Three-Dimensional Point Cloud Reconstruction and Morphology Measurement Method for Greenhouse Plants Based on the Kinect Sensor Self-Calibration

Guoxiang Sun [1,2,*] and **Xiaochan Wang** [1,2]

1   College of Engineering, Nanjing Agricultural University, Nanjing 210095, China; wangxiaochan@njau.edu.cn
2   Jiangsu Province Engineering Lab for Modern Facility Agriculture Technology & Equipment,
     Nanjing 210031, China
*   Correspondence: sguoxiang@njau.edu.cn; Tel.: +86-255-860-6585

Abstract: Plant morphological data are an important basis for precision agriculture and plant phenomics. The three-dimensional (3D) geometric shape of plants is complex, and the 3D morphology of a plant changes relatively significantly during the full growth cycle. In order to make high-throughput measurements of the 3D morphological data of greenhouse plants, it is necessary to frequently adjust the relative position between the sensor and the plant. Therefore, it is necessary to frequently adjust the Kinect sensor position and consequently recalibrate the Kinect sensor during the full growth cycle of the plant, which significantly increases the tedium of the multiview 3D point cloud reconstruction process. A high-throughput 3D rapid greenhouse plant point cloud reconstruction method based on autonomous Kinect v2 sensor position calibration is proposed for 3D phenotyping greenhouse plants. Two red–green–blue–depth (RGB-D) images of the turntable surface are acquired by the Kinect v2 sensor. The central point and normal vector of the axis of rotation of the turntable are calculated automatically. The coordinate systems of RGB-D images captured at various view angles are unified based on the central point and normal vector of the axis of the turntable to achieve coarse registration. Then, the iterative closest point algorithm is used to perform multiview point cloud precise registration, thereby achieving rapid 3D point cloud reconstruction of the greenhouse plant. The greenhouse tomato plants were selected as measurement objects in this study. Research results show that the proposed 3D point cloud reconstruction method was highly accurate and stable in performance, and can be used to reconstruct 3D point clouds for high-throughput plant phenotyping analysis and to extract the morphological parameters of plants.

Keywords: three-dimensional reconstruction; point cloud analysis; Kinect; RGB-D; greenhouse plants; high-throughput phenotyping

1. Introduction

The phenotype of a plant is determined or affected by genetic and environmental factors, and the structure, composition, and physical, physiological, and biochemical traits and properties of the plant reflect these factors during growth and development and at maturity [1]. Plant phenotypic data are an important basis for analyzing the relationship between genotype, environment, and phenotype. Plant phenotyping techniques severely lag behind research needs and have become a bottleneck that limits the development of molecular crop breeding and functional plant genomics [2–4]. With the rapid development of sensor and spectral imaging technologies, automated plant phenotyping can now be achieved by computer graphic and image processing. Thus, studying this technique is of great significance to the realization of the high-throughput, precise, and automated phenotyping of greenhouse plants [5,6].

Greenhouse environments are controllable. High-throughput greenhouse plant phenotyping systems are key instruments for analyzing the relationship between genes, the environment, and phenotypes, and are important tools for accelerating the development of molecular crop breeding and functional plant genomics. Extensive research has been conducted to investigate the three-dimensional (3D) geometric morphological phenotyping of greenhouse plants. Sensor technologies, including monocular vision, stereoscopic vision, depth cameras, laser scanning, two-dimensional (2D)/3D laser lidar, X-ray computed tomography (CT), and magnetic resonance imaging (MRI), have primarily been used to measure the 3D geometric morphologies of plants.

Monocular vision requires fixed measurement conditions in conjunction with a measurement model, and it can only achieve limited 3D morphological plant phenotyping [7,8]. Thus, monocular vision has relatively low applicability. Multiview reconstruction techniques based on monocular vision mainly include space carving [9], visual structure from motion system techniques [10,11], multiview photogrammetry [12,13], and multicamera synchronous reconstruction [14]. However, these techniques require many angles of view (AOVs) for measurement and are unable to meet the reconstruction efficiency requirement for high-throughput plant phenotyping. Line laser [15,16] and 2D light detection and ranging scanning have also been used to reconstruct 3D point clouds of plants [17]. However, these techniques are unable to meet the requirement of high-throughput plant phenotyping. X-ray CT and MRI can be used to achieve single-view reconstruction of 3D plant structures and are mainly used to phenotype plant root systems in soil [18]. Stereoscopic vision [19,20], depth camera (Kinect sensor or TOF camera) [21–25], and 3D laser lidar [26–28] sensors can only capture two-and-a-half-dimensional (2.5D) depth images at a single angles of view (AOV). Kinect sensor-based 3D plant reconstruction can be divided into single-view [29–31] and multiview reconstruction [32–34], the latter mainly using the iterative closest point (ICP) algorithm [24,25]. However, we first need to solve the problem regarding the rough registration of multiview point clouds, otherwise ICP cannot be used for accurate registration. Plants are 3D nonrigid structures with complex 3D geometric morphologies and sheltered leaves and fruits. As a result, true 3D images of plants cannot be captured, and the 3D reconstruction of plants cannot be achieved at a single AOV. Instead, multiview stereo reconstruction is required for plants [35–39]. Therefore, rapid multiview registration is key to achieving the high-throughput 3D phenotyping of greenhouse plants.

In this study, a high-throughput 3D rapid greenhouse plant point cloud reconstruction method based on autonomous Kinect v2 sensor position calibration is proposed for the 3D phenotyping of greenhouse plants. This method mainly addresses the issue of rapid multiview 3D point cloud registration. Two calibration labels are affixed to the single-axis precision turntable. Two red–green–blue–depth (RGB-D) images are acquired by a Kinect v2 sensor. The central point and normal vector of the axis of rotation of the turntable are calculated automatically. The coordinate systems of RGB-D images captured at various AOVs are unified based on the central point and normal vector of the axis of the turntable to achieve coarse registration. Then, the ICP algorithm is used to perform fine multiview point cloud registration [24,25,40], thereby achieving rapid 3D point cloud reconstruction of the greenhouse plant. Greenhouse tomato plants (GTPs) were selected as the measurement objects in this study. The accuracy of 3D point cloud reconstruction was quantitatively analyzed based on the Hausdorff distance between the reconstructed and manually measured point clouds. In addition, the errors in the 3D morphological reconstruction of the measurement objects were also analyzed. This study provides theoretical and technical support for the high-throughput 3D morphological measurement of greenhouse plants.

2. Materials and Methods

2.1. Structure and Principle of the Measurement System

The high-throughput greenhouse plant phenotyping system consists mainly of an imaging chamber, a light-emitting diode (LED) light source, a precision turntable, a Kinect v2 sensor, a dual-axis

slider, a controller, and a graphic workstation. The imaging chamber is made of aluminum section bars and has internal dimensions of 180 (length) × 120 (width) × 160 cm (height). The internal sides and bottom of the imaging chamber are covered with matte white films. The LED light source consists of two Philips LED bulbs (6400 K, 70 W). The precision turntable has dimensions of 31.4 cm (length) × 21.4 cm (width) × 6 cm (height), a disk diameter of 20 cm, a rotation range of 360°, and a transmission ratio of 180:1. The turntable operates via a turbine and worm driving mechanism and is driven by a 57BYG stepper motor. In addition, the turntable has a resolution of 0.0005° and a positioning accuracy of 0.01°. The phenotyping system is equipped with a Kinect v2 sensor, which consists of a color camera and a depth sensor. In addition, the Kinect v2 sensor captures RGB images at 1920 px × 1080 px and 30 frames per second (fps), depth images at 512 px × 424 px and 30 fps, and IR images at 512 px × 424 px. Moreover, the Kinect v2 sensor has a horizontal AOV of 70°, a vertical AOV of 60°, and the detection range was 0.50–4.50 m. The dual-axis slider, driven by a 57BYG stepper motor, has a positioning accuracy of 0.1 mm, a bearing capacity of 100 N, a horizontal range of 80 cm, and a vertical range of 80 cm. In the phenotyping system, an HW-36MT-3PG programmable logic controller is used, which consists of 36 input/output channels, of which 20 are input channels and 16 are output channels. The controller has three built-in PG speed-up/speed-down high-speed pulse output channels with a maximum frequency of 100 kHz and an RS232C communication interface. The graphic workstation has an Intel Xeon E5-1620 V4 4C/3.5 GHz processor, a Windows 10 Professional System, 32 GB of error-correcting code random-access memory, and an NVIDIA Quadro P2000 5G graphics card. There is a hybrid programming software environment on the workstation, consisting of Visual Studio 2015 and MATLAB 2017a. High-throughput greenhouse plant phenotyping system software was developed using the Kinect v2 Software Developer Kit and C + + wrapper functions for the Microsoft Kinect v2.

Figure 1a,b respectively show a color image and a depth image of a plant captured by the Kinect v2 sensor. Figure 1c shows a photograph of the precision turntable at the bottom of the imaging chamber. The turntable surface is covered by a blue film, to which two circular calibration labels with a diameter of 5 cm (one yellow and one red) are affixed.
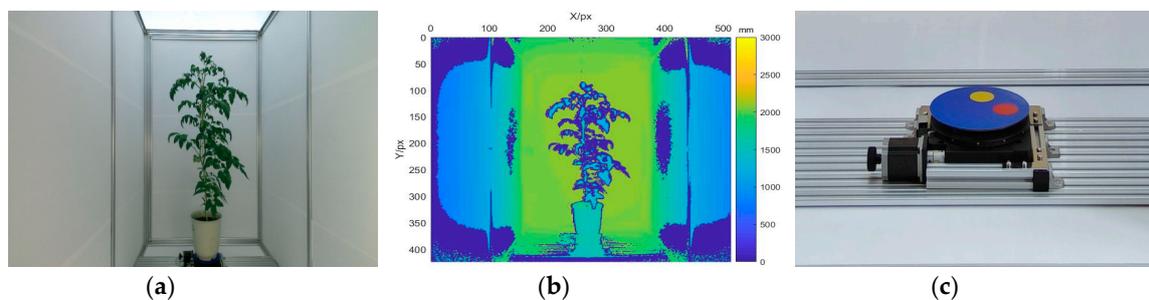

(**a**)   (**b**)   (**c**)

Figure 1. Images captured in the imaging chamber. (**a**) color image of a plant, (**b**) depth image of a plant, (**c**) precision turntable.

The working principle of the high-throughput greenhouse plant phenotyping system is as follows. Step 1: Initialization parameters are set, which mainly include the Kinect sensor position (the dual-axis slider initializes the measuring height and distance of the Kinect sensor in the Y-axis and Z-axis directions, respectively); the internal parameters of the Kinect sensor including the main point coordinates (c_x, c_y) and focal length (f_x, f_y); the region of interest bounding box including the horizontal X-axis, vertical Y-axis, and depth Z-axis; and the number of AOVs for 3D reconstruction of the plant (V_N). Step 2: Kinect sensor self-calibration: (i) an initial RGB-D image of the turntable is first captured, and then another RGB-D image is captured after rotating the turntable by 180°; (ii) an RGB-D image is converted to a 3D point cloud; (iii) the coordinates of the point clouds of the yellow and red calibration labels on the turntable surface are identified at the two AOVs; (iv) the barycenter of each calibration label is calculated; (v) the coordinates of the central point and the normal vector of the axis

of rotation of the turntable are calculated. Step 3: 3D reconstruction of the GTPs: (i) RGB-D images of the plant area captured at multiple AOVs based on V_N; RGB-D images of the plant are captured at various AOVs one by one; after an image is captured at each AOV, the precision turntable is rotated by $360°/V_N$. (ii) Based on the internal parameters of the Kinect sensor, the RGB-D images captured at various AOVs are converted to a 3D point cloud. (iii) The 3D point cloud is then subjected to a bounding box treatment (the selected plant area) and an outlier removal treatment (iv) Based on the coordinates of the central point of the turntable's axis of rotation, a displacement transformation is performed on the point cloud at each AOV; the central point of the turntable's axis of rotation is moved to the origin (0, 0, 0) of the coordinate system of the Kinect sensor. (v) Based on the normal vector of the turntable's axis of rotation, a rotation transformation is performed on the point cloud at each AOV, and thus the coordinate systems of the point clouds at various AOVs are unified. (vi) The point clouds at various AOVs are registered sequentially using the ICP algorithm. (vii) The 3D point cloud is down-sampled. Finally, a 3D point cloud is reconstructed for the GTP. Step 4: The morphological parameters of the GTPs are calculated: (i) the characteristic parameters of the GTPs including the height (H), the maximum width (W), the point cloud number (N_P), and the area of the canopy projected in the horizontal plane (S_{XOZ}) are calculated; (ii) according to the calibration model, the characteristic parameters of the GTPs are calculated, including the canopy volume (V) and fresh weight (FW).

2.2. Autonomous Calibration of the Kinect Sensor Position

The 3D morphology of a plant changes relatively significantly during the full growth cycle. As a result, it is necessary to frequently adjust the Kinect sensor position and consequently recalibrate the Kinect sensor during the full growth cycle of the plant, which significantly increases the tedium of the multiview 3D point cloud reconstruction process. Rapid multiview 3D point cloud registration is key to achieving the 3D point cloud reconstruction of plants. In this study, the point clouds of plants at various AOVs are acquired by single-axis rotation. We propose an autonomous Kinect sensor position calibration method. This method can be used to achieve rapid, coarse multiview point cloud registration. In addition, the ICP algorithm is employed to achieve fine multiview point cloud registration. Thus, the rapid reconstruction of the 3D point clouds of greenhouse plants is achieved. This method meets the cyclic, high-throughput 3D morphological measurement requirement for plants, and significantly simplifies the 3D point cloud reconstruction process.

The autonomous Kinect sensor position calibration method is as follows. An initial RGB-D image of the turntable is first captured (at the 0° position), and another RGB-D image is captured after rotating the turntable by 180° (at the 180° position). The two RGB-D images are then each converted to a 3D point cloud, as shown in Figure 2a,e, respectively. Based on the point cloud color threshold, the coordinates of the point clouds of the red and yellow calibration label areas are determined. Figure 2b,c show the identified red and yellow calibration label areas, respectively, when the turntable is at the 0° position. Figure 2f,g show the identified red and yellow calibration label areas, respectively, when the turntable is at the 180° position. Figure 2d,h show rendered images depicting the identification of the calibration labels when the turntable is at the 0° and 180° positions, respectively. The coordinates of the barycenter of the point cloud of each calibration label are calculated. The coordinates of the barycenters of the point clouds of the red and yellow calibration labels are denoted by R_1 (x, y, z), Y_1 (x, y, z) and R_2 (x, y, z), Y_2 (x, y, z), respectively. On this basis, the coordinates of the central point of the optical axis of rotation of the turntable are calculated: M (a_0, b_0, c_0) = mean (R_1, R_2, Y_1, Y_2); in addition, the normal vector (cross($M - R_1$, $M - Y_1$)) of the axis of rotation of the turntable is also calculated. After normalization, the normal vector P of the optical axis is obtained: $P = (a, b, c)$, as shown in Figure 2i.
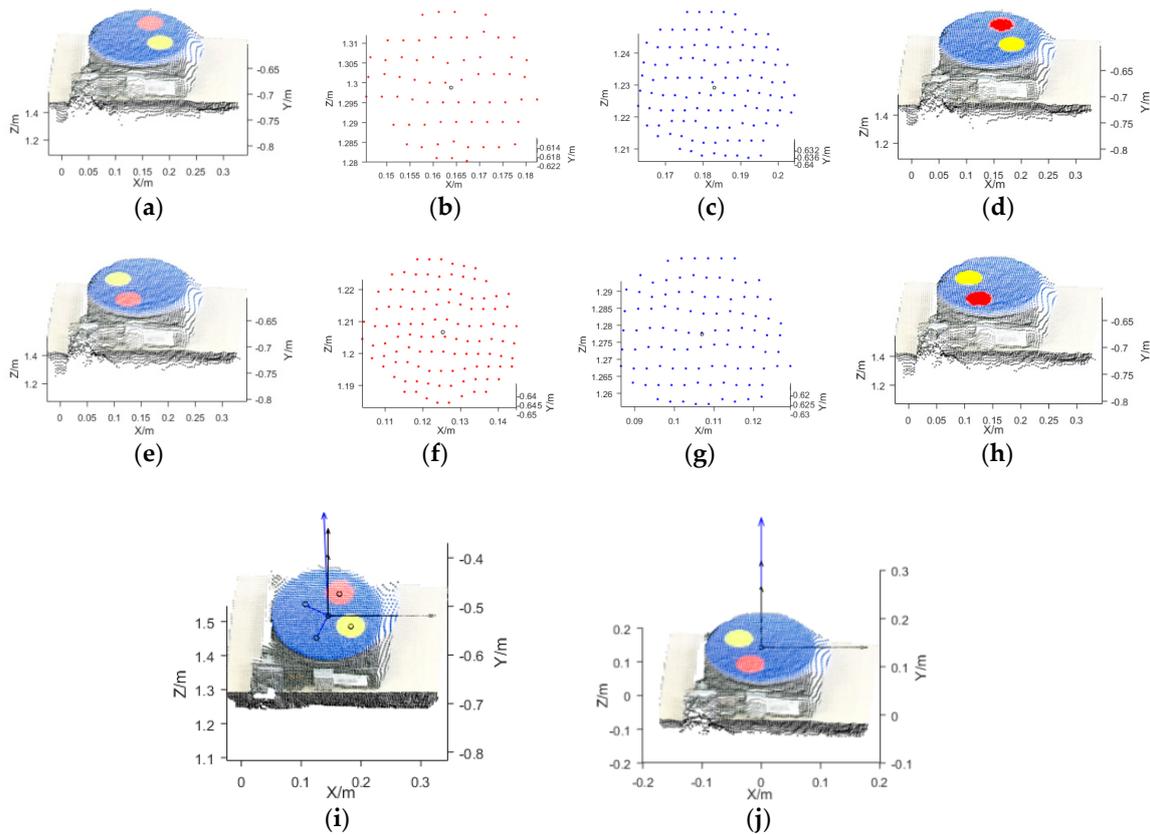
Figure 2. Calibration of the Kinect sensor position. (**a**) Segmentation of the point cloud of the turntable (0°); (**b**) Point cloud of the red calibration label (0°); (**c**) Point cloud of the yellow calibration label (0°); (**d**) Identification of the calibration labels (0°); (**e**) Segmentation of the point cloud of the turntable (180°); (**f**) Point cloud of the red calibration label (180°); (**g**) Point cloud of the yellow calibration label (180°); (**h**) Identification of the calibration labels (180°); (**i**) Normal vector of the axis of rotation of the turntable (original coordinates); (**j**) Normal vector of the axis of rotation of the turntable (after transformation).

Figure 3 shows the coordinate system of the Kinect sensor. Multiview point cloud coordinate systems are unified according to the following steps. Step 1: The point cloud at each AOV, $pointCloud_i$, is first translated by $M(a_0, b_0, c_0)$, and the central point of the axis of rotation is translated to the origin $(0, 0, 0)$ of the coordinate system of the Kinect sensor, as shown in Equation (1). Step 2: $pointCloud_i$ is rotated, and the normal vector of the axis of rotation is rotated to the Y-axis. Specifically, the normal vector $P(a, b, c)$ of the axis of rotation is first rotated around the X-axis by $-\alpha°$ to the XOY plane and then rotated around the Z-axis by $\beta°$ to the Y-axis, that is, the following operation is performed: $P \times R_x(-\alpha) \times R_z(\beta)$, as shown in Figure 3. $R_x(-\alpha)$ and $R_z(\beta)$ are then calculated using Equations (2) and (3), respectively. Then, based on the actual angle of the turntable at each AOV, the angle $\gamma°$ by which the axis of rotation is counterrotated is determined, and the counterrotation matrix $R_y(\gamma)$ is then determined as shown in Equation (4). The following rotation operation is performed on $pointCloud_i$: $pointCloud_i \times R_x(-\alpha) \times R_z(\beta) \times R_y(\gamma)$. The point cloud at the initial AOV, $pointCloud_1$, is the reference point cloud, which will not be counterrotated around the Y-axis. Figure 2i shows the coordinates of the initial point cloud of the turntable. Figure 2j shows the result obtained by moving the central point of the optical axis of the turntable to the origin of the coordinate system of the Kinect sensor and the normal vector of the axis of rotation of the turntable to the Y-axis.

$$pointCloud'_i(x_n, y_n, z_n) = pointCloud_i(x_n, y_n, z_n) - M(a_0, b_0, c_0) \tag{1}$$

$$R_x(-\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(-\alpha) & \sin(-\alpha) \\ 0 & -\sin(-\alpha) & \cos(-\alpha) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & b/\sqrt{b^2+c^2} & -c/\sqrt{b^2+c^2} \\ 0 & c/\sqrt{b^2+c^2} & b/\sqrt{b^2+c^2} \end{bmatrix} \tag{2}$$

$$R_z(\beta) = \begin{bmatrix} \cos\beta & \sin\beta & 0 \\ -\sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{b^2+c^2}/\sqrt{a^2+b^2+c^2} & a/\sqrt{a^2+b^2+c^2} & 0 \\ -a/\sqrt{a^2+b^2+c^2} & \sqrt{b^2+c^2}/\sqrt{a^2+b^2+c^2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

$$R_y(\gamma) = \begin{bmatrix} \cos\gamma & 0 & -\sin\gamma \\ 0 & 1 & 0 \\ \sin\gamma & 0 & \cos\gamma \end{bmatrix} \tag{4}$$

$$pointCloud''_i(x_n, y_n, z_n) = pointCloud'_i(x_n, y_n, z_n) \times R_x(-\alpha) \times R_z(\beta) \times R_y(\gamma) \tag{5}$$

Here, *pointCloud$_i$* is the point cloud at the ith AOV; (x_n, y_n, z_n) are the coordinates of the point cloud; n is the number of point clouds; $M(a_0, b_0, c_0)$ are the point coordinates of the turntable's axis of rotation; *pointCloud$'_i$* is the point cloud at the ith AOV that has been subjected to a displacement transformation; $R_x(-\alpha)$ is the rotation matrix by which the axis of rotation is rotated to the XOY plane; $R_z(\beta)$ is the rotation matrix by which the axis of rotation is rotated to the Y-axis; $R_y(\gamma)$ is the rotation matrix by which each AOV is rotated around the Y-axis with respect to the initial AOV; $P(a, b, c)$ is the normal vector of the axis of rotation; and *pointCloud$''_i$* is the point cloud at the ith AOV after the coordinate systems are unified.
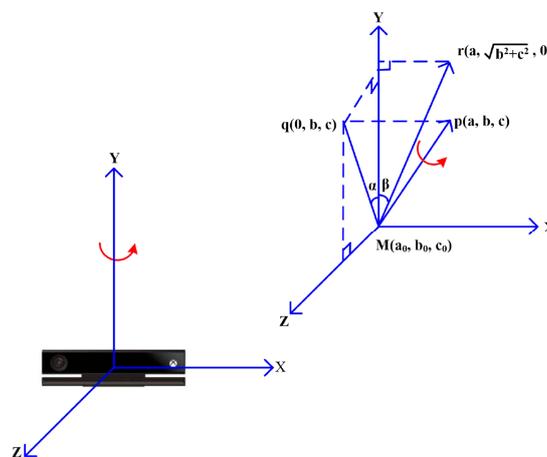

Figure 3. Schematic diagram of the unification of multiview point cloud coordinate systems.

2.3. Experimental and Data Analysis

To validate the proposed 3D point cloud reconstruction method, GTPs were selected as the measurement objects. Sixty GTPs cultivated from 15 March to 30 June 2019 were selected. These GTPs were 36.30–157.12 cm in height and 21.11–88.47 cm in width. The Kinect sensor captured 12 RGB-D images of each measurement object at AOV intervals of 30°. The *H*, *W*, *V*, and *FW* of each GTP were manually measured. The 3D point cloud reconstruction accuracy data for the GTPs were analyzed using the following method. The reference point cloud model of each GTP was scanned by an Artec Eva handheld 3D scanner (scanning accuracy 0.1 mm), and the 3D point cloud model was preprocessed by Artec studio software. Because the scanned point cloud model and the reconstructed point cloud model are in different 3D coordinate systems, they need to be aligned to the same 3D coordinate system before the reconstruction quality evaluation. In this study, the 3D scanning model was used as the target object, and the reconstruction model was used as the active object. The key feature points

were selected manually, and the coordinate systems of the two point cloud models were aligned. Based on the 3D scanned model, the reconstruction accuracy of the GTP point cloud models was analyzed. During the plant point cloud reconstruction experiment, the Kinect sensor was placed in 11 positions. In addition, different reconstruction views were combined, as shown in Table 1. The reconstruction method included three angles of view (V_3) and four angles of view (V_4), which was based on the reconstruction results of standard Styrofoam balls; the detailed analysis can be found in Appendix A. Statistical data included the distribution frequency of the set of distances (H_{RS}) between the reconstructed and reference point clouds of each GTP, the Hausdorff distance (HD) between the reconstructed and reference point clouds as shown in Equations (6–8), and the average (H_{avg}) of the HD set as shown in Equation (9), the standard deviation (SD) and coefficient of variation (CV) of the multiview measurement results including the H, W, N_P, S_{XOZ}, and V, as well as the coefficient of determination (R^2), root-mean-square error (RMSE), and the relative average deviation (RAD) of the manual measurement results of the H, W, V, and FW.

$$HD(A,B) = \max\{\max H(A,B), \max H(B,A)\} \tag{6}$$

$$H(A,B) = \underset{a \in A}{D}\left\{\underset{b \in B}{\min}\{d(a,b)\}\right\} \tag{7}$$

$$H(B,A) = \underset{b \in B}{D}\left\{\underset{a \in A}{\min}\{d(b,a)\}\right\} \tag{8}$$

$$H_{avg}(A,B) = \underset{a \in A}{avg}\left\{\underset{b \in B}{\min}\{d(a,b)\}\right\} \tag{9}$$

Table 1. Combination of different reconstruction views.

V_N	Abbreviation	AOV 1	AOV 2	AOV 3	AOV 4
V_3	V3-1	0°	120°	240°	
	V3-2	30°	150°	270°	
	V3-3	60°	180°	300°	
	V3-4	90°	210°	330°	
V_4	V4-1	0°	90°	180°	270°
	V4-2	30°	120°	210°	300°
	V4-3	60°	150°	240°	330°

Here, $HD(A, B)$ is the Hausdorff distance between point sets A and B; $H(A, B)$ is the set of distances of point set A relative to point set B; $H(B, A)$ is the set of distances of point set B relative to point set A; $H_{avg}(A, B)$ is the average of the set of distances of point set A relative to point set B; a is a point in point set A; and b is a point in point set B.

3. Results and Discussion

3.1. Reconstruction of GTP 3D Point Clouds

To examine the performance of the proposed 3D point cloud reconstruction method, 60 GTPs were selected as measurement objects. RGB-D images of each GTP were captured at AOV intervals of 30°. On this basis, a 3D point cloud of each GTP was reconstructed, and four morphological parameters, namely, H, W, V, and FW, were calculated. In addition, the H, W, V, and FW of each GTP were also manually measured. Moreover, 3D point clouds were reconstructed in a combination of different reconstruction views, as shown in Table 1. The errors in the morphological parameters were statistically calculated, and the accuracy of the point cloud reconstruction of the GTPs was analyzed.

As shown in Figure 4, a GTP 3D point cloud model was reconstructed from the RGB-D images of three views. The reconstruction process includes five main steps: point cloud preprocessing, point

cloud coarse registration, point cloud precise registration, plant canopy area selection, and point cloud down sampling. Point cloud preprocessing: First, based on the internal parameters of the Kinect sensor including the main point coordinates (c_x, c_y) and focal length (f_x, f_y), the RGB-D images at viewing angles of 0°, 120°, and 240° were converted into 3D point cloud images. The bounding box method was used to set the processing ranges of the horizontal X-axis, the vertical Y-axis, and the deep Z-axis, and the plant area was selected. The three-dimensional point cloud maps of 0°, 120°, and 240° are shown in Figure 4a–c, respectively. Point cloud coarse registration: According to the Kinect's self-calibrated rotation axis center coordinate M and rotation matrix of R_x and R_z, the 3D point cloud maps of 0°, 120°, and 240° were displaced, the center point coordinates were subtracted from all point cloud coordinates, and the point cloud coordinates after the displacement were multiplied by R_x and R_z. The 3D point cloud at a 0° viewing angle was selected as the reference coordinate system, and the point cloud coordinates at 120° and 240° were multiplied by R_y and reversely rotated around the Y axis by 120° and 240°. The 3D point cloud maps of 0°, 120°, and 240° after the normalization of the point cloud coordinate system are shown in Figure 4d–f, respectively. Point cloud precise registration: ICP registration was performed for the 0° point cloud and 120° point cloud, as shown in Figure 4g, and then the 0° and 120° registration results were subjected to the ICP registration with the 240° point to obtain the 3D point cloud map of GTP with the successful registration of the multiview point cloud, as shown in Figure 4h. Plant canopy area selection: Since the Y-axis coordinate of the bottom of the cultivation pot should be 0 after the displacement operation, according to the actual height of cultivation, a point with a point cloud Y value greater than the height of the cultivation pot can be used as the canopy area of the plant, as shown in Figure 4i; the RGB point cloud map is shown in Figure 4j. Point cloud down-sampling: Because there was a small amount of overlap in the multiview point cloud registration, the 3D point cloud was down-sampled by a 3D mesh filter, and the down-sampling result with a 3D mesh threshold of 5 mm is shown in Figure 4k.
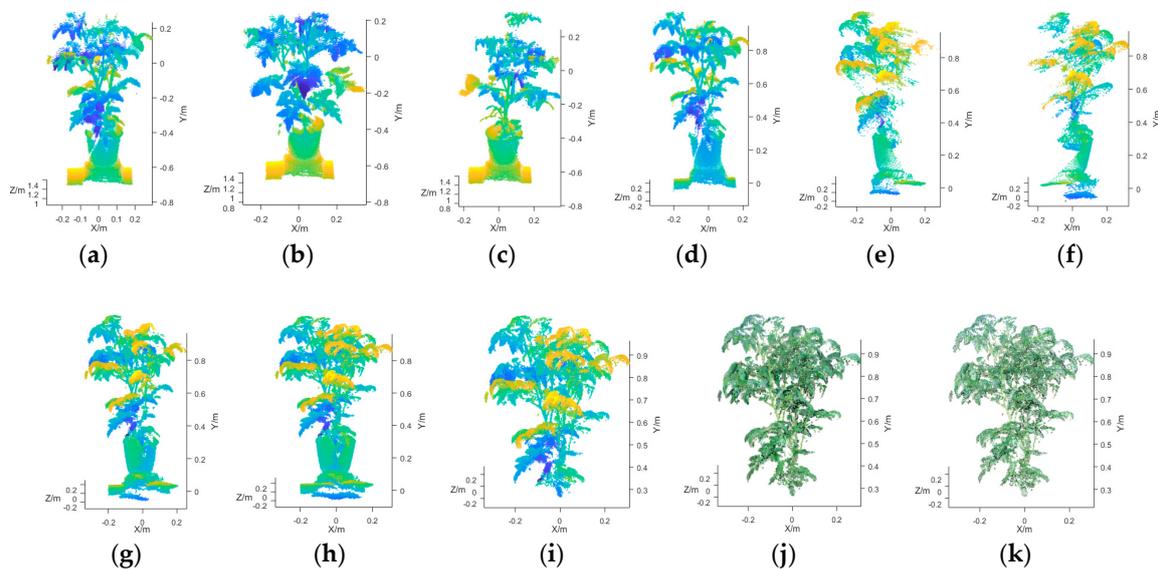

Figure 4. 3D point cloud reconstructions of greenhouse tomato plants (GTPs): (**a**) 0° point cloud; (**b**) 120° point cloud; (**c**) 240° point cloud; (**d**) 0° point cloud transformation; (**e**) 120° point cloud transformation; (**f**) 240° point cloud transformation; (**g**) iterative closest point (ICP) (0°, 120°); (**h**) ICP (0°, 120°, 240°); (**i**) canopy area (depth map); (**j**) canopy area (RGB); (**k**) point cloud down-sampling.

The multiview RGB-D images were used to reconstruct the 3D point cloud of the GTP, and the RGB-D images from different viewing angles around the Y-axis resulted in different coordinates of the reconstructed point cloud. The reconstruction results for the three views of V3-1, V3-2, V3-3, and V3-4 are shown in Figure 5a–d, with the point clouds of the three views represented by red, yellow, and blue. The reconstruction results for the four views angles of V4-1, V4-2, and V4-3 are shown in Figure 5e–g,

with the point clouds of the four views represented by red, yellow, blue, and gray. The point cloud depth map is shown in Figure 5h.
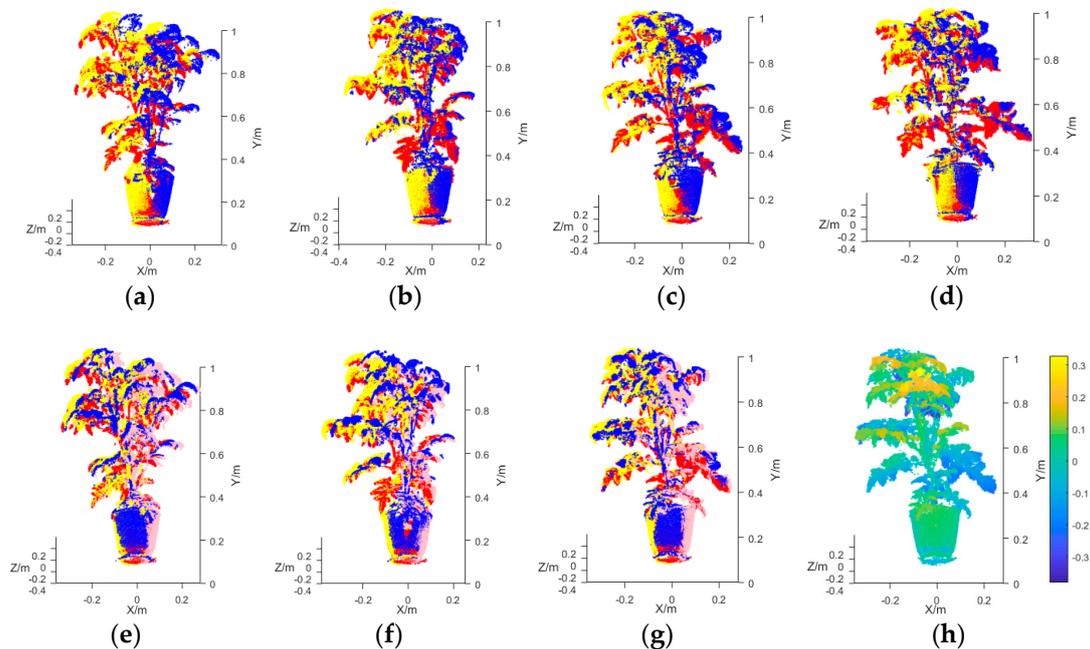

Figure 5. Results of the 3D point cloud reconstruction of a plant with different reconstruction methods: (**a**) V3-1; (**b**) V3-2; (**c**) V3-3; (**d**) V3-4; (**e**) V4-1; (**f**) V4-2; (**g**) V4-3; (**h**) depth map.

3.2. Accuracy Analysis of Point Cloud Reconstruction of the GTPs

To quantitatively describe the accuracy of 3D point cloud reconstruction of the GTPs, the distribution of the H_{RS} set of each of the 60 GTPs was statistically analyzed. Figure 6 shows the distribution of the H_{RS} sets of the GTPs numbered T25, T29, T48, and T56. Figure 7a shows the distribution of five groups (0 cm < H_{RS} ≤ 0.1 cm, 0.1 cm < H_{RS} ≤ 0.3 cm, 0.3 cm < H_{RS} ≤ 0.6 cm, 0.6 cm < H_{RS} ≤ 1.0 cm, and H_{RS} > 1.0 cm) of H_{RS} in each set.
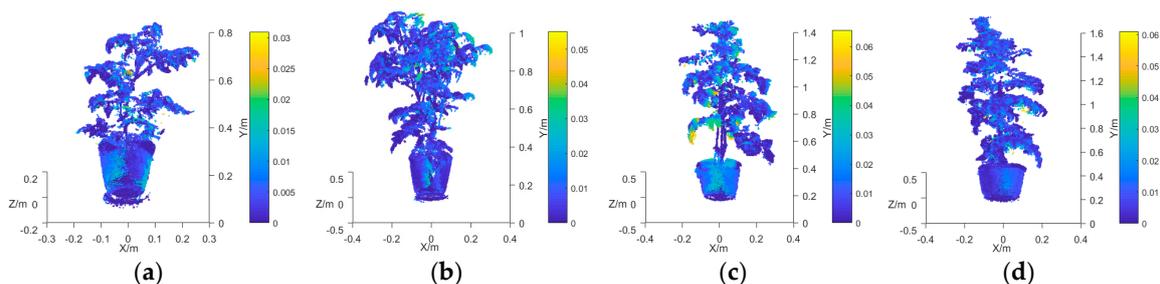

Figure 6. H_{RS} sets of the GTPs: (**a**) GTP T25; (**b**) GTP T29; (**c**) GTP T48; (**d**) GTP T56.

Figure 7b shows the metrics for assessing the accuracy of GTP 3D point cloud reconstruction, namely, *HD*, H_{avg}, and H_{std}, of the H_{RS} set. The average *HD*, H_{avg}, and H_{std} were 6.07, 0.46, and 0.54 cm, respectively. On average, the five groups of H_{RS}s (0 cm < H_{RS} ≤ 0.1 cm, 0.1 cm < H_{RS} ≤ 0.3 cm, 0.3 cm < H_{RS} ≤ 0.6 cm, 0.6 cm < H_{RS} ≤ 1.0 cm, and H_{RS} > 1.0 cm) accounted for 33.76%, 16.31%, 21.53%, 14.96%, and 13.44% of all the H_{RS}s, respectively. The statistical data show that, on average, 71.60% and 86.56% of the H_{RS}s in the H_{RS} sets were less than 0.6 cm and less than 1.0 cm, respectively. The average H_{avg} was 0.46 cm, suggesting that the reconstructed 3D point clouds of the GTPs were relatively highly accurate. The proposed 3D point cloud reconstruction method can be used to extract the 3D morphological parameters of plants.
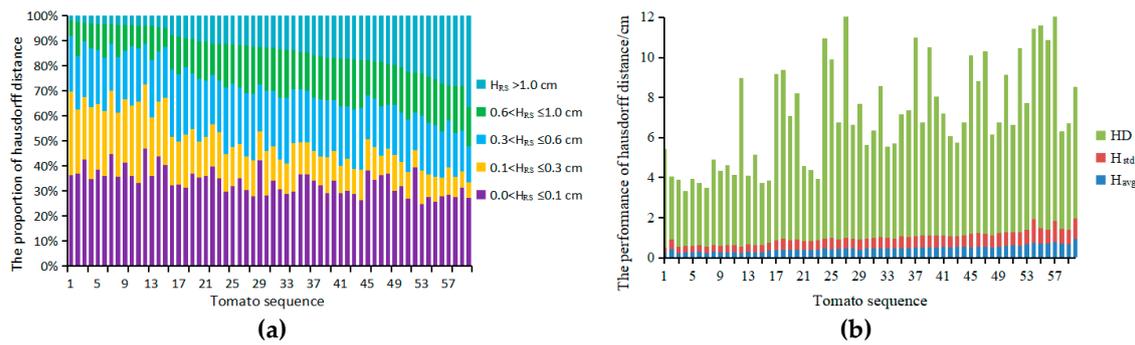
Figure 7. Analysis of point cloud reconstruction of the GTPs: (**a**) Distribution of the H_{RS} sets of the GTPs; (**b**) Metrics for assessing the accuracy of the GTP 3D point cloud reconstruction.

3.3. Calculation Method of 3D Point Cloud Morphological Characteristics

According to the GTPs' 3D cloud model, a calculation method was established for the H, W, S_{XOZ}, and V of the tomato canopy, and the calculation errors were compared with the values of the manual measurement. The H refers to the height of the tomato canopy, which is the vertical distance from the highest point of the plant to the upper surface of the cultivation pot. Due to the displacement of the 3D point cloud, the coordinates of the center of the rotary table are (0, 0, 0), so the maximum value of the Y-axis reflects the total height of the plant, as shown in Figure 8a, and the equation for calculating the H of the plant is shown in Equation (10). The morphology of the tomato's canopy is complex, and the width of the canopy from different views varies greatly, but the W of the canopy is unique and constant. The W refers to the maximum distance between the two points of the projection boundary on the horizontal XOZ plane. According to the Graham algorithm [41], the set of convex points on the XOZ plane was searched, as shown in Figure 8b, to calculate the W of the convex point set as the W of the canopy. The calculation equation is shown in Equation (11), and the connecting line for the W is shown in Figure 8b. Based on multiview RGB-D 3D reconstruction, the change of the initial reference point cloud would affect the projection morphology of the canopy point cloud on the XOY and YOZ planes, but the projection of the tomato canopy on the XOZ horizontal plane was not affected by the initial point cloud view. The XOZ horizontal plane projection boundary enclosing area is an invariant feature quantity, and the projection boundary enclosing area was calculated as shown in Equation (12).

$$H = Y_{\max} - H_{\text{flowerpot}} \tag{10}$$

$$W = \max\left(\sum_{i=1\sim m}^{j=i+1\sim m} \sqrt{(x_i - x_j)^2 + (z_i - z_j)^2}\right) \tag{11}$$

$$S_{XOZ} = \frac{1}{2}\sum_{i=1}^{m}[x_i z_{i+1} - z_i x_{i+1}] \tag{12}$$

Here, H is the canopy height of the plant, cm; $Y_{\max}$ is the maximum value on the Y-axis of the point cloud coordinates (cm); $H_{\text{flowerpot}}$ is the height of the cultivation pot (cm); W is the maximum width of the canopy (cm); S_{XOZ} is the projection area of the canopy on the XOZ plane (cm^2); m is the number of convex vertices; x_i is the x coordinate of the ith vertex; x_j is the x coordinate of the jth vertex; z_i is the z coordinate of the ith vertex; z_j is the z coordinate the jth vertex; x_{i+1} is the x coordinate of the $i + 1$th vertex; and z_{i+1} is the z coordinate of the $i + 1$th vertex.
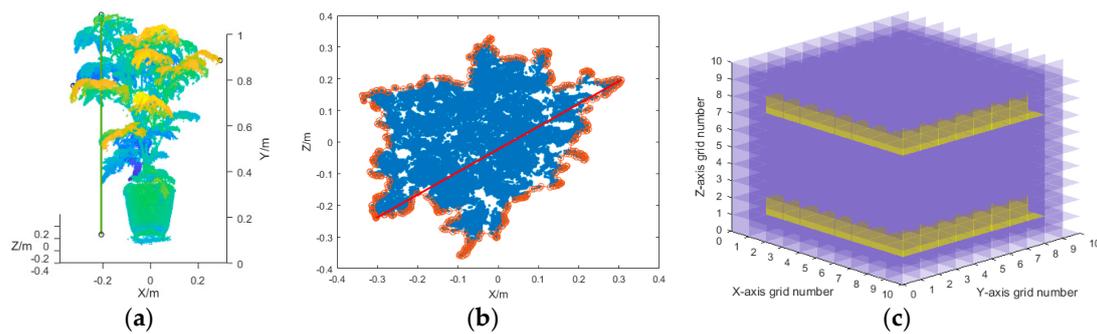
Figure 8. Calculation of canopy morphology parameters of a greenhouse tomato plant: (**a**) Total plant height; (**b**) Convex boundary point set of the canopy projection, and canopy maximum width; (**c**) Voxel division of the cuboid ($10 \times 10 \times 10$).

The 3D shape of a tomato canopy is complex. The method of the outer envelope convex volume cannot accurately describe the canopy volume. The large spaces in the convex volume will affect the accuracy of the volume measurement. This study established a method for the calculation of the canopy volume of a tomato plant, with the circumscribed cuboid of the tomato canopy divided into several cube voxels [42,43]. According to the voxel precision, the number of lattices of the X, Y, and Z axes of the cuboid could be determined. The $10 \times 10 \times 10$ voxel division diagram is illustrated in Figure 8c. First, the displacement operation of the canopy point cloud coordinates was performed. The minimum coordinates X_{min}, Y_{min}, and Z_{min} were respectively subtracted from the X, Y, and Z triaxial coordinates. Based on Equations (13)–(15), the cuboid was divided into voxels, and the 3D array of the voxels was initialized, with the initial value of each voxel being 0. According to Equation (16), the voxels of the point cloud were searched, and the voxel containing the point cloud was labeled as 1. The number of voxels containing the point cloud was counted, and the canopy volume was obtained based on the voxel precision.

$$N_x = (X_{max} - X_{min})/voxel \tag{13}$$

$$N_y = (Y_{max} - Y_{min})/voxel \tag{14}$$

$$N_z = (Z_{max} - Z_{min})/voxel \tag{15}$$

$$V = SUM(\sum_{i=1}^{n}(x_i, y_i, z_i) \in (G_{N_x}, G_{N_y}, G_{N_z})) \times voxel \tag{16}$$

Here, X_{max}, Y_{max}, and Z_{max} are the maximum values on the X-axis, Y-axis, and Z-axis of the canopy point cloud coordinates, respectively (m); X_{min}, Y_{min}, and Z_{min} are the minimum values on the X-axis, Y-axis, and Z-axis of the canopy point cloud coordinates, respectively (m); N_x, N_y, and N_z are the number of divided grids for the X-axis, Y-axis, and Z-axis, respectively; n is the number of point clouds; (x_i, y_i, z_i) are the point cloud coordinates; (G_{Nx}, G_{Ny}, G_{NZ}) is the number of 3D voxels; *voxel* is the voxel precision (m); and V is the canopy volume (m^3).

The point cloud down-sampling method (3D box filter threshold) and the voxel precision of the 3D point cloud model directly affect the accuracy of the tomato canopy volume. In this study, the point cloud model was down-sampled using 3D box filters of 2, 3.3, 5, and 8 mm, as shown in Figure 9a–d. At the same time, the point cloud model was connected to the outer cuboid for meshing, and 3D voxel models with voxel accuracies of $2 \times 2 \times 2$ mm^3, $3.3 \times 3.3 \times 3.3$ mm^3, $5 \times 5 \times 5$ mm^3, and $8 \times 8 \times 8$ mm^3 were established. The voxels of the point cloud were searched, and the voxels containing point clouds were labeled. As shown in Figure 9e–h, the number of voxels containing point clouds was counted, and the canopy volume was obtained based on the voxel precision.

In this study, the actual volume of the tomato canopy was measured by the water immersion method. The measurement method was as follows: pour some water into a graduated cylinder, record

the initial liquid volume V_0, and then completely immerse the tomato canopy below the liquid surface and record the liquid volume V_1 again. $V_1 - V_0$ is the actual measured volume V of the tomato canopy.
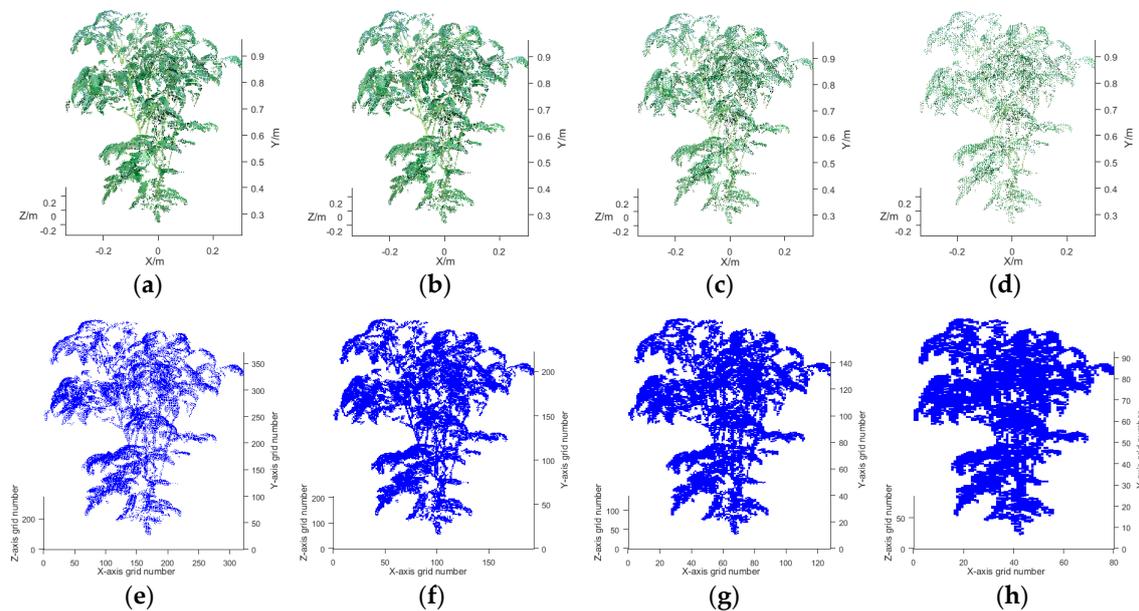

Figure 9. Calculation of the canopy volume of a greenhouse tomato plant: (**a**) Box filter: 2 mm; (**b**) Box filter: 3.3 mm; (**c**) Box filter: 5 mm; (**d**) Box filter: 8 mm; (**e**) Voxel: $2 \times 2 \times 2$ mm^3; (**f**) Voxel: $3.3 \times 3.3 \times 3.3$ mm^3; (**g**) Voxel: $5 \times 5 \times 5$ mm^3; (**h**) Voxel: $8 \times 8 \times 8$ mm^3.

3.4. Error Analysis of the Calculation Method of the 3D Point Cloud Morphological Features

According to the above calculation method of the 3D point cloud morphological features, the tomato point cloud model was reconstructed by measuring three views of V_3 and four views of V_4, with various view combinations for each measurement method, as shown in Table 1. Under different measurement modes, the SD and CV of the H, W, S_{XOZ}, and N_P morphological values were statistically analyzed. The CV and R^2 of the calculated tomato canopy volume V under different measurement methods and different voxel precisions with the actual measured volume were statistically analyzed.

Figure 10a–d shows the values of H, W, S_{XOZ}, and N_P for the canopies of 60 GTPs calculated by different measurement methods.

As shown in Table 2, in the V_3 measurement mode, for the 60 GTPs, H ranged from 18.73 to 131.13 cm, the average value (AVG) of SD was 0.37 cm, and the AVG of CV was 0.62%. In the V_4 measurement mode, H ranged from 18.74 to 130.35 cm, the AVG of the SD was 0.30 cm, and the AVG of CV was 0.50%. In the V_3 measurement mode, for the 60 GTPs, W ranged from 24.12 to 85.71 cm, the AVG of SD was 1.76 cm, and the AVG of CV was 3.25%. In the V_4 measurement mode, W ranged from 23.92 to 86.56 cm, the AVG of the SD was 1.57 cm, and the AVG of CV was 2.93%.

As shown in Table 2, in the V_3 measurement mode, for the 60 GTPs, S_{XOZ} ranged from 246.79 to 2771.60 cm^2, the AVG of SD was 71.24 cm^2, and the AVG of CV was 5.30%. In the V_4 measurement mode, S_{XOZ} ranged from 264.69 to 2926.02 cm^2, the AVG of SD was 64,030 cm^2, and the AVG of CV was 4.54%. In the V_3 measurement mode, N_P ranged from 2882.25 to 65,448.00, the AVG of SD was 947.44, and the AVG of CV was 3.83%. In the V_4 measurement mode, N_P ranged from 3489.67 to 84,132.00, the AVG of SD was 1042.10, and the AVG of CV was 3.06%.

The statistical data show that for the tomato morphological parameters calculated by the V_4 measurement method, the mean values of the SD and CV of H, W, S_{XOZ}, and N_P were smaller than those calculated by the V_3 measurement method, and the performance of the calculation of the morphological parameters was more stable.

Table 2. Effects of different measurement methods on the calculated values of the canopy morphological parameters.

| V_N	Measured Value	Calculated Value			SD			CV		
		Max	Min	Avg	Max	Min	Avg	Max	Min	Avg
V_3	H/cm	131.13	18.73	72.22	1.49	0.04	0.37	2.49%	0.03%	0.62%
	W/cm	85.71	24.12	54.78	6.73	0.24	1.76	10.22%	0.52%	3.25%
	S_{XOZ}/cm^2	2771.60	246.79	1488.58	252.41	7.82	71.24	14.21%	1.18%	5.30%
	N_P	65,448.00	2882.25	29,075.39	2924.91	81.94	947.44	13.88%	0.57%	3.83%
V_4	H/cm	130.35	18.74	72.25	1.18	0.02	0.30	2.00%	0.04%	0.50%
	W/cm	86.56	23.92	56.18	7.23	0.06	1.57	12.50%	0.17%	2.93%
	S_{XOZ}/cm^2	2926.02	264.69	1591.00	265.27	6.40	64.30	20.19%	0.60%	4.54%
	N_P	84,132.00	3489.67	37,532.43	4546.59	30.27	1042.10	9.60%	0.40%	3.06%
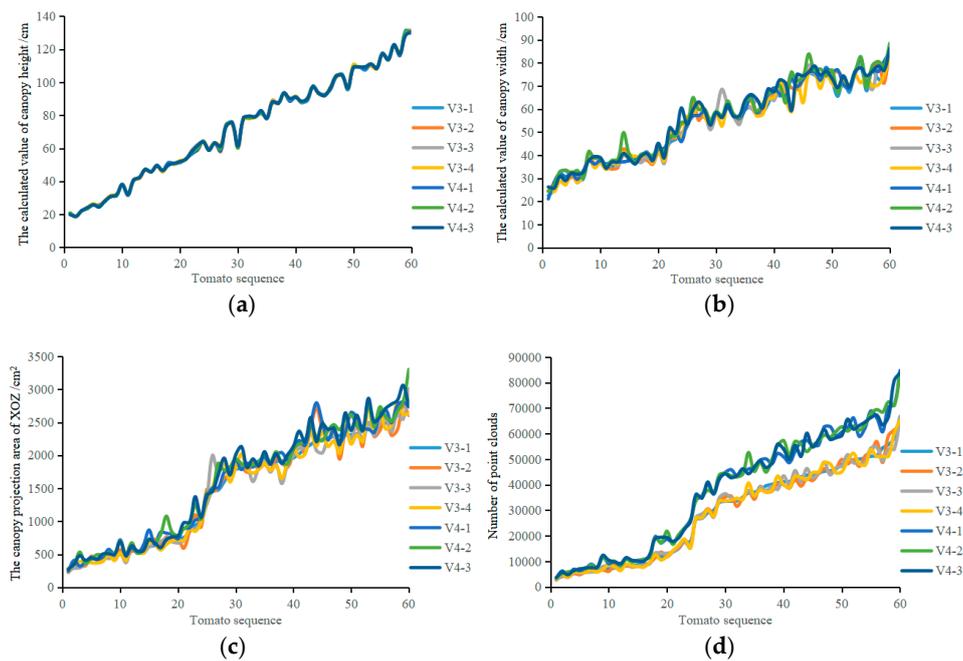
Figure 10. Canopy morphological parameters of GTPs: (**a**) canopy height H; (**b**) canopy maximum width W; (**c**) canopy projected area S_{XOZ}; (**d**) canopy point cloud number N_P.

In this study, the canopy volume V of the tomato plants was calculated using the V_3 and V_4 measurement modes with voxel precisions of 2, 3.3, 5, and 8 mm, as shown in Table 3.

Table 3. GTP canopy volume calculation method.

Calculation Method for Canopy Volume		Reconstruction Method	Voxel Precision	Calculation Method for Canopy Volume		Reconstruction Method	Voxel Precision
VA	VA-3	V_3	2 mm	VC	VC-3	V_3	5 mm
	VA-4	V_4	2 mm		VC-4	V_4	5 mm
VB	VB-3	V_3	3.3 mm	VD	VD-3	V_3	8 mm
	VB-4	V_4	3.3 mm		VD-4	V_4	8 mm

As shown in Figure 11a, the MIN of the CV was 0.00% in the VB-4 mode, and the MAX of the CV was 14.76% in the VD-4 mode. The statistical results of tomato canopy volume variability showed that the range of the mean CV was 3.25%–5.53% by different calculation methods, and the correlation of the coefficient of variation was not significant. However, with the increase in the voxel precision, the maximum coefficient of variation showed a significant increasing trend.
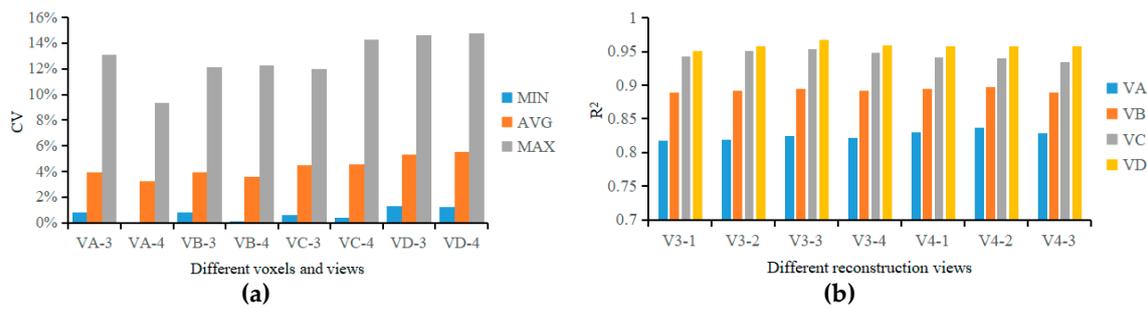
Figure 11. Performance parameters of canopy volume measurement. (**a**) CV of canopy volume, (**b**) R^2 for the calculated and measured values of canopy volume.

As shown in Figure 11b, in reconstruction modes V3-1, V3-2, V3-3, V3-4, V4-1, V4-2, and V4-3, with the increase in the voxel precision, the R^2 between the calculation values and the actual measurement values of the volume of the tomato canopy had a significant increasing trend. For different voxel precisions, the R^2 mean values of VA, VB, VC, and VD and the actual measured volume values were 0.8252, 0.8927, 0.9442, and 0.9586, respectively. In reconstruction modes V3-1, V3-2, V3-3, V3-4, V4-1, V4-2, and V4-3, the R^2 mean values of the calculation values and the actual measurement values of the volume were 0.8999, 0.9049, 0.9099, 0.9052, 0.9057, 0.9083, and 0.9025, respectively. The statistical data showed that the calculated volume of the tomato canopy could be significantly affected by the voxel precision but not by the reconstruction method.

3.5. Applicability Analysis of Geometrical Calculation Methods for Greenhouse Tomato Plants

To verify the applicability of the multiview RGB-D reconstruction method and the morphological feature parameter calculation method proposed in this study, the 3D point cloud reconstruction was performed on 60 GTPs, obtaining the RGB point cloud maps of the tomato plants shown in Figure 12a–f. The point cloud maps of the tomato plants reconstructed based on three views are shown in Figure 12g–l, and the point cloud maps of the tomato plants reconstructed based on four views are shown in Figure 12m–r. Additionally, the R^2, RMSE, and RAD of the calculated values of *H*, *W*, *V*, and *FW* were statistically analyzed with the values of the actual manual measurements.
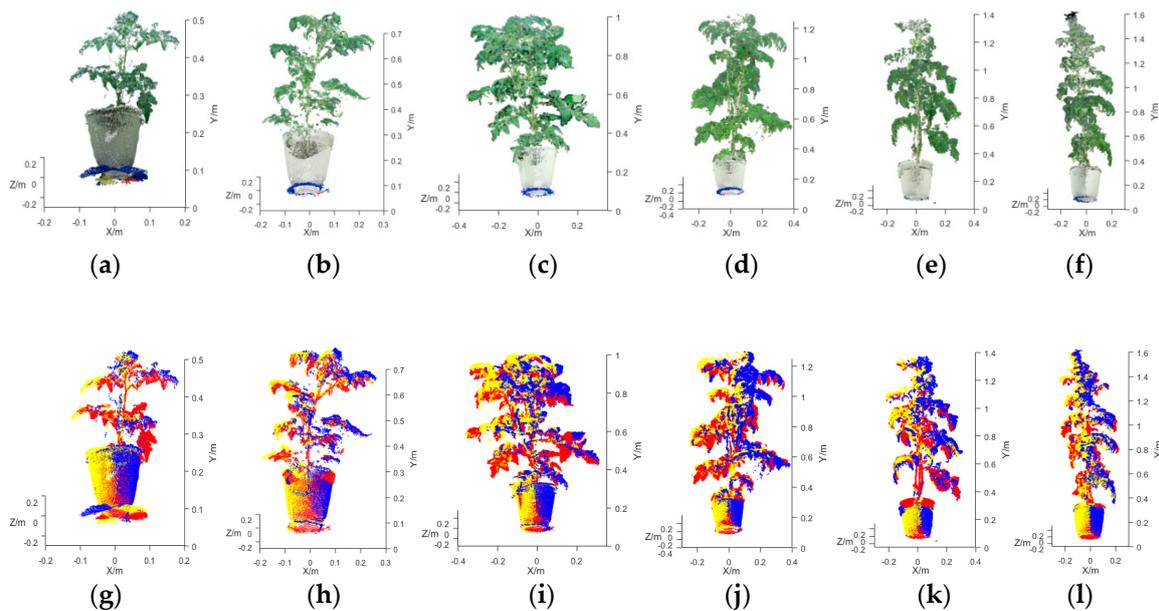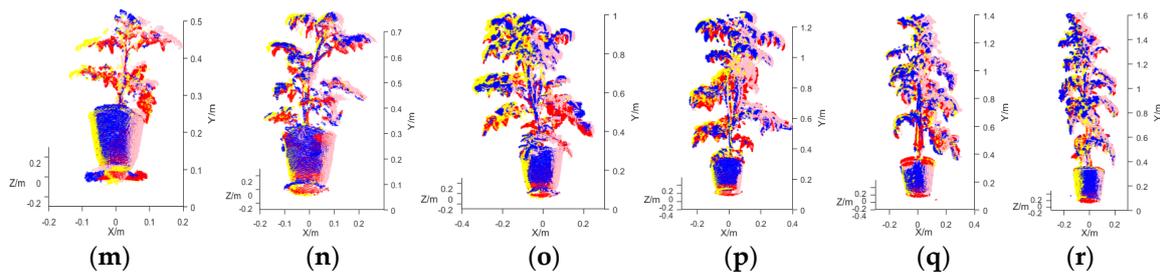

Figure 12. *Cont.*

Figure 12. 3D point cloud reconstruction maps of GTPs: (**a**) T9 point cloud; (**b**) T25 point cloud; (**c**) T29 point cloud; (**d**) T31 point cloud; (**e**) T48 point cloud; (**f**) T56 point cloud; (**g**) T9 (V3-1); (**h**) T25 (V3-1); (**i**) T29 (V3-1); (**j**) T31 (V3-1); (**k**) T48 (V3-1); (**l**) T56 (V3-1); (**m**) T9 (V4-1); (**n**) T25 (V4-1); (**o**) T29 (V4-1); (**p**) T31 (V4-1); (**q**) T48 (V4-1); (**r**) T56 (V4-1).

The correlation between the actual measurement values of the plant canopy H and W and the Kinect measurement values is shown in Figure 13a,b. The Kinect measurements of the plant canopy H and W could be calculated directly from the reconstructed point cloud, as shown in Equations (10) and (11).
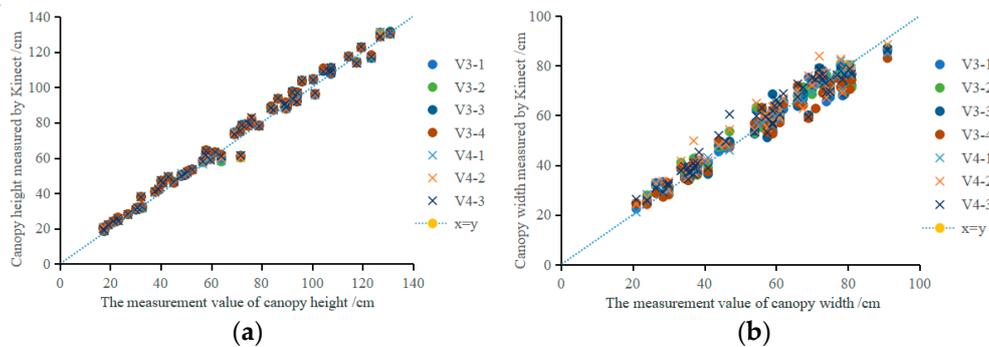

Figure 13. The results of canopy height and width: (**a**) calculated and measured values of canopy height; (**b**) calculated and measured values of canopy maximum width.

As shown in Table 4, in the V_3 measurement mode, the MIN, MAX, and AVG of the R^2 between the Kinect measurement value and the manual measurement value of the canopy H of the plants were 0.9883, 0.9897, and 0.9890, respectively; the MIN, MAX, and AVG of the RMSE were 0.30, 10.88, and 3.15 cm, respectively; and the MIN, MAX, and AVG of the RAD were 0.41%, 18.05%, and 5.53%, respectively. The MIN, MAX, and AVG of the correlation coefficient R^2 between the Kinect measurement value and the manual measurement value of the canopy width W of the plants were 0.9519, 0.9658, and 0.9587, respectively; the MIN, MAX, and AVG of the RMSE were 0.31, 7.76, and 3.30 cm, respectively; and the MIN, MAX, and AVG of the RAD were 0.49%, 18.07%, and 5.60%, respectively.

As shown in Table 4, in the V_4 measurement mode, the MIN, MAX, and AVG of the R^2 between the Kinect measurement value and the manual measurement value of the canopy H of the plants were 0.9880, 0.9894, and 0.9887, respectively; the MIN, MAX, and AVG of the RMSE were 0.37, 10.75, and 3.20 cm, respectively; and the MIN, MAX, and AVG of the RAD were 0.37%, 18.47%, and 5.59%, respectively. The MIN, MAX, and AVG of the correlation coefficient R^2 between the Kinect measurement value and the manual measurement value of the canopy width W of the plants were 0.9516, 0.9752, and 0.9597, respectively; the MIN, MAX, and AVG of the RMSE were 0.43, 8.84, and 3.49 cm, respectively; and the MIN, MAX, and AVG of the RAD were 0.59%, 20.44%, and 6.47%, respectively.

In this study, a canopy volume and fresh weight measurement model was established. Based on the V_3 and V_4 measurement methods, the average values of the tomato canopy H, W, S_{XOZ}, and V measured by Kinect were used as the input values, and the actual measured volume V of the canopy was used as the output to establish a multivariate stepwise regression model. In the V_3 measurement

mode, the canopy volume and fresh weight could be calculated by the equations $V_3 = 2.462 + 0.124 \times V\text{-}0.114 \times S_{XOZ}$ and $FW_3 = 6.576 + 0.106 \times V\text{-}0.114 \times S_{XOZ}$, with regression equation R^2 of 0.953 and 0.934, respectively. In the V_4 measurement mode, the canopy volume and fresh weight could be calculated by the equations $V_4 = 48.886 + 0.054 \times V\text{-}3.258 \times W + 1.242 \times H$ and $FW_4 = -20.56 + 0.049 \times V\text{-}0.081 \times S_{XOZ} + 1.069 \times H$, with regression equation R^2 of 0.955 and 0.939, respectively.

Table 4. Relationship between the calculated values and the measured values of the canopy morphological parameters by different measurement methods.

Measurement Method	Measurement Value	R^2			RMSE			RAD		
		MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG
V_3	H/cm	0.9883	0.9897	0.9890	0.30	10.88	3.15	0.41%	18.05%	5.53%
	W/cm	0.9519	0.9658	0.9587	0.31	7.76	3.30	0.49%	18.07%	5.60%
	V/cm^3	0.9190	0.9491	0.9297	3.81	59.53	23.95	1.41%	26.95%	9.77%
	FW/g	0.8906	0.9195	0.9056	2.09	67.97	20.24	1.10%	29.40%	10.62%
V_4	H/cm	0.9880	0.9894	0.9887	0.37	10.75	3.20	0.37%	18.47%	5.59%
	W/cm	0.9516	0.9752	0.9597	0.43	8.84	3.49	0.59%	20.44%	6.47%
	V/cm^3	0.9018	0.9341	0.9205	2.55	57.79	23.46	0.60%	24.75%	9.27%
	FW/g	0.9000	0.9225	0.9108	0.74	58.46	15.99	1.35%	23.09%	8.52%

As shown in Figure 14a,b, in the V_3 measurement mode, the R^2, RMSE, and RAD between the calculated values of the plant canopy V by the regression model and the manual measurement values were 0.9297, 23.95 cm^3, and 9.77%, respectively. The R^2, RMSE, and RAD between the calculated values of the plant canopy FW by the regression model and the manual measurement values were 0.9056, 20.24 g, and 10.62%, respectively. In the V_4 measurement mode, the R^2, RMSE, and RAD between the calculated values of the plant canopy V by the regression model and the manual measurement values were 0.9205, 23.46 cm^3, and 9.27%, respectively. The R^2, RMSE, and RAD between the calculated values of the plant canopy FW by the regression model and the manual measurement values were 0.9108, 15.99 g, and 8.52%, respectively.
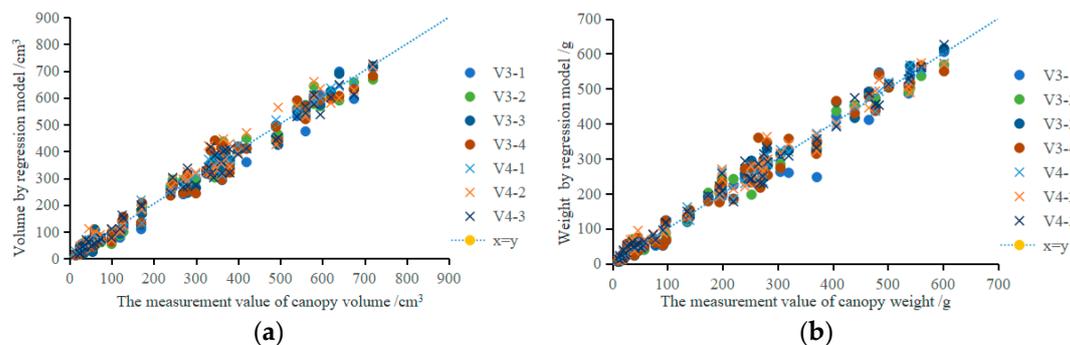

Figure 14. The results of canopy volume and fresh weight: (**a**) calculated and measured values of canopy volume; (**b**) calculated and measured values of canopy fresh weight.

The H and W of the plant were directly measured. The statistical data show that the error of the H measurement was the smallest, followed by the W, and the performance of the V_3 measurement mode was better than that of the V_4 measurement mode. The three main views could basically cover the entire canopy area. With the increase in the number of views, the probability of causing noise increased, resulting in an increase in the relative error of the direct measurement. The canopy volume and fresh weight of the plant were indirectly measured. The accuracy was affected by the measured values of the canopy morphological parameters of the plant and the calculation model. The relative errors were higher than those of H and W. Since the stepwise regression method was adopted in the calculation model, the inputs for the V_3 measurement mode included V and S_{XOZ}, while the inputs of the V_4 measurement mode included V, S_{XOZ}, and H. Therefore, the performance of the calculation

model based on the V_4 measurement mode was better than that based on the V_3 measurement mode. Of course, the construction method of the calculation model, the number of input morphological parameters, and whether it is a linear or nonlinear model will all affect the accuracy of the indirect measurement, which is not discussed in detail here.

4. Conclusions

This study proposed an autonomous Kinect sensor position calibration method. With only two RGB-D images of the turntable surface, a displacement matrix and a rotation matrix for unifying the coordinate systems of multiview point clouds can be obtained. The proposed method mainly addresses the rapid multiview point cloud registration issue, significantly simplifies the 3D point cloud reconstruction process of plants, and meets the full-growth-cycle high-throughput measurement requirement of plants. The average HD and H_{avg} between the reconstructed and reference point clouds of the GTPs were 6.07 and 0.46 cm, respectively. In addition, 71.60% and 86.56% of the H_{RS} in the H_{RS} sets were less than 0.6 and 1.0 cm, respectively. At the same time, the correlation and errors of the calculated values and the measured values of the canopy morphological parameters including H, W, V, and FW for 60 tomato plants were statistically analyzed. In the V_3 measurement mode, the RAD mean values were 5.53%, 5.60%, 9.77%, and 10.62%, respectively. In the V_4 measurement mode, the RAD mean values were 5.59%, 6.47%, 9.27%, and 8.52%, respectively.

The proposed 3D point cloud reconstruction method is highly accurate and stable in performance and can be used to reconstruct 3D point clouds for high-throughput plant phenotyping analysis and to extract the morphological parameters of plants. In addition, the proposed method can be used to extract many other 3D geometric morphological and phenotypic parameters.

Abbreviations

2D	two-dimensional
2.5D	two-and-a-half-dimensional
3D	three-dimensional
RGB	red–green–blue
RGB-D	red–green–blue–depth
CT	computed tomography
MRI	magnetic resonance imaging
AOV	angle of view
AOVs	angles of view
TOF	time of flight
ICP	iterative closest point
SSBs	standard Styrofoam balls
GTPs	greenhouse tomato plants
LED	light-emitting diode
fps	frames per second
V_N	the number of angles of view for 3D reconstruction of the plant
H	height
W	maximum width

N_P	point cloud number
S_{XOZ}	area of the canopy projected in the horizontal plane
V	canopy volume
FW	fresh weight
V_3	three angles of view
V_4	four angles of view
V_6	six angles of view
V_N	number of angles of view for 3D reconstruction of the plant
RAD	relative average deviation
CV	coefficient of variation
SD	standard deviation
AVG	average value
MAX	maximum value
MIN	minimum value
HD	Hausdorff distance
H_{avg}	average of the Hausdorff distance set
H_{std}	standard deviation of the Hausdorff distance set
H_{RS}	set of distances between the reconstructed and reference point clouds
H_{SR}	set of distances between the points of the reference and reconstructed point clouds
R^2	coefficient of determination
RMSE	root-mean-square error

Appendix A

In order to determine the angle interval and the number of views for the greenhouse tomato plants point cloud model reconstruction, the standard Styrofoam balls (SSBs) were selected as measurement objects. Four white SSBs with diameters of 30, 40, 50, and 60 cm were selected. The above 2.2 3D reconstruction method was used to reconstruct the SSBs point cloud model. The 3D point cloud reconstruction accuracy data for the SSBs were analyzed using the following method. During the SSBs point cloud reconstruction experiment, the Kinect sensor was placed in three positions (P_1, P_2, and P_3). In addition, three combinations of AOVs, namely, V_3 (0°, 120°, and 240°), V_4 (0°, 90°, 180°, and 270°), and V_6 (0°, 60°, 120°, 180°, 240°, and 300°) were used. Statistical data included the relative average deviation and coefficient of variation between the reconstructed and measured values of the diameter (D_X) in the horizontal (X-axis) direction, the diameter (D_Y) in the vertical (Y-axis) direction, the volume (*Vol*), the coverage (*Cr*) of the reconstructed point cloud, the distribution frequency of the set of distances (H_{RS}) between the reconstructed and reference point clouds of each SSB, the Hausdorff distance (*HD*) between the reconstructed and reference point clouds, and the average (H_{avg}) and standard deviation (H_{std}) of the *HD* set.

$$Cr = \frac{S_{HD}}{S_{standardball}} \times 100\% \tag{A1}$$

Here, *Cr* is the percentage of the surface area of the SSB covered by the reconstructed point cloud (%); $S_{standardball}$ is the surface area of the SSB (cm^2); and S_{HD} is the point cloud surface area for which the distance between the scanned and reconstructed point clouds is less than 5.00% of the diameter of the SSB (cm^2).

Figure A1a,b show an RGB color image and a depth image of an SSB, respectively. According to the 3D point cloud reconstruction process, 3D point clouds of the four SSBs with various diameters (30, 40, 50, and 60 cm) were reconstructed based on images captured by the Kinect sensor in three positions (P_1, P_2, and P_3) and at three combinations of AOVs (V_3, V_4, and V_6). In addition, a reference point cloud consisting of 90,000 points was constructed for each of the four SSBs. Figure A1c shows the reference point cloud of the SSB with a diameter of 60 cm.
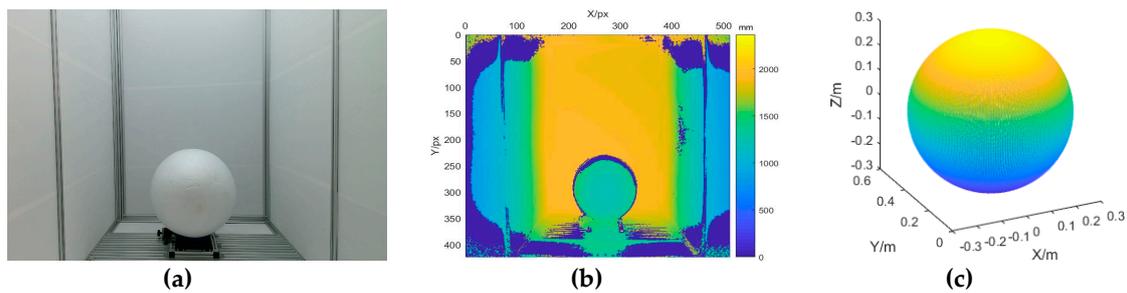
Figure A1. Reconstruction of a 3D point cloud of the standard Styrofoam ball (SSB) with a diameter of 60 cm: (**a**) color image of the SSB; (**b**) depth image of the SSB; (**c**) reference point cloud of the SSB.

Figure A2 shows the reconstructed 3D point clouds of the SSB and corresponding reconstruction accuracy analysis. Figure A2a shows the 3D point cloud of the SSB reconstructed based on RGB-D images captured at three AOVs (point clouds at 0°, 120°, and 240°, highlighted in red, yellow, and blue, respectively). The 3D point cloud was down-sampled using a 3D mesh filter. The 3D mesh threshold was set to 5 mm. Based on its outer boundary, the volume of the 3D point cloud of the SSB was calculated. As shown in Figure A2b, a set of distances (H_{SR}) between the points of the reference and reconstructed point clouds of the SSB was calculated. When $H_{SR} < 5.00\%$ of the diameter of the SSB, the point was considered to have been scanned by the sensor. In Figure A2c, the area of the SSB scanned by the sensor is marked. Based on Equation (6), the Cr of the reconstructed point cloud was calculated. Based on Equations (7)–(10), HD and H_{RS} of the SSB were calculated, H_{RS} as shown in Figure A2d.

Similarly, Figures A2e–h and A2i–l show the 3D point clouds of the SSB reconstructed based on images captured at four and six AOVs, respectively, as well as corresponding reconstruction accuracy analysis. In Figure A2e, the four AOVs (0°, 90°, 180°, and 270°) are highlighted in red, yellow, blue, and gray, respectively. In Figure A2i, the six AOVs (0°, 60°, 120°, 180°, 240°, and 300°) are highlighted in red, yellow, blue, gray, orange, and sky blue, respectively.

To quantitatively describe the accuracy of SSB 3D point cloud reconstruction, the distribution of the H_{RS} set of each of the SSBs with diameters of 30, 40, 50, and 60 cm and the corresponding point clouds reconstructed based on images captured by the Kinect sensor in three positions (P_1, P_2, and P_3) and at three, four, and six AOVs (V_3, V_4, and V_6) (a total of 36 combinations of measurement conditions) were statistically analyzed. The results are shown in Figure A3a. H_{RS}s were categorized into five groups for statistical analysis, namely, 0 cm $< H_{RS} \leq 0.2$ cm, 0.2 cm $< H_{RS} \leq 0.5$ cm, 0.5 cm $< H_{RS} \leq 0.8$ cm, 0.8 cm $< H_{RS} \leq 1.2$ cm, and $H_{RS} > 1.2$ cm.

Figure A3b shows the metrics for assessing the accuracy of 3D point cloud reconstruction of SSBs, namely, HD and the average (H_{avg}) and standard deviation (H_{std}) of the H_{RS} set. A comparison of the point clouds of the SSBs with diameters of 30, 40, 50, and 60 cm reconstructed based on images captured by the Kinect sensor in positions P_1, P_2, and P_3 at V_3, V_4, and V_6 and the corresponding respective reference point clouds shows that the average HDs were 2.77, 4.33, 5.41, and 6.38 cm, respectively; the average H_{avg}s were 0.64, 0.93, 1.26, and 1.14 cm, respectively; the average relative H_{avg}s were 2.13%, 2.33%, 2.52%, and 1.90%, respectively; and the average H_{std}s were 0.41, 0.64, 0.79, and 0.72 cm, respectively. The statistical data show that the excessively large HDs were caused by the noise in the point clouds. However, based on H_{avg} and relative H_{avg}, the average distance between the reconstructed and reference point clouds of the SSBs was less than 1.26 cm, and the error in the reconstructed point clouds was less than 2.52%. A comparison of the point clouds of the SSBs reconstructed based on images captured at V_3, V_4, and V_6 and the corresponding reference point clouds shows that the average HDs were 4.68, 4.38, and 5.12 cm, respectively; the average H_{avg}s were 0.98, 0.99, and 1.01 cm, respectively; and the average H_{std}s were 0.62, 0.64, and 0.65 cm, respectively. A comparison of the point clouds of the SSBs reconstructed based on images captured by the Kinect sensor in positions P_1, P_2, and P_3 and the corresponding reference point clouds shows

that the average *HD*s were 5.26, 5.29, and 3.81 cm, respectively; the average H_{avg}s were 1.08, 0.95, and 0.95 cm, respectively; and the average H_{std}s were 0.67, 0.64, and 0.61 cm, respectively. According to the statistical data, because *Cr* varied insignificantly between V_3, V_4, and V_6 and the reconstructed point clouds were down-sampled in the same way, the accuracy of point cloud reconstruction was not significantly affected by V_N or by the Kinect sensor position.
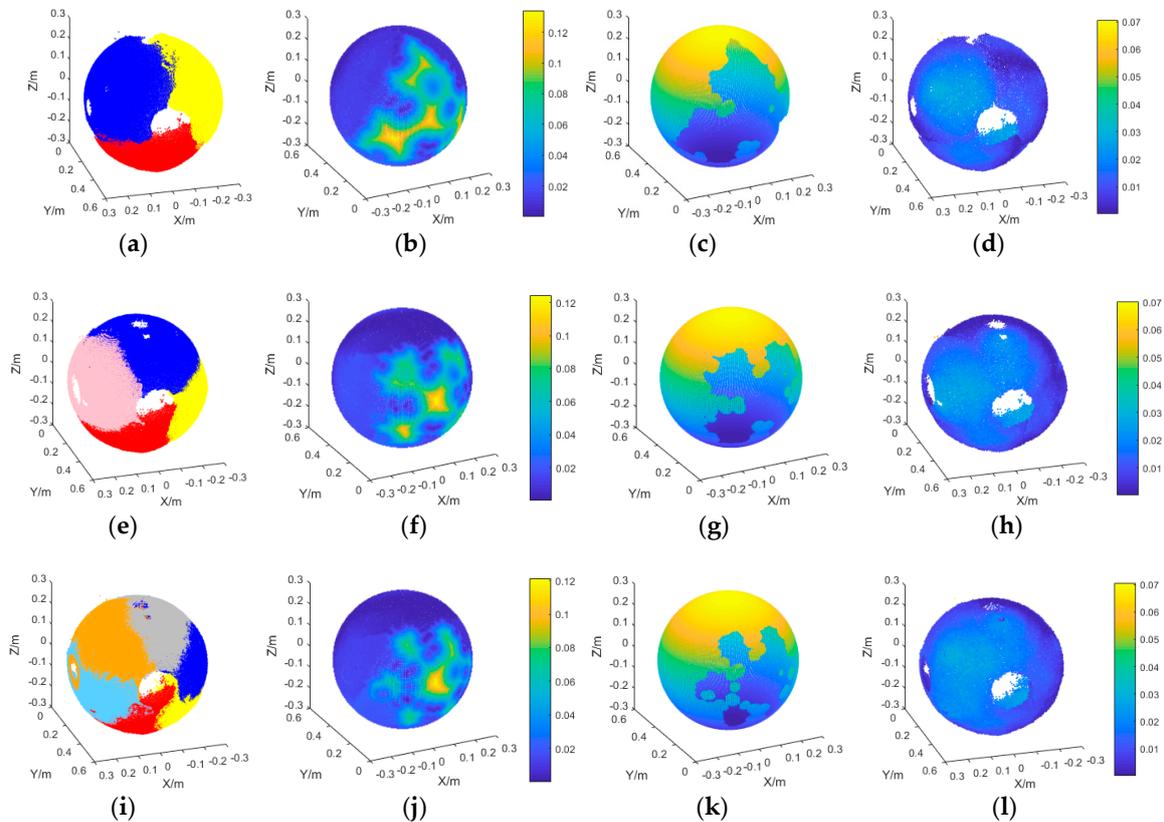

Figure A2. Reconstructed 3D point cloud of the SSB and reconstruction accuracy analysis. (**a**) point cloud of the SSB reconstructed based on images captured at three AOVs, (**b**) H_{SR} set, (**c**) area covered by the reconstructed point cloud, (**d**) H_{RS} set. (**a**–**d**) Point clouds reconstructed based on images captured at three AOVs. (**e**–**h**) Point clouds reconstructed based on images captured at four AOVs. (**i**–**l**) Point clouds reconstructed based on images captured at six AOVs.
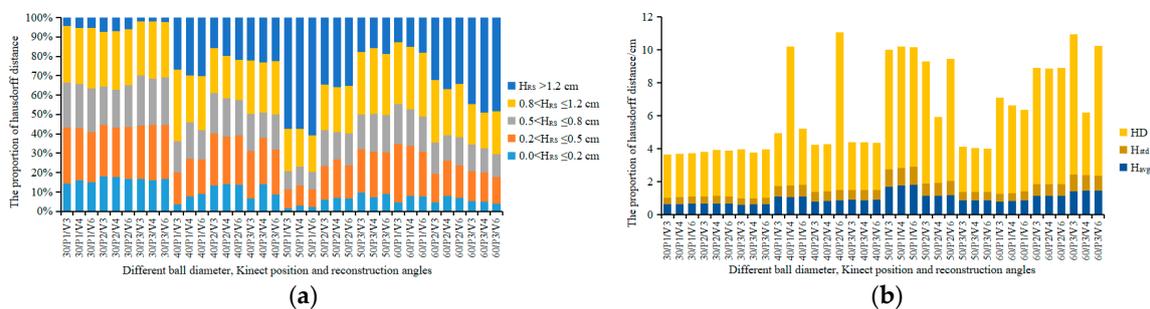

Figure A3. Analysis of point cloud reconstruction of the SSBs: (**a**) Distribution of the H_{RS} sets of the SSBs; (**b**) Metrics for assessing the accuracy of SSB 3D point cloud reconstruction.

Table A1 summarizes the statistical morphological measurement error data for the point clouds of the SSBs reconstructed based on images captured by the Kinect sensor in positions P_1, P_2, and P_3 and at V_3, V_4, and V_6. For the SSBs with diameters of 30, 40, 50, and 60 cm, the RADs for D_Y were 2.96%, 2.49%, 1.99%, and 1.97%, respectively; the coefficients of variation (CVs) for D_Y were 3.50%, 4.14%,

4.06%, and 4.76%, respectively; the RADs for D_X were 2.01%, 1.63%, 1.40%, and 1.61%, respectively; the CVs for D_X were 2.27%, 2.53%, 3.08%, and 4.42%, respectively; the RADs for *Vol* were 4.87%, 3.95%, 1.72%, and 5.02%, respectively; the CVs for *Vol* were 5.27%, 4.35%, 2.06%, and 5.18%, respectively; and the average *Cr*s were 92.81%, 89.85%, 89.91%, and 86.42%, respectively. The statistical data show that the measurement error in D_X was smaller than that in D_Y and that the CV for D_X was smaller than that for D_Y. This measurement error occurred mainly because some areas of the top and bottom of each SSB were not scanned, as shown in Figure A2. In addition, *Cr* decreased as the diameter of the SSB increased.

Table A1. Analysis of morphological measurements of SSBs that differ in diameter.

Ball Diameter/cm	D_Y		D_X		*Vol*		*Cr*/%
	RAD/%	CV/%	RAD/%	CV/%	RAD/%	CV/%	
30	2.96	3.50	2.01	2.27	4.87	5.27	92.81
40	2.49	4.14	1.63	2.53	3.95	4.35	89.85
50	1.99	4.06	1.40	3.08	1.72	2.06	89.91
60	1.97	4.76	1.61	4.42	5.02	5.18	86.42

Table A2 summarizes the RADs for D_Y, D_X, and *Vol* and average *Cr* of the point clouds of SSBs with diameters of 30, 40, 50, and 60 cm reconstructed based on images captured by the Kinect sensor in three positions (P_1, P_2, and P_3) and at three, four, and six AOVs (V_3, V_4, and V_6). For measurements taken at V_3, V_4, and V_6, the RADs for D_Y were 2.33%, 2.38%, and 2.34%, respectively; the RADs for D_X were 1.52%, 1.42%, and 2.05%, respectively; the RADs for *Vol* were 4.14%, 4.00%, and 3.52%, respectively; and the average *Cr*s were 85.45%, 90.17%, and 93.62%, respectively. The statistical data show that V_N did not significantly affect the error in the morphological measurement of the point cloud, but *Cr* increased significantly as V_N increased. For measurements taken in positions P_1, P_2, and P_3, the RADs for D_Y were 1.24%, 2.70%, and 3.12%, respectively; the RADs for D_x were 1.30%, 1.31%, and 2.37%, respectively; the RADs for *Vol* were 3.32%, 4.16%, and 4.19%, respectively; and the average *Cr*s were 87.10%, 87.60%, and 94.54%, respectively. The statistical data show that as the distance between the Kinect sensor and the measurement object decreased, *Cr* increased significantly, but the RADs for the morphological parameters also increased.

Table A2. Analysis of the morphological measurements of the SSBs (at various AOVs and in various positions).

Reconstruction Angle	RAD			*Cr*	Kinect Position	RAD			*Cr*/%
	D_Y/%	D_X/%	*Vol*/%			D_Y/%	D_X/%	*Vol*/%	
V_3	2.33	1.52	4.14	85.45	P_1	1.24	1.30	3.32	87.10
V_4	2.38	1.42	4.00	90.17	P_2	2.70	1.31	4.16	87.60
V_6	2.34	2.05	3.52	93.62	P_3	3.12	2.37	4.19	94.54

Based on the above analysis, for measurements taken at V_3, V_4, and V_6, the average *Cr*s were 85.45%, 90.17%, and 93.62%, respectively. The greater the number of perspectives, the greater the coverage, but the lower the reconstruction efficiency. Because the SSB is an entity measurement object, and the tomato plant is only partially blocked under each perspective, V_3 and V_4 reconstruction methods were selected to reconstruct the greenhouse tomato plants.

References

1. Pan, Y. Analysis of concepts and categories of plant phenome and phenomics. *Acta Agron. Sin.* **2015**, *41*, 175–186. [CrossRef]
2. Furbank, R.T.; Tester, M. Phenomics—Technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* **2011**, *16*, 635–644. [CrossRef]

3. Dhondt, S.; Wuyts, N.; Inzé, D. Cell to whole-plant phenotyping: The best is yet to come. *Trends Plant Sci.* **2013**, *18*, 428–439. [CrossRef] [PubMed]

4. Zhou, J.; Tardieu, F.; Pridmore, T.; Doonan, J.; Reynolds, D.; Hall, N.; Griffiths, S.; Chen, T.; Zhu, Y.; Wang, X.; et al. Plant phenomics: History, present status and challenges. *J. Nanjing Agric. Univ.* **2018**, *41*, 580–588. [CrossRef]

5. Rahaman, M.M.; Chen, D.; Gillani, Z.; Klukas, C.; Chen, M. Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Front. Plant Sci.* **2015**, *6*, 619. [CrossRef]

6. Perez-Sanz, F.; Navarro, P.J.; Egea-Cortines, M. Plant phenomics: An overview of image acquisition technologies and image data analysis algorithms. *Gigascience* **2017**, *6*, 1–18. [CrossRef] [PubMed]

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

8. Sun, G.; Li, Y.; Zhang, Y.; Wang, X.; Chen, M.; Li, X.; Yan, T. Nondestructive measurement method for greenhouse cucumber parameters based on machine vision. *Eng. Agric. Environ. Food* **2016**, *9*, 70–78. [CrossRef]

9. Kutulakos, K.N.; Seitz, S.M. A theory of shape by space carving. *Int. J. Comput. Vis.* **2000**, *38*, 199–218. [CrossRef]

10. Zheng, E.; Wu, C. Structure from motion using structure-less resection. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 2075–2083. [CrossRef]

11. Martinez-Guanter, J.; Ribeiro, Á.; Peteinatos, G.G.; Pérez-Ruiz, M.; Gerhards, R.; Bengochea-Guevara, J.M.; Machleb, J.; Andújar, D. Low-Cost Three-Dimensional Modeling of Crop Plants. *Sensors* **2019**, *19*, 2883. [CrossRef]

12. Andujar, D.; Calle, M.; Fernandez-Quintanilla, C.; Ribeiro, A.; Dorado, J. Three-dimensional modeling of weed plants using low-cost photogrammetry. *Sensors* **2018**, *18*, 1077. [CrossRef] [PubMed]

13. Rose, J.C.; Paulus, S.; Kuhlmann, H. Accuracy analysis of a multi-view stereo approach for phenotyping of tomato plants at the organ level. *Sensors* **2015**, *15*, 9651–9665. [CrossRef] [PubMed]

14. Zhang, Y.; Teng, P.; Shimizu, Y.; Hosoi, F.; Omasa, K. Estimating 3D leaf and stem shape of nursery paprika plants by a novel multi-camera photography system. *Sensors* **2016**, *16*, 874. [CrossRef] [PubMed]

15. Paulus, S.; Schumann, H.; Kuhlmann, H.; Léon, J. High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants. *Biosyst. Eng.* **2014**, *121*, 1–11. [CrossRef]

16. Yan, T.; Zhu, H.; Sun, L.; Wang, X.; Ling, P. Detection of 3-D objects with a 2-D laser scanning sensor for greenhouse spray applications. *Comput. Electron. Agric.* **2018**, *152*, 363–374. [CrossRef]

17. Reiser, D.; Vázquez-Arellano, M.; Paraforos, D.S.; Garrido-Izard, M.; Griepentrog, H.W. Iterative individual plant clustering in maize with assembled 2D LiDAR data. *Comput. Ind.* **2018**, *99*, 42–52. [CrossRef]

18. Van, V.A.; Tourell, M.C.; Koebernick, N.; Pileio, G.; Roose, T. Correlative visualization of root mucilage degradation using X-ray CT and NMRI. *Front. Environ. Sci.* **2018**, *6*, 32. [CrossRef]

19. Xiang, R.; Jiang, H.; Ying, Y. Recognition of clustered tomatoes based on binocular stereo vision. *Comput. Electron. Agric.* **2014**, *106*, 75–90. [CrossRef]

20. Xiong, X.; Yu, L.; Yang, W.; Liu, M.; Jiang, N.; Wu, D.; Chen, G.; Xiong, L.; Liu, K.; Liu, Q. A high-throughput stereo-imaging system for quantifying rape leaf traits during the seedling stage. *Plant Methods* **2017**, *13*, 7. [CrossRef]

21. Andújar, D.; Ribeiro, A.; Fernández-Quintanilla, C.; Dorado, J. Using depth cameras to extract structural parameters to assess the growth state and yield of cauliflower crops. *Comput. Electron. Agric.* **2016**, *122*, 67–73. [CrossRef]

22. Li, J.; Tang, L. Developing a low-cost 3D plant morphological traits characterization system. *Comput. Electron. Agric.* **2017**, *143*, 1–13. [CrossRef]

23. Su, Q.; Kondo, N.; Li, M.; Sun, H.; Riza, D.F.A.; Habaragamuwa, H. Potato quality grading based on machine vision and 3D shape analysis. *Comput. Electron. Agric.* **2018**, *152*, 261–268. [CrossRef]

24. Hu, Y.; Wang, L.; Xiang, L.; Wu, Q.; Jiang, H. Automatic non-destructive growth measurement of leafy vegetables based on kinect. *Sensors* **2018**, *18*, 806. [CrossRef]

25. Vázquez-Arellano, M.; Reiser, D.; Paraforos, D.S.; Garrido-Izard, M.; Burce, M.E.C.; Griepentrog, H.W. 3-D reconstruction of maize plants using a time-of-flight camera. *Comput. Electron. Agric.* **2018**, *145*, 235–247. [CrossRef]

26. Lin, Y. LiDAR: An important tool for next-generation phenotyping technology of high potential for plant phenomics? *Comput. Electron. Agric.* **2015**, *119*, 61–73. [CrossRef]

27. Thapa, S.; Zhu, F.; Walia, H.; Yu, H.; Ge, Y. A novel LiDAR-based instrument for high-throughput, 3D measurement of morphological traits *Maize* and *Sorghum*. *Sensors* **2018**, *18*, 1187. [CrossRef] [PubMed]

28. Hosoi, F.; Nakabayashi, K.; Omasa, K. 3-D modeling of tomato canopies using a high-resolution portable scanning lidar for extracting structural information. *Sensors* **2011**, *11*, 2166–2174. [CrossRef] [PubMed]

29. George, A.; Michael, L.; Radu, B. Rapid characterization of vegetation structure with a microsoft kinect sensor. *Sensors* **2013**, *13*, 2384–2398. [CrossRef]

30. Cui, J.; Zhang, J.; Sun, G.; Zheng, B. Extraction and Research of Crop Feature Points Based on Computer Vision. *Sensors* **2019**, *19*, 2553. [CrossRef]

31. Vit, A.; Shani, G. Comparing RGB-D Sensors for Close Range Outdoor Agricultural Phenotyping. *Sensors* **2018**, *18*, 4413. [CrossRef]

32. Dionisio, A.; César, F.; José, D. Matching the best viewing angle in depth cameras for biomass estimation based on poplar seedling geometry. *Sensors* **2015**, *15*, 12999–13011. [CrossRef]

33. Pezzuolo, A.; Guarino, M.; Sartori, L.; Marinello, F. A Feasibility Study on the Use of a Structured Light Depth-Camera for Three-Dimensional Body Measurements of Dairy Cows in Free-Stall Barns. *Sensors* **2018**, *18*, 673. [CrossRef] [PubMed]

34. Sun, G.; Wang, X.; Sun, Y.; Ding, Y.; Lu, W. Measurement Method Based on Multispectral Three-Dimensional Imaging for the Chlorophyll Contents of Greenhouse Tomato Plants. *Sensors* **2019**, *19*, 3345. [CrossRef] [PubMed]

35. Hu, P.; Guo, Y.; Li, B.; Zhu, J.; Ma, Y. Three-dimensional reconstruction and its precision evaluation of plant architecture based on multiple view stereo method. *Trans. Chin. Soc. Agric. Eng.* **2015**, *31*, 209–214. [CrossRef]

36. Fang, W.; Feng, H.; Yang, W.; Duan, L.; Chen, G.; Xiong, L.; Liu, Q. High-throughput volumetric reconstruction for 3D wheat plant architecture studies. *J. Innov. Opt. Health Sci.* **2016**, *9*, 1650037. [CrossRef]

37. Brichet, N.; Fournier, C.; Turc, O.; Strauss, O.; Artzet, S.; Pradal, C.; Welcker, C.; Tardieu, F.; Cabrera-Bosquet, L. A robot-assisted imaging pipeline for tracking the growths of maize ear and silks in a high-throughput phenotyping platform. *Plant Methods* **2017**, *13*, 96. [CrossRef] [PubMed]

38. He, J.Q.; Harrison, R.J.; Li, B. A novel 3D imaging system for strawberry phenotyping. *Plant Methods* **2017**, *13*, 93. [CrossRef] [PubMed]

39. Liu, S.; Acosta-Gamboa, L.M.; Huang, X.; Lorence, A. Novel low cost 3D surface model reconstruction system for plant phenotyping. *J. Imaging* **2017**, *3*, 39. [CrossRef]

40. Paul, B.J.; Neil, M.D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [CrossRef]

41. Graham, R.L. An efficient algorith for determining the convex hull of a finite planar set. *Inf. Process. Lett.* **1972**, *1*, 132–133. [CrossRef]

42. Grau, E.; Durrieu, S.; Fournier, R.; Gastellu-Etchegorry, J.P.; Yin, T. Estimation of 3D vegetation density with Terrestrial Laser Scanning data using voxels. A sensitivity analysis of influencing parameters. *Remote Sens. Environ.* **2017**, *191*, 373–388. [CrossRef]

43. Chen, X.; Chen, Y.; Gupta, K.; Zhou, J.; Najjaran, H. SliceNet: A proficient model for real-time 3D shape-based recognition. *Neurocomputing* **2018**, *316*, 144–155. [CrossRef]