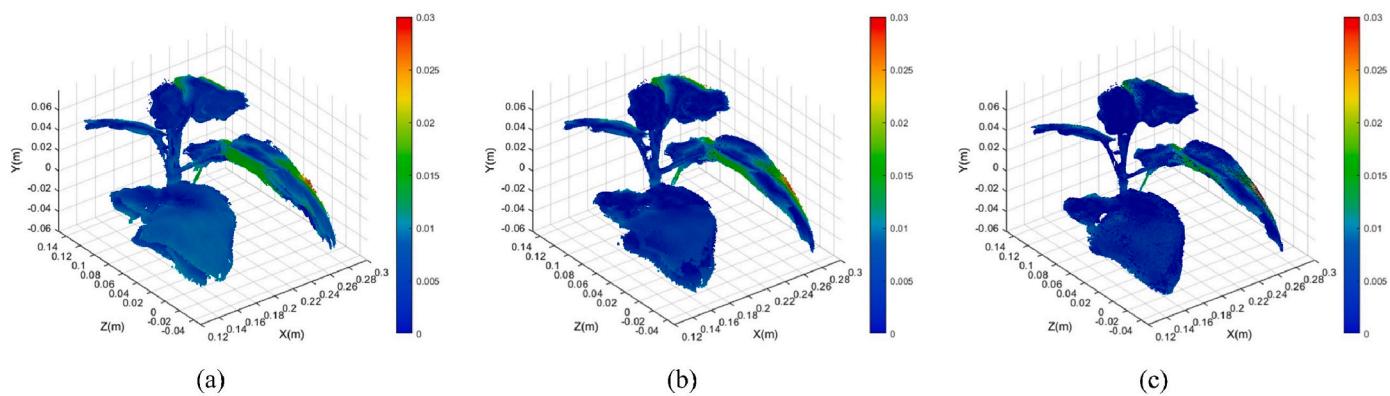
Fig. 13 displays the Hausdorff distance heatmaps for the three methods which  $h_{avg}$  is 5.279 mm, 4.253 mm, and 4.209 mm. Moreover, the stitching time is a crucial metric of interest in high-throughput 3D reconstruction. The stitching times for the three methods are 0.054 s, 73.23 s, and 0.04 s, respectively.

Calculate the projected point-contour distance for all reconstructed 3D models and compute the average value for each distance (Fig. 12(e)). From Table 1, the projection point-contour distance  $D_p$  and  $D_c$  are  $1.9749 \pm 0.8803$  pixels and  $0.3185 \pm 0.1722$  mm, respectively.

In this study, 3D models are reconstructed for a total of 20 seedlings of seven different varieties. Among them, one seedling from each variety



**Fig. 12.** The results of the seedling 3D model: (a) seedling colour image, (b) 3D model before denoising, (c) 3D model after denoising, with missing stem points indicated within the dashed box, (d) Complete 3D model after completion, (e) 2D Projected contour  $P_{f-c}^v$  of 3D model and 2D contour  $P_{m-c}^v$  of the seedling mask. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

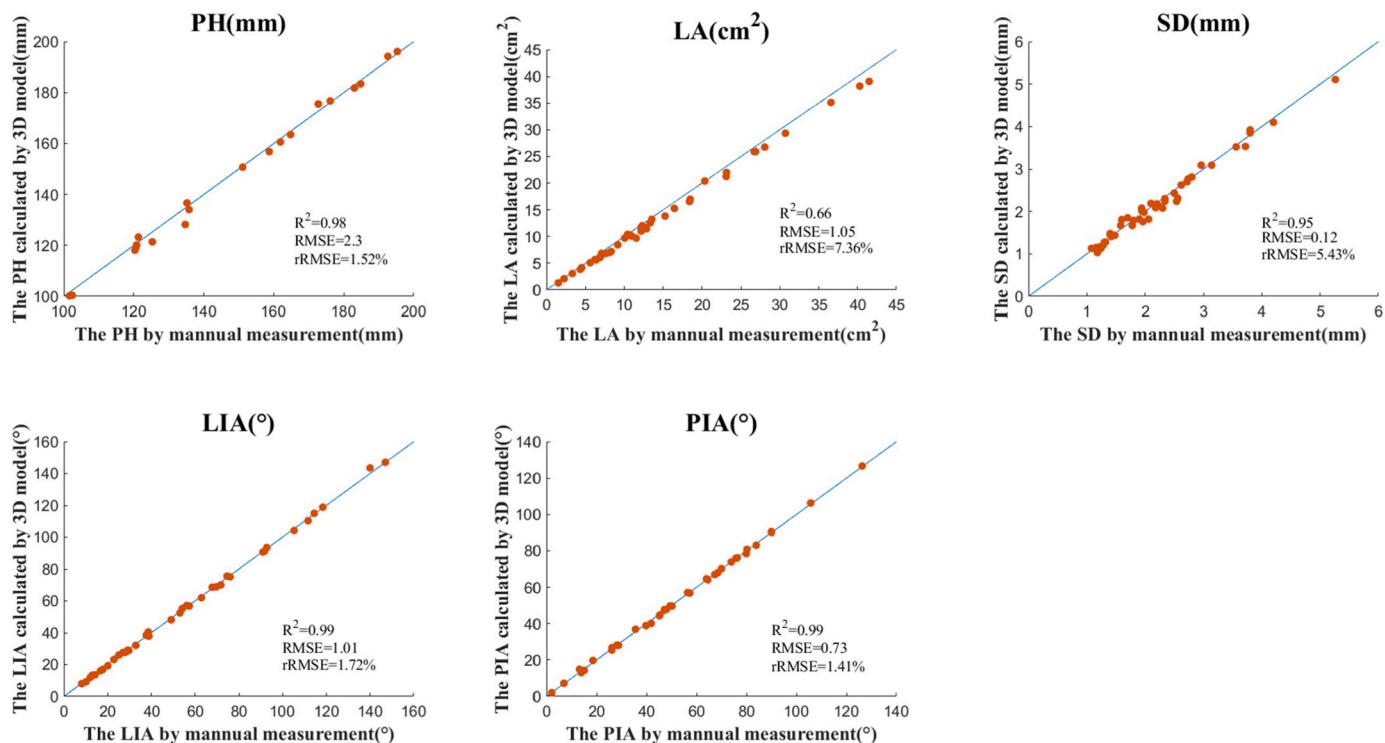


**Fig. 13.** Heat map for the Hausdorff distance using (a) calibrated rotation axis, (b) ICP algorithm, (c) our proposed optimised calibration results.

**Table 1**  
Performance evaluation of 3D Reconstruction using projected point-contour distance.

Species	Number	$D_p$ (pixel)	$k_{pt}$ (mm pixel $^{-1}$ )	$D_c$ (mm)
Black Round Eggplant	4	1.9950	0.1118	0.2230
Long Dragon Eggplant	4	1.0976	0.1523	0.1672
Purple Eggplant	4	1.5104	0.1713	0.2587
Hangzhou Pepper	2	1.9953	0.1948	0.3887
Long Green Pepper	2	3.4867	0.1783	0.6217
Okra	2	2.6929	0.1605	0.4322
Cucumber	2	1.0461	0.1319	0.1380
Average error	/	$1.9749 \pm 0.8803$	/	$0.3185 \pm 0.1722$

undergoes comprehensive phenotype data collection. The remaining 13 plants are used for *PH* measurements only, to enrich the *PH* data. The correlation between ground truth values of seedling and simulated values from 3D model is illustrated in Fig. 14. *PH* is calculated from 20 plants. *LA* and *LIA* is calculated from 43 leaves. *PIA* is calculated from 38 stems. *SD* is calculated from 46 stems. The calculated  $R^2$  between simulated and ground truth values is 0.98 for *PH*, 0.66 for *LA*, 0.95 for *SD*, 0.99 for *LIA*, and 0.99 for *PIA*. When the data distribution is biased towards one side of the 1:1 line, it may yield a poorer  $R^2$  result. However, the other two indicators for *LA* demonstrate its better measurement accuracy. The RMSE and rRMSE values for *PH*, *LA*, *SD*, *LIA*, and *PIA* (*PH*: 2.3 mm and 1.52%; *PIA*: 0.73° and 1.41%; *LA*: 1.05 cm $^2$  and 7.63%; *LIA*: 1.01° and 1.72%; *SD*: 0.12 mm and 5.43%) provide a direct insight into the specific error magnitude between simulated and ground truth values, demonstrating the potential for substituting manual measurements. The rRMSE are 1.52% for *PH*, 7.36% for *LA*, 5.43% for *SD*, 1.72% for *LIA*, and 1.41% for *PIA*. The values calculated from 3D models show high accuracies when compared to the relevant manual



**Fig. 14.** Compared results between the ground truth values by manual measurement and the calculated results from 3D model, the blue line represents the 1:1 line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

measurements.

#### 4. Discussion

In this study, we established a low-cost 3D reconstruction platform and accurate multi-view 3D reconstruction pipeline. Initially, the platform was calibrated using a rotation axis calibration.

optimisation algorithm to determine the rotation axis position and direction. Once calibrated, the results could consistently apply to subsequent reconstruction processes, as long as the fixed position of the camera remained unchanged. Subsequently, the platform automatically captured multi-view image sequences of the seedlings. The rotation axis was then utilised to stitch point clouds from different view. Post-processing of the stitched initial point cloud involved the point cloud filtering algorithm and a the small-size stem point cloud completion algorithm to reconstruct a 3D model with clear contours and complete structure. Finally, plant phenotypic parameters were measured based on the 3D model. The entire process of 3D model reconstruction and phenotypic measurement is non-destructive by optical imaging and computer processing. Consequently, plants can be measured to monitor morphological changes during the whole life. This platform offers a non-destructive, high-precision, and efficient measurement method for seedling phenotypic research, facilitating the tracking of seedling growth status.

##### 4.1. Rotation axis calibration

The rotation axis calibration optimisation algorithm was an improvement over existing rotation axis calibration method. In previous studies involving multi-view point cloud stitching for 3D reconstruction, researchers typically employed rotation axis calibration along with rotation angles to achieve rough point cloud stitching. Subsequently, they used the ICP registration algorithm to refine the stitching of all point clouds, ultimately obtaining the 3D model (Li et al., 2012). We conducted an analysis of the results from both rough stitching and ICP registration in terms of stitching accuracy and computational time. The results revealed that rough stitching had lower accuracy but significantly lower computational time, while ICP registration had higher accuracy but consumed a lot of time. To solve the problem between low accuracy in rough stitching and significant time-cost of ICP registration, a calibration optimisation method is proposed to enhance the accuracy of rotation axis calibration, achieving a balance between stitching time and accuracy that is comparable to ICP registration. In experimental comparisons with traditional methods, our approach exhibited a slightly improved stitching accuracy of 4.209 mm compared to the 4.253 mm achieved through ICP registration. Furthermore, our method achieved a much shorter stitching time of only 0.04 s, in stark contrast to the 73.23 s required by the ICP registration. These results demonstrated that our method addresses the issue of inadequate rotation axis calibration accuracy leading to larger stitching errors. Moreover, our approach achieved accuracy equivalent to ICP registration while significantly enhancing reconstruction efficiency. However, it should be noted that if the camera's relative position to the sliding platform changes due to external factors, recalibration and optimisation of the rotation axis are necessary. This process could be extremely time-consuming, with the current optimisation algorithm of the rotation axis taking around 2 h. As a result, the maintenance time cost of our equipment was relatively high.

##### 4.2. Denoising and completion

There are many studies about denoising of point cloud from RGB-D camera. Lao et al. (2019) analysed the ranging errors of RGB-D cameras on different distances from the wall during the maize plant reconstruction process. They found that the ranging results on the wall at the same distance fluctuated around the true value, with errors increasing with distance. Additionally, Teng et al. (2021) pointed out inherent

limitations of RGB-D cameras in terms of structure and imaging principles, leading to significant noise in point cloud acquisition. They demonstrated point cloud images captured in a single acquisition, showing more noise at the stem and resulting in distorted stems. This type of noise could adversely affect the accuracy of traditional ICP registration, leading to increased registration errors. Therefore, various filtering algorithms were needed to preprocess point clouds. In this paper, noise is categorised as two types: outliers (Fig. 7 (a)), which were scattered in distribution, and anomalies (Fig. 7 (b)), which were distributed on the surface of the point cloud model and were caused by measurement errors. Lao et al. (2019) used pass-through filtering and statistical filtering to denoise the maize point cloud, removing all outliers from the maize model. However, the effect of removing anomalies was not significant, as the reconstructed maize model indicated the presence of many anomalies in the stem, making the stem of the model appear thicker than the true maize stem in the colour image. Liang et al. (2020) proposed a colour-based statistical filtering method to filter coloured maize point clouds, removing outliers. However, anomalies still could not be effectively removed, resulting in blurred edges in the 3D point cloud. Teng et al. (2021) proposed a neighbourhood maximum filtering algorithm, which demonstrated better denoising effects compared to statistical filtering algorithms. Liu et al. (2023) employed traditional statistical filtering for point cloud denoising, resulting in some residual outliers in the peanut model, and removed sparsely distributed stems as outliers. These researchers' filtering methods required presetting thresholds. The selection of thresholds needed to strike a balance between removing as many outliers as possible and retaining as few non-outliers as possible, often requiring iterative experimentation to determine, which was time-consuming. Therefore, we proposed a threshold-free filtering method that could be directly applied to point cloud filtering for various objects. Our filtering method constrained the point cloud based on the true plant contour information from multi-view colour images, achieving good denoising effects for both outliers and anomalies. Additionally, it avoided misclassifying sparse plant point clouds within the contour as noise, resulting in point cloud models with more realistic structural information. Our testing on multiple objects (Fig. 12 (b)(c)) demonstrated the efficacy and robustness of this denoising method.

The typical error of RGB-D cameras within 1m was less than or equal to 5 mm, and this error could lead to either an overestimation or underestimation of stem thickness in plant models compared to reality. We found in our experiments that this error had a significant impact on the reconstruction of seedling stems, not only altering stem thickness but also causing changes in the spatial structure of the stems, resulting in stem distortion. After filtering with denoising methods, points representing stem portions that did not conform to the true spatial structure of plants were removed as noise, leading to incomplete point cloud models of stem portions. Liu et al. (2023) applied statistical filtering to denoise the reconstructed peanut model, resulting in stem point cloud loss. Similarly, Xie et al. (2023) reported significant stem loss in the reconstruction results of perilla and tomato. Additionally, seedling specimens exhibited simple textures and are challenging to extract feature points. Even when utilising SV and SFM to reconstruct 3D models, there remained an issue of missing stem, directly affecting the measurement of phenotypic data such as SD and PIA (Nguyen et al., 2015). Therefore, to reconstruct a complete plant model, it is necessary to use algorithms to fill in the missing areas of the point cloud model. Jiang et al. (2022) predicted and repaired stem point clouds based on the positions of the main stem and petiole without using real plant structural information. Their method can be applied to the repair of some ornamental point clouds but is not suitable for phenotype measurements. Lu et al. (2020) employed a method of view selection to actively choose viewpoints with less occlusion and a large amount of plant structural information for point cloud reconstruction, thereby improving the completeness of the point cloud at its source. However, it could not address the issue of point cloud loss caused by filtering. The proposed small-size stem point cloud

completion algorithm utilised the real contour information of plants from colour images to locate and repair missing areas in the point cloud, ensuring the authenticity of the repaired stem point cloud. This algorithm not only addressed point cloud loss during the acquisition stage but also resolved the issue of point cloud loss caused by filtering. Moreover, our method only required three inputs: the 3D model in need of repair, multi-view plant masks, and the camera pose relationships between different views. Therefore, this algorithm could be extended to address the issue of stem segment loss in 3D models obtained using multi-view 3D reconstruction methods.

#### 4.3. Phenotypic measurement

The projection point-contour distances ( $D_p$  and  $D_c$ ) of the 3D reconstruction results in this paper are  $1.9749 \pm 0.8803$  pixels and  $0.3185 \pm 0.1722$  mm, respectively. These values indicate that the projected contour of the 3D model from various viewpoints aligns closely with the plant contour in the colour images, with an error of less than 2 pixels in the images. In real scale, the error is less than 1 mm. This implies that the 3D model can effectively replace the real plant for phenotype measurements. Subsequently, we compared the phenotypic measurement results of the 3D model with manually destructive measurements, calculating the errors (Plant height:  $R^2 = 0.98$ , RMSE = 2.3 mm, rRMSE = 1.52%; Petioles inclination angle:  $R^2 = 0.99$ , RMSE = 0.73°, rRMSE = 1.41%; Leaf area:  $R^2 = 0.66$ , RMSE = 1.05 cm<sup>2</sup>, rRMSE = 7.63%; Leaf inclination angle:  $R^2 = 0.99$ , RMSE = 1.01°, rRMSE = 1.72%; Stem diameter:  $R^2 = 0.95$ , RMSE = 0.12 mm, rRMSE = 5.43%). The relatively lower  $R^2$  value for LA was because the reconstructed leaf 3D model constrained by the 2D contour of the leaf was generally smaller than the actual leaf. The rRMSE  $\leq 10\%$  suggested that the reconstruction outcomes in this study had a certain reference value for plant phenotype measurements. Finally, comparison of phenotypic measurement between our method and recently reported works is investigated. However, due to the different plant species and growth stages study by various researchers, as well as variations in the 3D reconstruction systems and equipment used, and different evaluation metrics for measurement accuracy, a comprehensive summary of recent researchers' phenotype measurement results, 3D reconstruction methods, and equipment used is provided (Table 2). All the researchers listed in Table 2 conducts their reconstructions in indoor environments, focusing on potted plants. To compare measurement errors for plants of different species and growth stages, MAPE is employed as an evaluation metric, as it can normalise the measurement errors. For researchers who do not report MAPE values, this study calculates MAPE based on the measurement data provides in their papers. Rossi et al. (2022) presents the most comprehensive and accurate phenotype measurement results among the listed researchers. Their data is applied in the analysis of

water stress in tomato. Comparing directly with their results, the measurement accuracy of PH, PIA, LA, and LIA improved by 57.69%, 45.88%, 3.24%, and 2.30%, respectively. This was because our method used 2D image contour information to constrain the 3D model, improving the integrity and spatial distribution realism of the 3D model. However, since leaf shapes were more diverse compared to stems, the restoration of stem 3D model could not be directly applied to restoration of leaf 3D model, leading to incomplete contours in the 3D model edges of leaves and limited improvement in measurement accuracy.

#### 4.4. Limitations

There are some limitations for our proposed method. Image segmentation performance will affect the denoising and repair accuracy. In our study, 120 RGB images of 1920 × 1080 resolution in a uniformed lighting environment were collected and trained for segmentation model. The results indicate that the model can achieve automatic seedling mask segmentation, effectively replacing manual segmentation. This achievement lays the foundation for the efficient automated 3D reconstruction process in this paper, but the model can't be used for outdoor scene. Furthermore, our method relied on the 2D contour of seedlings to constrain the 3D model. If a seedling had a complex leaf structure leading to unclear contour segmentation in the 2D image, it could cause the contour constraints for these leaf point clouds to be ineffective. For instance, in the coloured image of a cucumber seedling (Fig. 12 a1), it can be observed that each leaf has numerous small cracks. These cracks, being intricate and difficult to segment, led to discrepancies between the 3D model of the reconstructed result and the actual seedling. Additionally, for plants in their growth stage with complex canopies, our method could not capture colour and depth information for the plant's non-visible surfaces. As a consequence, our method could not reconstruct completely 3D models of plants with intricate canopies. Similar issues were also present in other 3D reconstruction methods such as laser scanning, SFM, and SV. Some researchers had attempted to address this by generating obscured branches (Liu et al., 2021), achieving satisfactory visual outcomes. However, the generated branches exhibited significant discrepancies from the actual structure and were unsuitable for phenotype research. Therefore, achieving a complete and clear-contour 3D model reconstruction for densely canopied plants remained a challenge (Ma et al., 2019). The camera's field of view and the distance between the camera and the plant in the device can impact the reconstruction size and accuracy. In this study, the field of view of the consumer-grade depth camera was fixed. As the distance between the camera and the plant increased, the method could be applied to larger plants, but it might have led to a decrease in the ranging accuracy of the depth camera. This paper focuses on studying

**Table 2**  
Comparison of various camera-based 3D reconstruction systems for plants.

Study	Camera System	Camera View	Measures	Techniques	Accuracy (MAPE)
Thuy Tuong Nguyen et al., 2015	10 high-resolution SLR cameras organised into 5 stereo pairs; 2 structured lights	Multiview for full 3D reconstruction	Plant height, leaf width and length, internode distance	SV	Plant height error: 8.1%; Leaf width error: 3.76%; Leaf length error: 4.87%; Internode distance error: 7.28%.
Xiaodan Ma et al., 2019	Kinect	Single view	Plant height, Canopy breadth	RGB-D	Plant height error: 5.1%; Canopy breadth error: 2.65%
Riccardo Rossi et al., 2022	Nikon® D810 36.3-mega-pixel digital camera	Multiview for full 3D reconstruction	Plant height, petiole inclination, single-leaf area, total leaf area, single-leaf angle	SFM	Plant height error: 3.12%; Petiole inclination error: 3.64%; Single-leaf area error: 7.4%; Single-leaf angle error: 2.17%.
Yadong Liu et al., 2023	Kinect v2	Multiview for full 3D reconstruction	Plant height, width, length and Volume	RGB-D	Plant height error: 4.12%; Plant width error: 4.83%; Plant length Error: 10.75%; Plant Volume Error: 10.83%.
Our	Realsense L515 LiDAR camera	Multiview for full 3D reconstruction	Plant height, petiole inclination, single-leaf area, single-leaf angle, stem cross-sectional area	RGB-D	Plant height error: 1.32%; Petiole inclination error: 1.97%; Single-leaf area error: 7.16%; Single-leaf inclination error: 2.12%; Stem cross-sectional area error: 4.68%.

the phenotypes of seedling-stage plants under arbitrary growth conditions. Therefore, the distance between the depth camera and the plant is set to be 25 cm to obtain high measurement accuracy. Using equation (1), the maximum height of the plant that could be captured was calculated to be 26 cm, and the maximum width was 35 cm. However, some branches and leaves of the plant might have drooped, causing some leaves to extend beyond the imaging range of the camera from certain angles, resulting in incomplete plant images that do not reach the theoretical values, as shown in the imaging result of the fifth plant in Fig. 11. Therefore, the object size is restricted to be less than 20 cm and 25 cm for height and width, respectively.

## 5. Conclusion

This study proposes an accurate multi-view 3D reconstruction for seedling based on projected contour consistency constraint. The results demonstrate an improvement in accuracy compared to conventional stitching methods involving coarse stitching and ICP registration. Moreover, once the rotation axis calibration is completed, subsequent multi-view point clouds can be stitched using the calculated camera pose matrices from the rotation axis, eliminating the need for ICP registration and thus enhancing the efficiency of 3D reconstruction. Additionally, our point cloud denoising algorithm leverages colour images containing more precise plant structural information to constrain the point cloud. This constraint-based approach effectively removes noise and outliers not adhering to the constraints, resulting in a more realistic appearance of the seedling's 3D structure and subsequently improving subsequent phenotype measurement accuracy. Then, unlike traditional methods such as statistical filtering and spatial filtering, our approach eliminates the need for iterative experimentation to find effective parameters. However, due to inherent accuracy limitations of depth cameras, our seedling reconstruction process reveals distortions in the 3D models of small-size stems, deviating significantly from the true seedling structure. After point cloud denoising, the stem 3D model becomes incomplete, consequently resulting in the loss of parameters such as stem diameter and petiole inclination angle. To address this, a small-size stem point cloud completion algorithm is investigated to restore the incomplete stem point cloud and achieve a complete 3D model.

Our goal is to establish a fully automated 3D seedling phenotyping and measurement experimental platform, providing phenotyping researchers with efficient tools. However, the current state of automated phenotype measurement algorithms based on 3D models requires further refinement. Therefore, future efforts will focus on enhancing automated phenotype measurement algorithms to elevate the level of automation in high-throughput phenotype analysis. Moreover, by incorporating cameras with various pose, the reconstruction of plants with more intricate canopy structures can be achieved by capturing a more comprehensive plant contour.

## Funding

This research was supported by Pioneer and Leading Goose R&D Program of Zhejiang (2022C02026), Zhejiang Provincial Natural Science Foundation of China under Grant No. (LTGY23F030001, LY23F010008), National Key R&D Program of China (2022YFD2000100).

## Data availability

The datasets used and/or analysed during the current study are available from the corresponding author upon request.

## CRediT authorship contribution statement

**Qingguang Chen:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition,

Conceptualization. **Shentao Huang:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Shuang Liu:** Software, Investigation. **Mingwei Zhong:** Software, Methodology. **Guohao Zhang:** Investigation, Data curation. **Liang Song:** Resources, Methodology, Investigation. **Xinghao Zhang:** Resources, Methodology, Data curation. **Jingcheng Zhang:** Supervision, Software, Funding acquisition. **Kaihua Wu:** Methodology, Investigation, Conceptualization. **Ziran Ye:** Software, Resources. **Dedong Kong:** Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.biosystemseng.2024.05.011>.

## References

- Bao, Y., Tang, L., Breitzman, M. W., Salas, F. M. G., & Schnable, P. S. (2019). Field-based robotic phenotyping of sorghum plant architecture using stereo vision. *Journal of Field Robotics*, 36(2), 397–415. <https://doi.org/10.1002/rob.21830>
- Despotovic, M., Nedic, V., Despotovic, D., & Cvetanovic, S. (2016). Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. *Renewable and Sustainable Energy Reviews*, 56, 246–260. <https://doi.org/10.1016/j.rser.2015.11.058>
- Forero, M. G., Murcia, H. F., Méndez, D., & Betancourt-Lozano, J. (2022). LiDAR platform for acquisition of 3D plant phenotyping database. *Plants*, 11(17), 2199. <https://doi.org/10.3390/plants11172199>
- Gibbs, J. A., Pound, M., French, A. P., Wells, D. M., Murchie, E., & Pridmore, T. (2018). Plant phenotyping: An active vision cell for three-dimensional plant shoot reconstruction. *Plant physiology*, 178(2), 524–534. <https://doi.org/10.1104/pp.18.00664>
- Harandi, N., Vandenberge, B., Vankerschaver, J., Depuydt, S., & Messem, A. V. (2023). How to make sense of 3D representations for plant phenotyping: A compendium of processing and analysis techniques. *Plant Methods*, 19, 60. <https://doi.org/10.1186/s13007-023-01031-z>
- Jay, S., Rabatel, G., Hadoux, X., Moura, D., & Gorretta, N. (2015). In-field crop row phenotyping from 3D modeling performed using Structure from Motion. *Computers and Electronics in Agriculture*, 110, 70–77. <https://doi.org/10.1016/j.compag.2014.09.021>
- Jiang, J., Li, L., & Wang, M. (2022). Research on plant stem complement based on L1-medial skeletal extraction. *Journal of Nanjing Forestry University (Natural Sciences Edition)*, 46(1), 40. <https://doi.org/10.12302/j.issn.1000-2006.202110040>
- Jin, S., Sun, X., Wu, F., Su, Y., Li, Y., Song, S., Xu, K., Ma, Q., Baret, F., Jiang, D., Ding, Y., & Guo, Q. (2021). Lidar sheds new light on plant phenomics for plant breeding and management: Recent advances and future prospects. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171, 202–223. <https://doi.org/10.1016/j.isprsjprs.2020.11.006>
- Lao, C., Yang, H., Li, P., & Feng, Y. (2019). 3D reconstruction of maize plants based on consumer depth camera. *Transactions of the Chinese Society for Agricultural Machinery*, 50(7), 222–228. <https://doi.org/10.6041/j.issn.1000-1298.2019.07.024>
- Li, Z., Guo, R., Li, M., Chen, Y., & Li, G. (2020a). A review of computer vision technologies for plant phenotyping. *Computers and Electronics in Agriculture*, 176, Article 105672. <https://doi.org/10.1016/j.compag.2020.105672>
- Li, M., Shao, M. R., Zeng, D., Ju, T., Kellogg, E. A., & Topp, C. N. (2020b). Comprehensive 3D phenotyping reveals continuous morphological variation across genetically diverse sorghum inflorescences. *New Phytologist*, 226(6), 1873–1885. <https://doi.org/10.1111/nph.16533>
- Li, H., Shen, H., & Cheng, Y. (2012). 3D reconstruction method based on turntable multiple-view registration. *Journal of Computer Applications*, 32(12), 3365. <http://www.joca.cn/EN/10.3724/SP.J.1087.2012.03365>.
- Li, Y., Su, Y., Hu, T., Xu, G., & Guo, Q. (2018). Retrieving 2-D Leaf angle distributions for deciduous trees from terrestrial laser scanner data. *IEEE Transactions on Geoscience and Remote Sensing*, 56, 4945–4955. <https://doi.org/10.1109/TGRS.2018.2843382>
- Li, Y., Wu, X., Xu, W., Sun, Y., Wang, Y., Li, G., & Xu, P. (2021). High-throughput physiology-based stress response phenotyping: Advantages, applications and prospective in horticultural plants. *Horticultural Plant Journal*, 7(3), 181–187. <https://doi.org/10.1016/j.hpj.2020.09.004>
- Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., Liu, G., & Jin, L. (2020c). Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, 161–172. <https://doi.org/10.1016/j.isprsjprs.2020.02.013>

- Liang, X., Zhou, F., Chen, H., Liang, B., Xu, X., & Yang, W. (2020). Three-dimensional maize plants reconstruction and traits extraction based on structure from motion. *Transactions of the Chinese Society of Agricultural Engineering*, 51(6), 209–219.
- Liu, Z., Wu, K., Guo, J., Wang, Y., Deussen, O., & Cheng, Z. (2021). Single image tree reconstruction via adversarial network. *Graphical Models*, 117, Article 101115. <https://doi.org/10.1016/j.gmod.2021.101115>
- Liu, Y., Yuan, H., Zhao, X., Fan, C., & Cheng, M. (2023). Fast reconstruction method of three-dimension model based on dual RGB-D cameras for peanut plant. *Plant Methods*, 19(1), 17. <https://doi.org/10.1186/s13007-023-00998-z>
- Lu, X., Ono, E., Lu, S., Zhang, Y., Teng, P., Aono, M., Shimizu, Y., Hosoi, F., & Omasa, K. (2020). Reconstruction method and optimum range of camera-shooting angle for 3D plant modeling using a multi-camera photography system. *Plant Methods*, 16(1), 1–16. <https://doi.org/10.1186/s13007-020-00658-6>
- Ma, X., Wei, B., Guan, H., & Yu, S. (2022). A method of calculating phenotypic traits for soybean canopies based on three-dimensional point cloud. *Ecological Informatics*, 68, Article 101524. <https://doi.org/10.1016/j.ecoinf.2021.101524>
- Ma, X., Zhu, K., Guan, H., Feng, J., Yu, S., & Liu, G. (2019). High-throughput phenotyping analysis of potted soybean plants using colorized depth images based on a proximal platform. *Remote Sensing*, 11(9), 1085. <https://doi.org/10.3390/rs11091085>
- Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2021). Nerf: Representing scenes as neural radiance fields for view synthesis. *Communications of the ACM*, 65(1), 99–106. <https://doi.org/10.1145/3503250>
- Nguyen, T. T., Slaughter, D. C., Maloof, J. N., & Sinha, N. (2016). Plant phenotyping using multi-view stereo vision with structured lights. *Autonomous SPIE Commercial + Scientific Sensing and Imaging*, 1–9. <https://doi.org/10.1117/12.2229513>. SPIE, Baltimore, MA, USA.
- Nguyen, T. T., Slaughter, D. C., Max, N., Maloof, J. N., & Sinha, N. (2015). Structured light-based 3D reconstruction system for plants. *Sensors*, 15(8), 18587–18612. <https://doi.org/10.3390/s150818587>
- Noshita, K., Murata, H., & Kirie, S. (2022). Model-based plant phenomics on morphological traits using morphometric descriptors. *Breeding Science*, 72(1), 19–30. <https://doi.org/10.1270/jbsbs.21078>
- Piovesan, A., Vancauwenbergh, V., Van De Looverbosch, T., Verboven, P., & Nicolaï, B. (2021). X-ray computed tomography for 3D plant imaging. *Trends in Plant Science*, 26 (11), 1171–1185. <https://doi.org/10.1016/j.tplants.2021.07.010>
- Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., & Jagersand, M. (2020). U2-Net: Going deeper with nested U-structure for salient object detection. *Pattern Recognition*, 106, Article 107404. <https://doi.org/10.1016/j.patcog.2020.107404>
- Qiu, R., Miao, Y., Zhang, M., & Li, H. (2021). Detection of the 3D temperature characteristics of maize under water stress using thermal and RGB-D cameras. *Computers and Electronics in Agriculture*, 191, Article 106551. <https://doi.org/10.1016/j.compag.2021.106551>
- Qiu, R., Zhang, M., & He, Y. (2022). Field estimation of maize plant height at jointing stage using an RGB-D camera. *The Crop Journal*, 10(5), 1274–1283. <https://doi.org/10.1016/j.cj.2022.07.010>
- Rodrigues, O. (1840). Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire. *Journal de Mathématiques Pures et Appliquées*, 5, 380–440. [http://www.numdam.org/item?id=JMPA\\_1840\\_1\\_5\\_380\\_0](http://www.numdam.org/item?id=JMPA_1840_1_5_380_0).
- Rossi, R., Costafreda-Aumedes, S., Leolini, L., Leolini, C., Bindi, M., Zaldei, A., & Moriondo, M. (2022). Implementation of an algorithm for automated phenotyping through plant 3D-modeling: A practical application on the early detection of water stress. *Computers and Electronics in Agriculture*, 197, Article 106937. <https://doi.org/10.1016/j.compag.2022.106937>
- Rossi, R., Leolini, C., Costafreda-Aumedes, S., Leolini, L., Bindi, M., Zaldei, A., & Moriondo, M. (2020). Performances evaluation of a low-cost platform for high-resolution plant phenotyping. *Sensors*, 20(11), 3150. <https://doi.org/10.3390/s20113150>
- Simbeye, D. S., Mkiramweni, M. E., Karaman, B., & Taskin, S. (2023). Plant water stress monitoring and control system. *Smart Agricultural Technology*, 3, Article 100066.
- Sultan, S. E. (2003). Phenotypic plasticity in plants: A case study in ecological development. *Evolution and Development*, 5(1), 25–33. <https://doi.org/10.1046/j.1525-142X.2003.03005.x>
- Sun, G., & Wang, X. (2019). Three-dimensional point cloud reconstruction and morphology measurement method for greenhouse plants based on the kinect sensor self-calibration. *Agronomy*, 9(10), 596. <https://doi.org/10.3390/agronomy9100596>
- Syed, T. N., Jizhan, L., Zhou, X., Zhao, S., Yan, Y., Mohamed, S. H. A., & Lakhiar, I. A. (2019). Seedling-lump integrated non-destructive monitoring for automatic transplanting with Intel RealSense depth camera. *Artificial Intelligence in Agriculture*, 3, 18–32. <https://doi.org/10.1016/j.aiia.2019.09.001>
- Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., & Bennett, M. (2017). Plant phenomics, from sens-or-s to knowledge. *Current Biology*, 27(15), 770–783. <https://doi.org/10.1016/j.cub.2017.05.055>
- Teng, X., Zhou, G., Wu, Y., Huang, C., Dong, W., & Xu, S. (2021). Three-dimensional reconstruction method of rapeseed plants in the whole growth period using RGB-D camera. *Sensors*, 21(14), 4628. <https://doi.org/10.3390/s21144628>
- Wang, J., Zhang, Y., & Gu, R. (2020). Research status and prospects on plant canopy structure measurement using visual sensors based on three-dimensional reconstruction. *Agriculture*, 10(10), 462. <https://doi.org/10.3390/agriculture10100462>
- Xie, P., Du, R., Ma, Z., & Cen, H. (2023). Generating 3D multispectral point clouds of plants with fusion of snapshot spectral and RGB-D images. *Plant Phenomics*, 5, 40. <https://doi.org/10.34133/plantphenomics.0040>
- Zhou, J., Fu, X., Schumacher, L., & Zhou, J. (2018). Evaluating geometric measurement accuracy based on 3d reconstruction of automated imagery in a greenhouse. *Sensors*, 18(7), 2270. <https://doi.org/10.3390/s18072270>