

Research Article

Research on Crop 3D Model Reconstruction Based on RGB-D Binocular Vision

Zhiyan Ma , Haidi Wan , and Xiong Gan 

Institute of Agricultural Machinery, Hubei University of Technology, Wuhan 430068, China

Correspondence should be addressed to Zhiyan Ma; 20071017@hbut.edu.cn

Received 3 March 2023; Revised 29 May 2023; Accepted 20 June 2023; Published 14 July 2023

Academic Editor: Tomàs Margalef

Copyright © 2023 Zhiyan Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Taking maize seedlings as the object, the implementation of crops 3D reconstruction based on RGB-D binocular vision and the selection of some key parameters are investigated in this research. First, multiple images are taken from different angles around the target. By mapping the maize seedling region coordinate values after the Otsu algorithm and global threshold segmentation to the corresponding depth image, the depth data of the maize seedling region can be obtained accurately. An improved mean filter is proposed to adaptively fill the holes in the depth image. Then, the different point clouds with the fixed step angle of the maize seedling are registered and fused. Finally, after the fusion point cloud is simplified, the 3D model of crops can be reconstructed. Experimental results show that the simplification effect of the octree algorithm is better than that of the voxel grid filter. Among all the step angles, the reconstruction error of the step angle with 60° is the smallest. Under this condition, the height error between the model and the maize seedling is 2.22%, and the error in stem diameter is 11.67%.

1. Introduction

As technology continues to evolve, the development of smart agriculture has brought many new technologies and solutions to modern agriculture, and 3D reconstruction technology [1, 2] is one of them. 3D reconstruction of crops can be used not only to measure the phenotypic parameters of the target but also to visualize the object in virtual 3D space [3]. Crop 3D reconstruction technology [4, 5] has gradually become a research hotspot in this field. At present, crop 3D reconstruction has several methods, mainly rule-based [6], 3D scanner [7, 8], digitizer [9, 10], and vision-based image reconstruction [11, 12]. Compared with the 3D reconstruction methods based on scanners and digitizers, vision-based 3D reconstruction has the advantages of low cost and a certain accuracy guarantee, which has been extensively studied and applied in recent years [13].

There have emerged many innovative companies and teams in the field of smart agriculture, both domestically and internationally, exploring and innovating in the area of intelligent agriculture. Gibbs et al. [14] proposed an active vision cell, which consists of a camera-mounted robot arm, a combined software interface, and a novel surface reconstruction algorithm. Due to

the active visual framework of the application and the automatic selection of key parameters for surface reconstruction, this pipeline can be applied to any plant species or morphology. Han and Burks [15] took images of the orange canopy from multiple angles with a camera and finally reconstructed the canopy surface by using the Plücker coordinate system. Shirazi et al. [16] used an active stereo vision-based 3D perception system to acquire 3D models of human body parts and surgical tools. To evaluate the proposed system, the article performed a 3D scan of a cardboard box as an object. The results showed errors in height and width measurements to be 9.4 and 23 μm , respectively, compared to the 3D scan results. Peng et al. [17] developed an SFM method based on binocular vision to acquire the physical parameters of plants and constitute the 3D model of the plant. Experimental results show that the mean errors of the measured sizes are all less than 2%. Dong et al. [18] discuss a sea wave measurement method with binocular vision. Using a binocular camera, an accurate 3D model of the ocean waves captured by the camera can be established without contacting the waves themselves. The metric 3D coordinates of the waves can also be measured in both the world coordinate system and the actual camera. The experimental results show that this

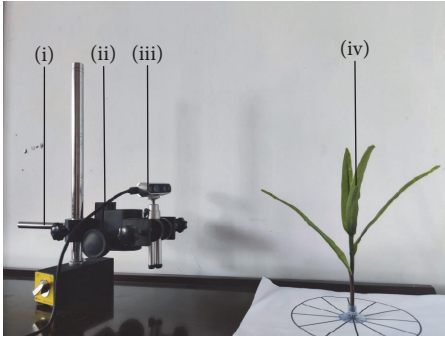


FIGURE 1: Hardware experimental platform for crop 3D reconstruction: (i) camera tripod; (ii) USB cable; (iii) RGB camera; (iv) maize seedling.

method can achieve the expected results. Wu et al. [19] combined deep learning and classical image processing algorithms to calculate the number of banana bunches in two periods and designed a software for estimating the weight of banana fruits during the harvest period. The results show that during the bud removal period, the target segmentation MIoU is 0.878, the average pixel accuracy is 0.936, and the final bunch detection accuracy reaches 86%. During the harvest period, bunch detection is very challenging, with an accuracy of 76%, and the final overall bunch counting accuracy is 93.2%. Tang et al. [20] aim to explore the techniques for improving fruit detection methods in complex environments. For the common types of complex backgrounds in outdoor orchards, the improvement measures are divided into two categories: optimization before and after image sampling. By comparing the results of these methods, the future development trend of fruit detection optimization techniques in complex backgrounds is described.

This paper focuses on the implementation of crop 3D model reconstruction [21] based on RGB-D binocular vision and the selection of some key parameters in the process. Two sets of depth images and RGB images were taken from a fixed step angle around the target by the camera. The 3D point clouds were obtained through preprocessed depth images and RGB images. Furthermore, the point clouds of different angles of maize seedlings were coarsely and finely registered and simplified. The best set of model reconstruction parameters was obtained by comparing the errors of each step angle.

2. Materials and Methods

Vision-based 3D model reconstruction used visual sensors to obtain images, and then the images were processed to obtain a 3D model by a computer [22]. Intel RealSenseD435 RGB-D binocular camera [23] was utilized in the experiment, which has RGB images and depth modules [24]. Two cameras that have the same intrinsic parameters are arranged in parallel with identical focal lengths. The parallax depth image was obtained directly by using the depth module. As shown in Figure 1, the experimental platforms include a PC, binocular camera, experimental crop, and some other auxiliary devices. The software platforms and kits are VS2019, Matlab2019, pcl1.11.0, OpenCV4.5.0, and Cloudcompare2.12.1.

The 3D reconstruction process based on binocular vision was illustrated with the modeling of a maize seedling, including image acquisition, image preprocessing, point cloud construction, and crop modeling. The implementation process is shown in Figure 2.

2.1. Image Acquisition. The depth images and RGB images of the maize seedling were obtained. An origin alignment operation was required because the spatial coordinate system of RGB image and depth image is different. To make a full-view reconstruction of the maize seedling, multiple images from different view angles had to be taken over the maize seedling. The experiment took the method of rotating the target around a fixed camera. Comparative experiments were carried out from four methods of 90° , 60° , 45° , and 30° step angles by using the aligned camera.

2.2. Image Preprocessing. Image preprocessing is mainly prepared for the acquisition of point clouds. By mapping the maize seedling region coordinate value after the Otsu algorithm [25] and global threshold segmentation to the corresponding depth image, the depth data of the maize seedling region could be obtained accurately. The impact of the environment might cause some data to be missed in the depth image. In order to reduce it, an improved zero-point culling mean filter was proposed to adaptively fill the holes in the depth image.

2.3. Point Cloud Construction. After image preprocessing, the 3D point cloud in each angle could be recovered by converting all pixels in the crop region and their corresponding depth values.

2.4. Crop Modeling. The binocular camera could not get complete information about the maize seedling at an angle due to the inevitable shielding of blades. Therefore, the experiment needed to register and fuse the point clouds of different angles [26, 27]. It was necessary to simplify and smooth the fused point clouds due to the large number of points. Finally, the final point clouds could be reconstructed through the greedy projection triangulation algorithm [28, 29].

3. Implementation and Results of 3D Model Reconstruction

3.1. Image Acquisition. Both RGB images and depth images are obtained by an RGB-D camera. In order to reduce the influence of external factors, the images of maize seedlings are taken in front of a white wall. The camera is about half a meter away from the maize seedling. The maize seedling is placed on a plate with an eighth-equal circle drawing. Then, a set of RGB images and depth images are captured from eight angles by rotating maize seedling with a fixed camera, and the image size is 640×480 .

The 45° step angle is taken as an example, and the RGB images captured at 0° , 45° , 90° , 135° , 180° , 225° , 270° , and 315° angles are shown in Figures 3(a), 3(b), 3(c), 3(d), 3(e), 3(f), 3(g), and 3(h), respectively.

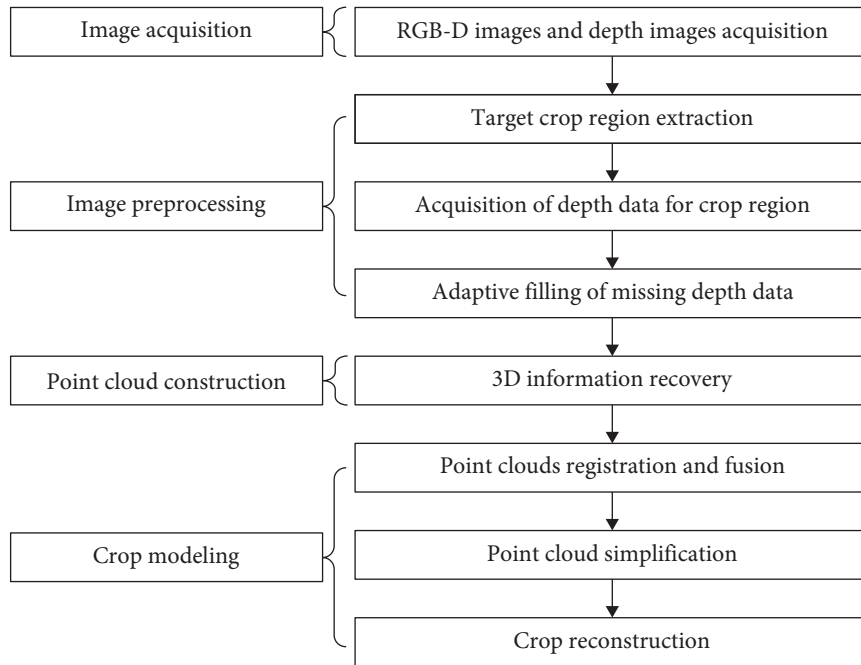


FIGURE 2: Vision-based 3D reconstruction process of the crop.

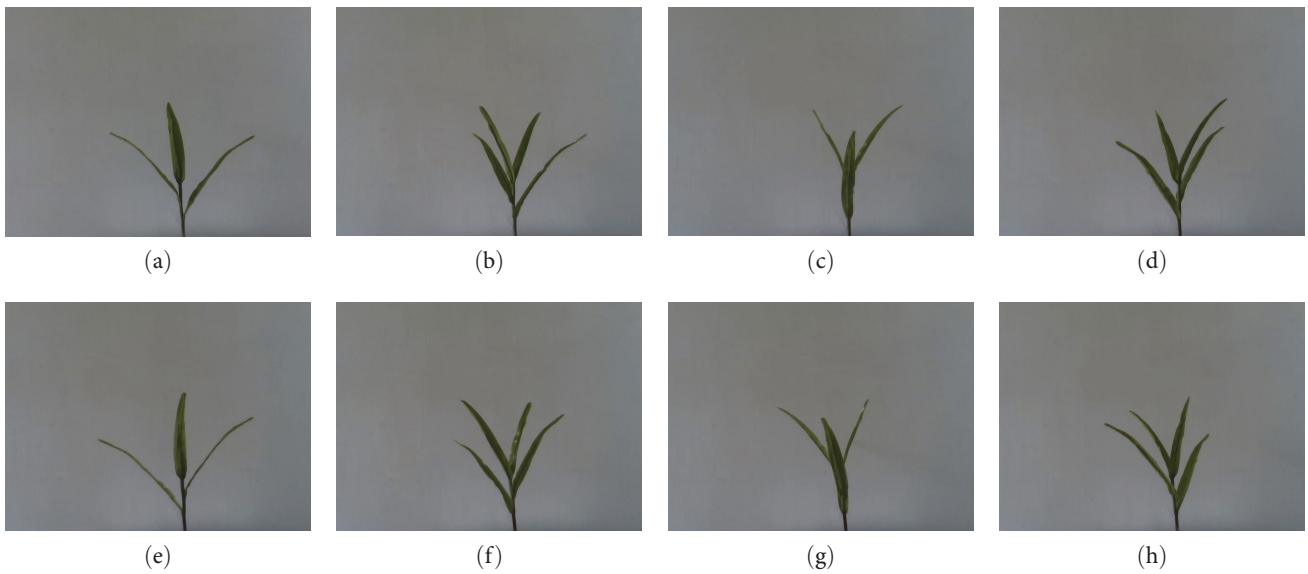


FIGURE 3: RGB images of maize seedlings in 45° step angle: (a) the RGB image captured at a 0° angle; (b) the RGB image captured at a 45° angle; (c) the RGB image captured at a 90° angle; (d) the RGB image captured at a 135° angle; (e) the RGB image captured at a 180° angle; (f) the RGB image captured at a 225° angle; (g) the RGB image captured at a 270° angle; (h) the RGB image captured at a 315° angle.

3.2. Image Preprocessing

3.2.1. *Target Crop Region Extraction.* Otsu algorithm selects the global optimal threshold by maximizing the between-class variance. The pixels of image are divided into two classes C_0 with gray levels $[0, 1, 2, \dots, t]$ and C_1 with gray levels $[t + 1, \dots, M - 1]$ by the threshold t . The gray-level probability distributions for the two classes are as follows:

$$\omega_0 = \sum_{i=0}^t P_i, \tag{1}$$

$$\omega_1 = \sum_{i=t+1}^{M-1} P_i, \tag{2}$$

$$\omega_0 + \omega_1 = 1. \tag{3}$$

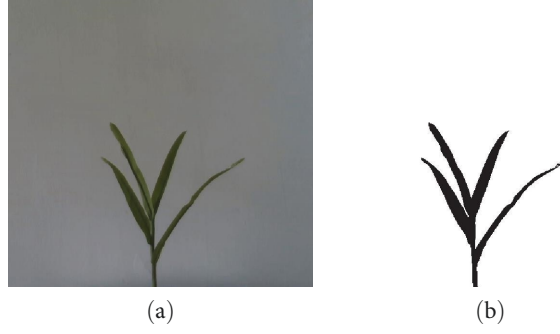


FIGURE 4: Target region extraction of RGB image: (a) RGB image; (b) image after global threshold segmentation.

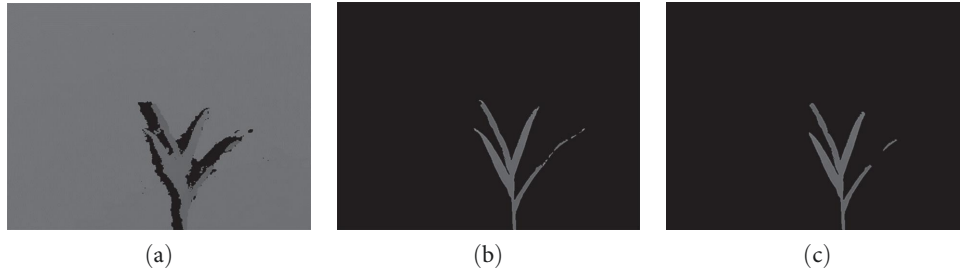


FIGURE 5: Acquisition of depth data for crop region: (a) the raw depth data; (b) the depth data in the target region; (c) the depth data after removing error points.

The means of class C_0 and C_1 are

$$\mu_0 = \sum_{i=0}^t i \times Pr(i|\omega_0) = \frac{1}{\omega_0} \sum_{i=0}^t i \times P_i, \quad (4)$$

$$\mu_1 = \sum_{i=t+1}^{M-1} i \times Pr(i|\omega_1) = \frac{1}{\omega_1} \sum_{i=t+1}^{M-1} i \times P_i. \quad (5)$$

The total mean of gray levels is denoted by μ

$$\mu = \omega_0 \times \mu_0 + \omega_1 \times \mu_1. \quad (6)$$

The between-class variance is as follows:

$$\sigma^2 = \omega_0 \times (\mu_0 - \mu)^2 + \omega_1 \times (\mu_1 - \mu)^2. \quad (7)$$

The optimal threshold t is chosen by maximizing the between-class variance.

$$t = \arg_{t \in \{0, 255\}} \max(\sigma^2(t)). \quad (8)$$

The RGB image before segmentation is represented as Figure 4(a). After obtaining the best threshold t , it is easy to extract the maize seedling region with the global threshold. The pixels lower than the threshold value are set to zero, and the remaining pixels are set to 255, as shown in Figure 4(b).

$$g(x, y) = \begin{cases} 0, & f(x, y) \leq t \\ 255, & f(x, y) > t. \end{cases} \quad (9)$$

3.2.2. Acquisition of Depth Data for Crop Region. The original depth image is represented as Figure 5(a). The maize seedling segmentation coordinates completed in the binary image are mapped to the corresponding step angle depth image, and the depth data for the maize seedling region can be obtained. At the same time, the depth values of the background are set to zero, and only the depth values in the crop region are retained, as shown in Figure 5(b). Due to the accuracy errors of the camera, the depth values of some pixels of the crop are incorrect in the depth image. Thus, the wrong depth values are set to zero by finding a cutoff value, as shown in Figure 5(c).

3.2.3. Adaptive Filling of Missing Depth Data. Because the depth value of the pixels on the template may be zero, using the mean filter in the experiment directly will cause bigger errors in the process of hole-filling. Furthermore, the depth of other nonhole pixels will also be changed. In this paper, an improved mean filter is proposed to fill such holes.

In the traditional mean filtering, the gray value $g(x_i, y_j)$ of the current point is as follows:

$$\begin{aligned} g(x_i, y_j) &= \frac{1}{(2n+1)^2} \sum_{\xi, \eta} f(x_{i+\xi}, y_{j+\eta}) \\ &= \frac{1}{9} [f(x_i + y_j) + f(x_{i+1} + y_{j+1}) + f(x_i + y_{j+1}) \\ &\quad + f(x_{i+1} + y_j) + f(x_i + y_{j-1}) + f(x_{i-1} + y_j) \\ &\quad + f(x_{i+1} + y_{j-1}) + f(x_{i-1} + y_{j-1}) + f(x_{i-1} + y_{j+1})]. \end{aligned} \quad (10)$$

In the improved mean filtering, the number of zero points in the template is denoted by t , the gray value $g(x_i, y_j)$ of the zero point is as follows:

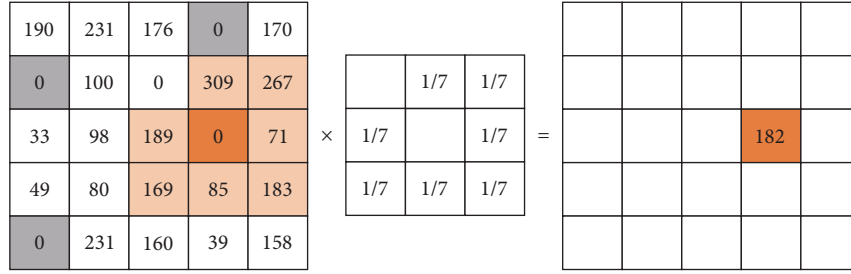


FIGURE 6: Adaptive filling of the improved mean filter.



FIGURE 7: Comparison of hole-filling effect in-depth image: (a) depth image by mean filter; (b) depth image by improved mean filter.

$$\begin{cases} \text{if } : g(x_i, y_i) = 0 \\ g(x_i, y_j) = \frac{1}{(2n + 1)^2 - t_{\xi, \eta}} \sum_{\xi, \eta} f(x_{i+\xi}, y_{j+\eta}) \end{cases} \quad (11)$$

The depth image of the region where the target is identified is traversed. As shown in Figure 6, first, hole pixels with zero depth are detected and removed from the template of the hole pixel, and a new value is assigned to the hole pixel by using the average of the remaining pixels, which considerably reduces the errors.

As shown in Figure 7(a), after the mean filtering, the depth image becomes blurred, and some pixels in the non-target region are filled with values. So, hole-filling by mean filter does not work well. Compared with the method of mean filter, the improved adaptive mean filter effectively fills the holes without affecting the depth of other pixels and protects the image details very well, as shown in Figure 7(b).

3.3. Point Cloud Construction. All the target 3D point coordinates construct their point cloud. They can be calculated directly from the depth and RGB images. This process is the conversion from a 2D depth image to a 3D point cloud. And the transformation relationship of Formula (1) can be shown as follows:

$$\begin{cases} p \cdot z = \text{double}(d)/\text{factor} \\ p \cdot x = (\text{col} - cx) \times p \cdot z/fx \\ p \cdot y = (\text{row} - cy) \times p \cdot z/fx \end{cases} \quad (12)$$

(fx, fy): The focal length of the camera in the directions;
 factor: The scale factor of the camera, take 1.0;
 (cx, cy): The center of the camera's aperture;



FIGURE 8: A set of point cloud data of maize seedling.

(col, row): The coordinate of pixel in the depth image;
 double (d): The depth value of the (col, row) pixel in the depth image;
 Figure 8 shows the point cloud of a processed depth image.

3.4. Crop Modeling

3.4.1. Point Clouds Registration and Fusion. The point cloud from one angle provides only partial information about the maize seedling. The images taken from the adjacent locations have an overlapping region. In order to obtain the complete information of the maize seedling, these point clouds from the same step angle should be registered to a unified coordinate system. There is a fixed angle difference between the point clouds of different angles of the same step in the experiment. However, the distance and position of the two-point clouds in fine registration should be short enough. The registration of point clouds includes two steps: coarse registration and fine registration. The two-point clouds are spatially close to each other after coarse registration. Then the coarsely registered point clouds can be fine-registered by using iterative closest point (ICP) algorithm with lower errors.

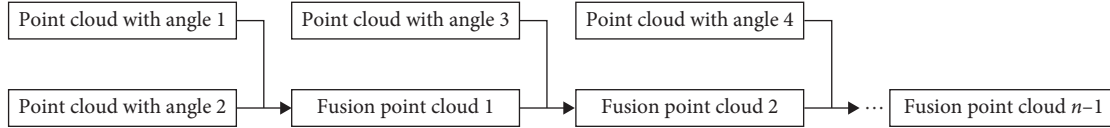


FIGURE 9: The process of point clouds registration and fusion.

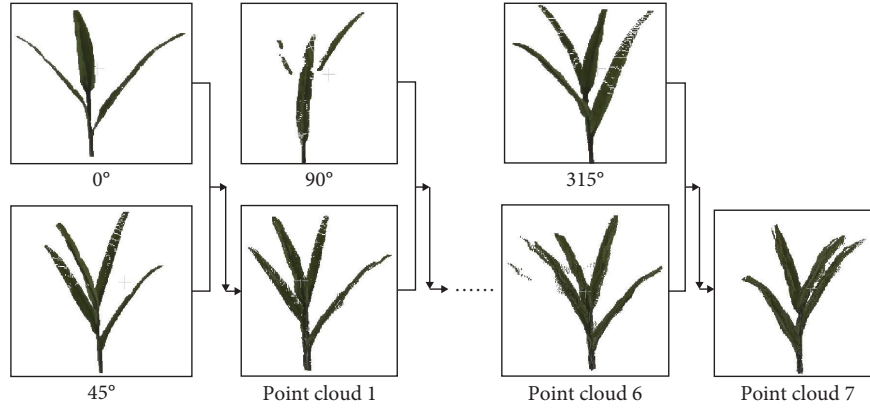


FIGURE 10: Point clouds registration process.

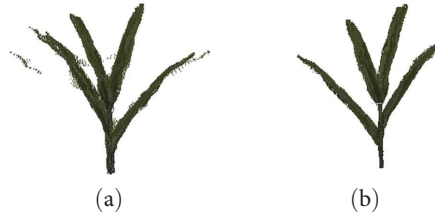


FIGURE 11: The clipped point cloud: (a) before the clipped; (b) after the clipped.

The process of registration and fusion is shown in Figure 9. The source clouds are from the depth images of n divide the circumference angle equally ($n=4, 6, 8, 12$) that have been preprocessed. Taking the 45° step angle as an example, the eight-point clouds from 0° to 315° are respectively referred to as angle 1, angle 2, ..., and angle 8. Point cloud with angle 1 and point cloud with angle 2 are registered and fused into a new point cloud which calls fusion point cloud 1, and then fusion point cloud 1 and point cloud with angle 3 are processed by coarse registration, fine registration, and fusion. As shown in Figure 10, at least fusion point cloud 7 is obtained by sequentially merging the eight-point clouds.

In the process of point cloud registration [26, 27], the key point is to solve the transformation matrix parameters between the two point clouds, and the important prerequisite for solving the transformation matrix parameters is to obtain the corresponding point pair sets. Coarse registration manually selects at least three pairs of corresponding feature points as registration primitives and calculates the registration transformation matrix. ICP registration iterates from an initial value until the parameter sequence converges to meet the

minimum requirement of the objective function. ICP algorithm is a nonlinear least squares problem. Selecting the appropriate threshold, numbers of iterations, and point cloud overlap to obtain the transformation matrix.

As shown in Figure 11(a), the final point cloud usually has some obvious errors due to camera depth errors and point cloud registration errors. It is necessary to manually clip and delete the data with depth errors in the registered point clouds to reduce the measurement errors. The processed point cloud is shown in Figure 11(b).

3.4.2. Point Cloud Simplification. In the experiment, the final point cloud has more than 50,000 points after registration and fusion. The point cloud includes many redundant points that are not needed for measurement. Two typical voxel methods, VoxelGrid, and octree, are selected to simplify the point cloud for comparison.

Because the point cloud is unordered, we find the maximum and minimum values of the point cloud in the X , Y , and Z axes to construct the minimum space-bounding volume cube to partition the point cloud. The octree divides the minimum space bounding volume cube into eight nodes

TABLE 1: Octree-based point cloud simplification algorithm.

Algorithm 1 Octree-based point cloud simplification algorithm

Input: Original point cloud (P)
Output: Point cloud of Octree simplification (O)

1. Set the octree resolution (voxel size)
2. Divide O into voxel cubes (C) by voxel segmentation until the specified voxel size is reached
3. Calculate the voxel center of C
4. Traverse all the voxel centers
5. For: $i = 0, 1, \dots, \text{voxel_centers.size}() - 1$
6. Construct a new point cloud by voxel centers

TABLE 2: VoxelGrid-based point cloud simplification algorithm.

Algorithm 2 VoxelGrid-based point cloud simplification algorithm

Input: Original point cloud (P)
Output: Point cloud of VoxelGrid simplification (V)

1. Set the VoxelGrid resolution (voxel size)
2. Create the filtering object: sor
3. Start filtering: sor.filter

Divide P into point cloud blocks (B) by VoxelGrid filter
Calculate the barycenter of B

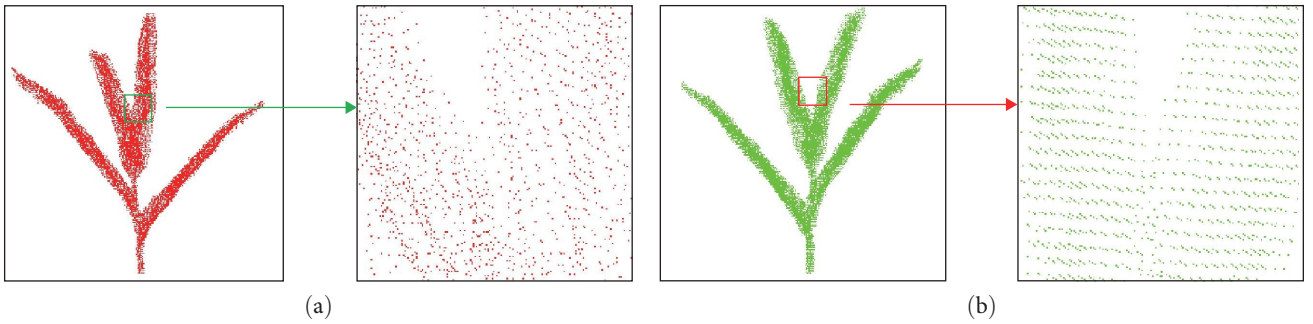


FIGURE 12: Comparison of two methods for point cloud simplification: (a) VoxelGrid simplification result and local detail; (b) Octree simplification result and local detail.

averagely, and each node is divided based on this rule until the specified division level is reached. The points within each voxel are represented by the voxel centers. With the octree method, the simplified point cloud consists of voxel centers, as shown in Table 1.

The VoxelGrid simplification method can divide the point cloud into small cloud blocks. At the same time, all points in the cloud blocks are replaced with the barycenter, as shown in Table 2.

As shown in Figures 12(a) and 12(b), the point cloud is more regular after the octree method. Therefore, the method of octree voxel center approximation is used to simplify the scattered point cloud.

3.4.3. Crop Construction. Finally, the greedy projection-based triangulation algorithm [28] is used to reconstruct the final crop model. The greedy projection triangulation algorithm is an algorithm that quickly establishes the topological relationship of point cloud through the triangular grid structure. Greedy projection triangulation is an algorithm to quickly triangulate the original point cloud. The overall optimal result is obtained by local optimal processing when the problem is solved using a greedy algorithm [29] based on growth.

In 3D space of the point cloud, the topology cannot be directly established, so the point cloud needs to be projected to a 2D plane, and then a topological relationship is established on it. The point cloud in 3D space and K neighbors of

the point cloud are projected into a plane; that is, it maps the point cloud from 3D to 2D, and it uses the 2D Delaunay triangulation growth algorithm to connect the triangular grids for the 2D point cloud mapped to the plane.

Assuming the normal vector of the point $N_0(x_0, y_0, z_0)$ is $m = (A, B, C)$, the tangent plane at point $N(x, y, z)$ crossing N_0 is $\cos \alpha$.

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \quad (13)$$

In order to project points in space into the 2D tangent plane Π , the projection matrix method is generally used. The projection of the 3D points on the tangent plane is obtained by a series of operations such as translation and rotation.

The projection matrix $T_{M_{\Pi}}$ is as follows:

$$T_{M_{\Pi}} = T_c \cdot R_x \cdot R_y. \quad (14)$$

The translation transformation matrix T_c is as follows:

$$T_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ X_0 & Y_0 & Z_0 & 1 \end{bmatrix}. \quad (15)$$

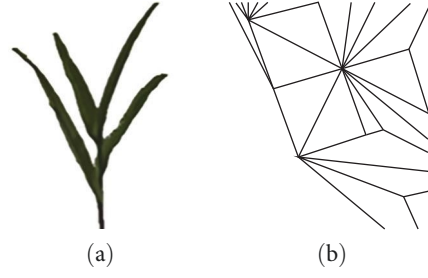


FIGURE 13: 3D reconstruction result and detail: (a) 3D rendering effect of the reconstructed seedling corn; (b) local enlargement of the mesh.

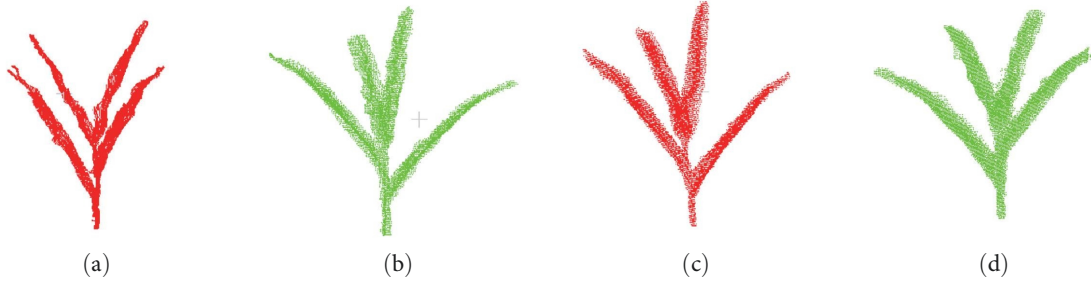


FIGURE 14: Final point cloud of four step angles: (a) the final point clouds of the step of 90°; (b) the final point clouds of the step of 60°, (c) the final point clouds of the step of 45°; (d) the final point clouds of the step of and 30°.

R_x represents a α degree rotation around x -axis.

$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (16)$$

R_y represents a θ degree rotation around y -axis.

$$R_y = \begin{bmatrix} \cos \theta & 0 & -\sin \theta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \theta & 0 & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (17)$$

The projection of the point $q(x_i, y_i, z_i)$ on the tangent plane Π is as follows:

$$[x'_i, y'_i, z'_i, 1] = T_{M_{\Pi}} \cdot [x_i, y_i, z_i, 1]^T. \quad (18)$$

The method selects a sample triangular piece as the initial surface, then continuously expands the boundary of the surface, and finally constructs a complete triangular mesh surface so that the point cloud in 2D space has a topological relationship. Then the topological relationship of the 2D point cloud is mapped to the original 3D point cloud, and the topological relationship of the original point cloud 3D space is constructed. Until a complete mesh topological relationship is constructed, the 3D reconstruction of the object surface is realized, as shown in Figures 13(a) and 13(b).

4. Discussion

Repeating the above steps, Figure 14(a)–14(d) corresponds to the final point clouds with the step of 90°, 60°, 45°, and 30°. The point cloud data are incomplete, the phenotypic information of maize seedling is missing, and the error is relatively large in the scheme of 90° step angle. The phenotypic information of maize seedling at 60°, 45°, and 30° step angles are relatively complete.

Phenotypic measurements of 60°, 45°, and 30° step angles are compared. The plant height and stem diameter of maize seedling are measured. Then the experimental results are compared with the manual measurements, as shown in Table 3.

5. Conclusions

- (1) In this paper, the implementation of crops 3D reconstruction based on RGB-D binocular vision and the selection of some key parameters are investigated. The 3D point clouds are obtained through pre-processed depth images and RGB images. Furthermore, the point clouds are coarsely and finely registered and simplified. Finally, the greedy projection-based triangulation algorithm is used to reconstruct the final crop model. The best set of model reconstruction parameters was obtained by comparing the errors of each step angle.
- (2) An improved adaptive mean filter is proposed by removing the zero point of the template in the dynamic mean filter; the holes of the depth image in the maize seedling region are filled adaptively. Compared with the method of traditional mean filter,

TABLE 3: Phenotype measurements and errors calculation of maize seedling.

Measurement object	Step angle (°)	Manual measurement (cm)	System measurement (cm)	Relative error (%)
Overall height	60	20.20	20.65	2.22
	45		19.78	2.07
	30		20.60	1.98
Stem diameter angle 1	60	0.6	0.53	11.67
	45		0.7	16.67
	30		0.65	8.33
Stem diameter angle 2	60	0.6	0.53	11.67
	45		0.52	13.30
	30		1.00	66.67

the improved adaptive mean filter effectively fills the hole without affecting the depth of other pixels.

- (3) By comparing four step angles schemes, the phenotypic information of the 90° step angle scheme is seriously missing, and the stem diameter error of the 30° step angle is serious. Compared with the 45° step angle, the error of the 60° step angle is relatively small; therefore, the 60° step angle of the four 3D reconstruction schemes in the experiment has the highest accuracy. Under this condition, the height error between the model and the maize seedling is 2.22%, and the error of stem diameter is 11.67%. The research needs of 3D reconstruction of maize seedling can be satisfied.
- (4) The 3D reconstruction method discussed in this article is effective for modeling single plants under simple experimental conditions. However, when it comes to complex population crops in field conditions, its effectiveness is limited due to the presence of textures or colors in the background that is similar to those of the crops, as well as depth errors caused by lighting and occlusion. Going forward, research will focus on exploring 3D reconstruction techniques for population crops in complex growth environments to gain insights into crop growth.

While the 3D reconstruction method described in this paper can reconstruct a model for a specific moment in a crop's growth cycle, it has not yet achieved full modeling of the entire growth process of the target crop or established a corresponding relationship with the crop growth model. Subsequent research will be required to achieve full 3D reconstruction throughout the plant's lifecycle. This entails further development of a crop's 3D model based on time-series data combined with plant growth rules, which will aid agricultural managers in monitoring and managing crops more efficiently.

Data Availability

The data used to support the findings of this study can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Key R&D Program subproject of China (grant no. 2018YFD0701002-03).

References

- [1] R. Zhu, S. Li, Y. Sun et al., "Research advances and prospects of crop 3D reconstruction technology," *Smart Agriculture*, vol. 3, pp. 94–115, 2021.
- [2] X. Chen, Q. Wu, and S. Wang, "Research on 3D reconstruction based on multiple views," in *2018 13th International Conference on Computer Science & Education (ICCSE)*, pp. 1–5, IEEE, Colombo, Sri Lanka, 2018.
- [3] M. G. Kantarci, B. Gökberk, and L. Akarun, "A Survey of 3D object reconstruction methods," in *2022 30th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4, IEEE, Safranbolu, Turkey, 2022.
- [4] B. Gowda, B. C. Wünsche, and A. Shaw, "Reconstruction or retrieval? Investigating neural 3D reconstruction," in *36th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 1–6, IEEE, Tauranga, New Zealand, 2021.
- [5] Y. Yin, H. C. Zhang, and J. Q. Zheng, "Three dimensional reconstruction system of plant based on binocular stereo vision," *Journal of Chinese Agricultural Mechanization*, vol. 42, no. 3, pp. 129–135, 2021.
- [6] L. Xie, H. Hu, Q. Zhu et al., "Combined rule-based and hypothesis-based method for building model reconstruction from photogrammetric point clouds," *Remote Sensing*, vol. 13, no. 6, Article ID 1107, 2021.
- [7] V. V. Gumaste, R. Deepa, K. R. V. Kumari, and V. Kavitha, "Optical tracking system for auto-registration of multiple views of a 3D scanner for 3D shape reconstruction of components," in *2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET)*, pp. 1–6, IEEE, Patna, India, 2022.
- [8] A. Dawda and M. Nguyen, "Defects detection in highly specular surface using a combination of stereo and laser reconstruction," in *2020 35th International Conference on*

- Image and Vision Computing New Zealand (IVCNZ)*, pp. 1–6, IEEE, Wellington, New Zealand, 2020.
- [9] W. S. Chowdhury, G. Lu, and M. M. Hossain, “Three-dimensional reconstruction and measurement of avian eggs through digital imaging,” in *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, pp. 1–5, IEEE, Dubrovnik, Croatia, 2020.
- [10] G. Sansoni, C. Cattaneo, M. Trebeschi et al., “Scene-of-crime analysis by a 3-dimensional optical digitizer: a useful perspective for forensic science,” *The American journal of forensic medicine and pathology*, vol. 32, no. 3, pp. 280–286, 2011.
- [11] W. Oh, H. Yoo, T. Ha, and S. Oh, “Vision-based 3D reconstruction using a compound eye camera,” in *2021 21st International Conference on Control, Automation and Systems (ICCAS)*, pp. 418–423, IEEE, 2021.
- [12] X. Lin, J. Wang, and C. Lin, “Research on 3D reconstruction in binocular stereo vision based on feature point matching method,” in *2020 IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pp. 551–556, IEEE, Dalian, China, 2020.
- [13] H. V. Le, “A structure-from-motion method for 3-D reconstruction of moving objects from multiple-view image sequences,” in *2004 International Conference on Image Processing, 2004. ICIP '04*, vol. 3, pp. 1955–1958, IEEE, Singapore, 2004.
- [14] J. A. Gibbs, M. P. Pound, A. P. French, D. M. Wells, E. H. Murchie, and T. P. Pridmore, “Active vision and surface reconstruction for 3D plant shoot modelling,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 6, pp. 1907–1917, 2020.
- [15] S. H. Han and T. F. Burks, “3D reconstruction of a citrus canopy,” in *Proceedings of the 2009*, American Society of Agricultural and Biological Engineers, Reno, Nevada, 2009.
- [16] M. A. Shirazi, D. Khan, M. Affan, H. A. Poonja, M. S. A. Shah, and R. Uddin, “Active stereo vision based 3D reconstruction for image guided surgery,” in *2021 International Conference on Robotics and Automation in Industry (ICRAI)*, pp. 1–5, IEEE, Rawalpindi, Pakistan, 2021.
- [17] Y. Peng, M. Yang, G. Zhao, and G. Cao, “Binocular-vision-based structure from motion for 3-D reconstruction of plants,” *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, Article ID 8019505, 2022.
- [18] X. Dong, L. Duan, L. Niu, Y. An, and L. Gong, “Binocular vision based 3D reconstruction of ocean waves and position coordinate measurement,” in *2022 International Symposium on Electrical, Electronics and Information Engineering (ISEEIE)*, pp. 177–182, ISEEIE, Chiang Mai, Thailand, 2022.
- [19] F. Wu, Z. Yang, X. Mo et al., “Detection and counting of banana bunches by integrating deep learning and classic image-processing algorithms,” *Computers and Electronics in Agriculture*, vol. 209, Article ID 107827, 2023.
- [20] Y. Tang, J. Qiu, Y. Zhang et al., “Optimization strategies of fruit detection to overcome the challenge of unstructured background in field orchard environment: a review,” *Precision Agriculture*, vol. 24, pp. 1183–1219, 2023.
- [21] X. Tian, R. Liu, Z. Wang, and J. Ma, “High quality 3D reconstruction based on fusion of polarization imaging and binocular stereo vision,” *Information Fusion*, vol. 77, pp. 19–28, 2022.
- [22] Y. Tang, M. Chen, X. Wu et al., “Three-dimensional reconstruction and monitoring of large-scale structures via real-time multi-vision system,” in *Artificial Intelligence Algorithms and Applications*, K. Li, W. Li, H. Wang, and Y. Liu, Eds., vol. 1205 of *Communications in Computer and Information Science*, pp. 442–457, Springer, Singapore, 2020.
- [23] T. X. Zhang, S. Huang, Y. F. Li, and M. C. Fu, “Key techniques for vision based 3D reconstruction: a review,” *Acta Automatica Sinica*, no. 46, pp. 631–652, 2020.
- [24] X. S. Li, K. Y. Qin, P. Yao, J. Yu, W. J. Wu, and L. Chen, “3D surface reconstruction based on binocular vision,” in *2014 IEEE International Conference on Mechatronics and Automation*, pp. 1861–1865, Tianjin, China, 2014.
- [25] L. Xiao, H. Ouyang, and C. Fan, “Otsu’s thresholding method based on plane intercept histogram and geometric analysis,” *The International Arab Journal of Information Technology*, vol. 17, no. 5, 2020.
- [26] J. Liu, P. Guo, and X. Sun, “An automatic 3D point cloud registration method based on biological vision,” *Applied Sciences*, vol. 11, no. 10, Article ID 4538, 2021.
- [27] P. Yu, Y. Yang, A. Tian et al., “An improved ICP point cloud registration algorithm based on three-points congruent sets,” in *2021 3rd International Conference on Artificial Intelligence and Advanced Manufacture (AIAM)*, pp. 407–411, IEEE, Manchester, United Kingdom, 2021.
- [28] J. Liu, D. Bai, and L. Chen, “3-D point cloud registration algorithm based on greedy projection triangulation,” *Applied Sciences*, vol. 8, no. 10, Article ID 1776, 2018.
- [29] L. Huang, S. Li, A. Zhu, X. Fan, C. Zhang, and H. Wang, “Non-contact body measurement for Qinchuan cattle with LiDAR sensor,” *Sensors*, vol. 18, no. 9, Article ID 3014, 2018.