2024 12th International Conference on Information Systems and Computing Technology (ISCTech) | 979-8-3503-7986-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISCTech63666.2024.10845263

Phenotypic Parameters Extraction of *Panax notoginseng* with NeRF

1st Ruoxi Wang
Faculty of Modern Agricultural Engineering
Kunming University of Science and Technology
Kunming, China
wrx2102022@163.com

2nd Liang Dai
Faculty of Modern Agricultural Engineering
Kunming University of Science and Technology
Kunming, China
947730301@qq.com

3rd Qiliang Yang
Faculty of Modern Agricultural Engineering
Kunming University of Science and Technology
Kunming, China
yangqilianglovena@163.com

4th Ling Yang*
Faculty of Information Engineering and Automation
Kunming University of Science and Technology
Kunming, China
yangling@kust.edu.cn

***Abstract*—The growth stage significantly influences both the yield and quality of *Panax notoginseng*. Accurate plant phenotypic parameters are crucial for the precise management of *P. notoginseng* cultivation. Currently, traditional methods like manual measurements and tools are used for collecting phenotypic information, but they are inefficient and costly. Manual measurements are prone to subjective biases and can potentially harm plants irreversibly. While 2D images can capture plant phenotypic information, they often suffer from incomplete data and lack precision, especially for plants with complex structures like *P. notoginseng*. This limitation makes it challenging to achieve comprehensive and accurate measurements. To address these challenges, this study proposes a novel approach using a Neural Radiance Field (NeRF) for extracting phenotypic parameters from *P. notoginseng*. By capturing video and multi-view image sequences of *P. notoginseng*, we were able to achieve high-fidelity 3D rendering of the plants and extract point cloud data from them. This approach enabled accurate measurement of plant height and leaf area parameters. The results demonstrate promising accuracy, with an average percentage error of 1.76% for plant height and 1.73% for leaf area based on the point cloud measurements obtained using NeRF. This method leverages advanced computational techniques to overcome the limitations of traditional 2D imaging methods, offering a more comprehensive and precise means of phenotypic characterization for complex plant structures like *P. notoginseng*.**

Keywords-component; Panax notoginseng; Image Sequence; Neural Radiation Field; 3D reconstruction; plant phenotype

I. INTRODUCTION

Plant phenotypic information (plant height, leaf area, etc.) is an important indicator to reflect the current growth status of plants [1]. By obtaining and analyzing this phenotypic information, growers can adjust their planting strategies in time to effectively improve the yield of crops. Therefore, rapidly and accurately acquiring phenotypic parameters of *Panax notoginseng* plants is crucial for achieving precise management of *P. notoginseng* cultivation. Currently, the collection of plant phenotypic information is mainly realized by traditional manual measurements and diverse measurement tools, which have limited efficiency and high cost. Manual measurement is easily affected by individual subjectivity and is limited to a small range each time, which is not conducive to obtaining a wide range of phenotypic parameters [2], [3]. With the development of imaging sensors and the increasing progress of image processing technology, new opportunities have been brought to plant phenotype technology. In recent years, many researchers have used the plant phenotype technology of two-dimensional images to monitor the phenotype of maise, apple and tomato and achieved good performance [4], [5], [6]. However, more research on two-dimensional technology is in the monitoring and classification of plant phenotypes, and there are few studies on the morphological phenotypes of plants. Moreover, the two-dimensional image itself is limited in processing the spatial structure information of plants. *P. notoginseng* is a multi-leaf plant with a complex shape structure. It is difficult for 2D images to capture plants' actual geometric and spatial characteristics accurately. To solve this problem, deep learning-based Neural Radiance Fields (NeRF) [7] have been widely used in 3D reconstruction in recent years. NeRF can reconstruct high-fidelity, detail-rich 3D scenes from sparse image sets by training a neural network to learn the scene's continuous volumetric density function. It performs well and remains robust even under complex lighting and occlusion. NeRF generates smooth and continuous volumetric renderings, providing a more realistic visual experience, especially when dealing with dynamic scenes and complex geometries.

Since 2020, NeRF has advanced viewpoint synthesis techniques in computer vision and graphics. Initially introduced by Mildenhall et al., NeRF leverages volume rendering to synthesize high-quality images from new viewpoints. This method samples five-dimensional coordinates along camera rays to map color and density information onto the image plane [7]. Building on NeRF's foundation, subsequent research has expanded its capabilities. Zhang et al. developed NeRF++, focusing on optimizing 360-degree object capture [8]. At the same time, Wang et al. simplified NeRF training for frontal scenes, enabling high-quality 3D reconstruction from 2D images without prior camera parameter knowledge [9]. In 2022, Tancik et al. introduced Block-NeRF to facilitate 3D reconstruction of large-scale environments [10], while Müller et al. proposed Instant-NGP, which reduces network size and computational cost through multi-resolution hash coding of feature vectors [11]. The latest advancement includes NeRF-Ag by Zhang, based on implicit neural

representation [12]. This model generates high-precision 3D mesh models and photorealistic rendering of multi-scale scenes like orchards. It improves texture detail representation and modelling accuracy compared to traditional methods, showcasing significant progress in realistic scene synthesis and representation.

We propose a method for 3D reconstruction of *P. notoginseng* plants using NeRF technology to solve the problem of missing information in 2D images due to branch and leaf occlusion and self-occlusion. By extracting the point cloud from the 3D rendering model, we aim to obtain the 3D information of *P. notoginseng* plants quickly and efficiently, thus enhancing the accuracy of its phenotypic data measurement. The main contributions of this paper are as follows:

- We successfully proposed a method to obtain phenotypic parameters of *P. notoginseng* using three-dimensional reconstruction techniques.
- We propose a new NeRF-based method for extracting phenotypic parameters of *P. notoginseng* plants. After 3D rendering by NeRF, point clouds are then extracted from the rendered model and utilized to achieve accurate measurements of *P. notoginseng* phenotypic parameters, which significantly improves the accuracy compared to traditional 3D methods.

II. DATA COLLECTION

We acquired image data from the *P. notoginseng* plantation of Kunming University of Science and Technology and used RGB camera and video frame extraction methods for image collection. We chose the perspective of tilting 45° to shoot overhead and controlled the angle every 5°, 10°, 15°, 20°, 30°, 45°, and 60° with a 360° dial. We acquired images using a video frame extraction script for the other angles. After the acquisition, we checked and removed poor-quality images, converted them uniformly to JPG format and maintained a consistent resolution.

III. METHODS

A. NeRF Basic Model

Neural Radiance Fields (NeRF) is a deep learning-based technique for 3D scene reconstruction and new perspective synthesis. It utilizes a Multilayer Perceptron (MLP)[13] network to learn an implicit scene representation, enabling high-quality image rendering with new perspectives. During training, NeRF converts 3D spatial coordinates into feature vectors via position coding to enhance the MLP network's ability to process 3D coordinates and capture high-frequency details. NeRF uses volumetric rendering techniques to simulate light propagation and employs the MLP network to predict the color and density of each sample point by integrating and estimating the color of image pixels. Once trained, the MLP network can render the scene in a new perspective without the traditional explicit geometry reconstruction step.

To further enhance the rendering quality, NeRF can perform subsequent processes such as denoising, sharpening, and beam method levelling to optimize the detail performance and improve the accuracy and visual realism of the rendered results. These steps enable NeRF to provide a more detailed and realistic representation of the 3D scene. As shown in Fig. 1. NeRF uses the volume rendering formula. In essence, NeRF uses its MLP network H_θ to describe the mapping space coordinates (x, y, z) and the corresponding observation direction vector d between the density a of each point in the 3D scene and the directional emission color $c = (r, g, b)$. Establish a mapping relationship between density and directional emission color from spatial coordinates and observation direction to points. The relationship can be expressed as:

$$H_\theta(x, y, z, d) \rightarrow (c, \sigma) \quad (1)$$

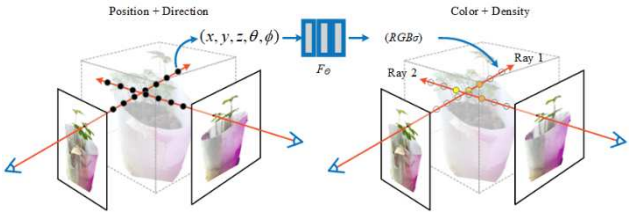

Figure 1 NeRF Overview

B. Extraction of Plant Phenotyping Parameters

In this part, we introduce the method of extracting plant phenotypic parameters in detail. First, we use Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [14] to remove irrelevant background from the image. Then, we use the K-means clustering algorithm to segment the leaves and stems of plants. Finally, by processing 3D point cloud data, we further extract the phenotypic parameters of plants.

DBSCAN algorithm is a density-based spatial clustering method suitable for segmenting complex-shaped 3D point clouds. It forms clusters by identifying high-density regions in the data and treats low-density regions as noise or background. DBSCAN does not require a preset number of clusters and performs effectively in application scenarios where background removal or identification of specific objects is needed. The algorithm's performance and final clustering results are affected by two key parameters: epsilon (eps) and minPts.

In the next stage, the leaves and stems of the plant are segmented. Because the planting of *P. notoginseng* is dense, the background of the photos we obtained is very complex. This scenario may make the Expectation-Maximization (EM) algorithm complicated and unstable when estimating the parameters of the Gaussian mixture model. The K-means algorithm has a loose assumption on data distribution and is more suitable for processing such complex image data. The k-means clustering algorithm is an efficient unsupervised learning technology. The purpose is to divide the data set X into k clusters so that the objects within each cluster are as similar as possible and the objects between different clusters are as different as possible. In this process, firstly, k objects are randomly selected as the initial cluster centers C_j , and then each sample is divided into the nearest cluster according to the distance between the sample and each cluster center. The distance between the object X_i and the cluster center C_j is usually measured by Euclidean distance. The calculation formula is as follows:

$$d(X_i, C_i) = \sqrt{\sum_{t=1}^{m} \left(x_{it} - c_{jt}\right)^2} \quad (2)$$

In formula (2), x_{it} is the attribute value of the object X_i in the t dimension, and C_{jt} is the coordinate value of the cluster center C_j in the same dimension. After all objects are assigned to the corresponding cluster, the algorithm updates the center of each cluster to become the mean of all objects in the cluster on each dimension. The update of cluster center C_j is shown in Formula 3. $|C_j|$ is the number of objects in the cluster C_j. By repeating the allocation and update steps until the convergence condition is satisfied.

$$C_{jt} = \frac{1}{|C_j|} \sum_{X_i \in C_j} x_{it} \quad (3)$$

For the parameter extraction of plant stem height, the 3D point cloud information of a single plant is first collected, and coordinate correction is performed to ensure that the growth direction of the plant is positively aligned with the *Z*-axis of the spatial right-angled coordinate system and that the projection of the base of the plant falls on the *XY* coordinate plane. In this coordinate framework, the absolute value of the difference between the *Z*-axis coordinates of the plant point cloud along the stalk direction and the *Z*-axis coordinates of the basal plane is the desired plant height.

Greedy Projection Triangulation (GPT) algorithm is an efficient point cloud data processing method, which is widely used in plant leaf area measurement. By connecting the three-dimensional points into triangles, GPT gradually constructs the surface and accurately reconstructs the blade geometry. The algorithm uses a greedy strategy to select the optimal triangle each time, usually by projecting new points to the existing surface to expand. By maximizing the triangle area or minimizing the distortion, the surface is smooth and continuous. GPT can effectively process noisy or sparse point cloud data and improve the accuracy of area extraction. The experimental results show that GPT is efficient and robust under large-scale data sets, and provides accurate leaf measurement, which provides important support for plant morphology research and agricultural applications.

IV. EXPERIMENTS

A. Implementation environment and details

The computer hardware used in the experiment is a 12th Gen Intel (R) Core (TM) i9-12900K processor, max/min CPU frequency: 5200.0000 MHz / 800.0000 MHz. The Operating system is Ubuntu 22.04.2 LTS, and the RAM is 128 G. The camera used is a CanonEOS600D with an 18-megapixel CMOS sensor.

B. Results and Discussion

We carried out experiments on the effect of 3D reconstruction based on NeRF. Since the three methods of Instan-NGP [15], NeRFacto and NeRFacto-big have similar impacts on the reconstruction of Panax notoginseng, we further compared their time-consuming reconstruction. Then, we use the shortest time-consuming method to explore the best interval angle for reconstruction. After the image reconstruction, we further verified the image segmentation and phenotypic parameter extraction.

The time-consuming nature of 3D reconstruction is an essential index for evaluating the performance of the model. According to the data in Table I, the time required to reconstruct single and multiple *P. notoginseng* under 72 data sets is shown. It can be seen from the table that whether it is a single plant or multiple plants, the time-consuming gap between the Instant-NGP and NeRFacto models is small, and the NeRFacto model is the fastest. However, the training time of the NeRFacto-big model is more than ten times that of the other two methods.

In the experiment, we chose the NeRFacto model with the shortest reconstruction time to explore the best angle interval for reconstructing *P. notoginseng* plants. We took two sets of images for each *P. notoginseng*: one group was taken manually, and the other group was an image sequence extracted from video frames. Due to the difficulty of accurately controlling the small angle unit, the manually captured image sequence only covers intervals of 5 °, 10 °, 15 °, 20 °, 30 °, 45 ° and 60 °. The comparative analysis of Fig. 2 and Fig. 3 shows that the manually acquired image sequences reconstructed the 3D point cloud very well at a 5° shooting interval. As the shooting interval increases, the number of point clouds gradually decreases, and problems such as point cloud voids begin to appear at 10°, and the problem is gradually aggravated. The image sequences the frame extraction program acquired performed best at 5° and 10° shooting intervals. Still, at 15°, there was a lot of noise and severe problems with detail performance, including point cloud voids and missing critical parts.

TABLE I COMPARISON OF THREE RECONSTRUCTION DURATIONS BASED ON THE NERF MODEL

Plant	NeRF model	Time/s
single	Instant-NGP	642
	NeRFacto	603
	NeRFacto-big	6600
multiple	Instant-NGP	570
	NeRFacto	523
	NeRFacto-big	6720

Interval 5° Interval 10° Interval 15° Interval 20°

Figure 2 3D reconstruction results of images captured manually at different interval angles

Interval 5° Interval 10° Interval 15° Interval 20°

Figure 3 3D Reconstruction results of images obtained by frame sampling at different interval angles

After image reconstruction, the stems and leaves of *P. notoginseng* were segmented. The results processed by the DBSCAN clustering algorithm are shown in Fig. 4: (a) is the initial untreated point cloud, and(b) is obtained after point cloud preprocessing. Then, the planter background is removed by down-sampling and DBSCAN clustering, and finally, the segmented plant point cloud is received, as shown in Fig. 4 (c). The background was separated through these steps, and the point cloud data of stems and leaves were retained and down-sampled. The segmentation results obtained after debugging the parameters are shown in Fig. 5. The appropriate K-means clustering algorithm can accurately segment the original point cloud of *P. notoginseng* plants into leaves and stems. Then, re-segmentation can separate each leaf.

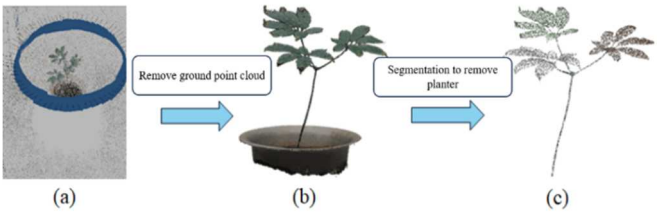

Figure 4 Removal of ground and background point clouds

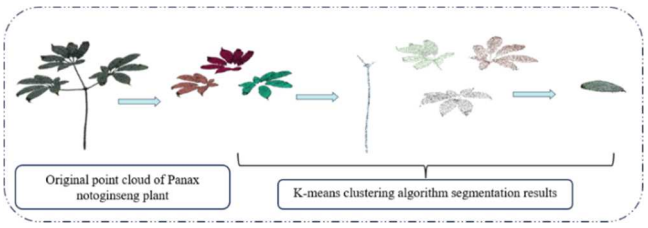

Figure 5 Schematic diagram of point cloud segmentation results of *Panax notoginseng* plants

After the segmentation of stems and leaves, we used the NeRFacto method based on NeRF to accurately measure the leaf area and plant height of *P. notoginseng*. Table II presents the extracted leaf area parameters from the point cloud data, indicating a determination coefficient (R^2) of 0.9826, a Root Mean Square Error (RMSE) of 0.57 cm^2, and an average Mean Absolute Percentage Error (MAPE) of 1.73%. These metrics underscore the method's high accuracy and reliability in extracting plant phenotypic parameters. Additionally, Table III displays the results of plant height measurements using the NeRFacto 3D reconstruction method for *P. notoginseng*, with an R^2 of 0.9998, an RMSE of 0.51 cm, and an average MAPE of 1.76%.

Finally, we compared the results of plant height and leaf area obtained by the NeRFacto model and SFM-MVS 3D reconstruction method, and the comparison results are shown in Table IV. The accuracy of the NeRFacto model in extracting plant height and leaf area parameters was significantly better than that of the SFM-MVS method. The calculated determination coefficient R^2, root mean square error RMSE, and mean absolute percentage error MAPE showed that the NeRFacto method was 1.01 % and 5.95 % higher than the SFM-MVS method.

V. CONCLUSION

We studied extracting phenotypic parameters from P. notoginseng plants, precisely measuring plant height and leaf area using the NeRFacto method, which was compared against manual measurements. The NeRFacto method demonstrated high accuracy and reliability: for leaf area, the R^2 was 0.9826, RMSE was 0.57 cm^2, and MAPE was 1.73%; for plant height, the R^2 was 0.9998, RMSE was 0.51 cm, and MAPE was 1.76%. In contrast, results from the SFM-MVS method showed slightly lower performance: for leaf area, R^2 was 0.84, RMSE was 2.48 cm^2, and MAPE was 7.68%; R^2 for plant height was 0.9992, RMSE was 0.56 cm, and MAPE was 2.77%. Our future research will focus on enhancing the anti-interference capabilities of 3D reconstruction methods to support more complex scenarios in crop 3D reconstruction and parameter extraction. This will provide crucial technical support for precision agriculture and crop breeding research.

TABLE II MEASUREMENT RESULTS OF LEAF AREA OF PANAX NOTOGINSENG PLANTS BASED ON THE NERFACTO METHOD

Plant number	*true values (cm)*	*NeRFacto values (cm)*	*Absolute percentage error (%)*	*Plant number*	*true values (cm)*	*NeRFacto values (cm)*	*Absolute percentage error (%)*
1	25.8	26.0	0.78	16	25.2	25.5	1.19
2	29.6	30.2	2.02	17	27.4	26.8	2.19
3	33.7	34.6	2.67	18	33.6	32.9	2.08
4	27.7	27.2	1.80	19	33.2	32.5	2.11
5	21.8	21.6	0.92	20	31.2	30.7	1.60
6	18.5	18.7	1.08	21	26.5	26.3	0.75
7	17.9	17.3	3.35	22	27.7	28.1	1.44
8	25.5	25.9	1.57	23	28.4	29.0	2.11
9	28.3	28.2	0.35	24	32.9	33.3	1.22
10	29.2	29.9	2.40	25	31.6	31.3	0.95
11	31.2	31.6	1.28	26	26.1	25.5	2.30
12	32.8	32.4	1.22	27	27.3	25.9	5.13
13	26.6	27.1	1.88	28	29.0	28.9	0.35
14	29.8	29.2	2.01	29	31.4	31.1	1.00
15	24.9	24.2	2.81	30	32.9	33.3	1.22

TABLE III RESULTS OF PANAX NOTOGINSENG PLANT HEIGHT MEASUREMENTS BASED ON THE NERFACTO METHOD

Plant number	*true values (cm)*	*NeRFacto values (cm)*	*Absolute percentage error (%)*
1	20.8	20.6	0.96
2	24.0	22.9	4.58
3	23.0	22.4	2.61
4	28.7	28.4	1.04
5	26.5	26.9	1.51
6	27.4	27.1	1.09
7	23.5	22.8	2.98
8	21.8	21.6	0.92
9	26.4	26.1	1.14
10	27.8	27.6	0.72

TABLE IV COMPARISON OF ALGORITHM PERFORMANCE

Methods	**Phenotype**	**R^2**	**RMSE (cm)**	**MAPE (%)**
SFM-MVS	***plant height***	0.99	0.56	2.77
	leaf area	0.84	2.48	7.68
NeRFacto	***plant height***	0.99	0.51	1.76
	leaf area	0.98	0.57	1.73%

ACKNOWLEDGMENT

This work was financially supported by the National Natural Science Foundation of China (No. 51979134), the Applied Basic Research Key Project of Yunnan (No. 202201AS070034), and the Yunnan Science and Technology Talent and Platform Program (No. 202305AM070006).

REFERENCES

[1] C. Zhao et al., "Crop Phenomics: Current Status and Perspectives," Front Plant Sci, vol. 10, p. 714, Jun. 2019, doi: 10.3389/fpls.2019.00714.

[2] W. Yang et al., "Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives," Molecular Plant, vol. 13, no. 2, pp. 187–214, Feb. 2020, doi: 10.1016/j.molp.2020.01.008.

[3] M. Legendre and G. S. Demirer, "Improving crop genetic transformation to feed the world," Trends in Biotechnology, vol. 41, no. 3, pp. 264–266, Mar. 2023, doi: 10.1016/j.tibtech.2022.12.002.

[4] D. Dodig et al., "Image-Derived Traits Related to Mid-Season Growth Performance of Maize Under Nitrogen and Water Stress," Front. Plant Sci., vol. 10, p. 814, Jun. 2019, doi: 10.3389/fpls.2019.00814.

[5] H. Cheng and H. Li, "Identification of apple leaf disease via novel attention mechanism based convolutional neural network," Front Plant Sci, vol. 14, p. 1274231, Oct. 2023, doi: 10.3389/fpls.2023.1274231.

[6] A. Fuentes, S. Yoon, S. Kim, and D. Park, "A Robust Deep-Learning-Based Detector for Real-Time Tomato Plant Diseases and Pests Recognition," Sensors, vol. 17, no. 9, p. 2022, Sep. 2017, doi: 10.3390/s17092022.

[7] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: representing scenes as neural radiance fields for view synthesis," Commun. ACM, vol. 65, no. 1, pp. 99–106, Jan. 2022, doi: 10.1145/3503250.

[8] K. Zhang, G. Riegler, N. Snavely, and V. Koltun, "NeRF++: Analyzing and Improving Neural Radiance Fields." arXiv, Oct. 21, 2020. Accessed: Jun. 23, 2024. [Online]. Available: http://arxiv.org/abs/2010.07492

[9] Z. Wang, S. Wu, W. Xie, M. Chen, and V. A. Prisacariu, "NeRF--: Neural Radiance Fields Without Known Camera Parameters." arXiv, Apr. 06, 2022. Accessed: Jun. 23, 2024. [Online]. Available: http://arxiv.org/abs/2102.07064

[10] M. Tancik et al., "Block-NeRF: Scalable Large Scene Neural View Synthesis," in 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA: IEEE, Jun. 2022, pp. 8238–8248. doi: 10.1109/CVPR52688.2022.00807.

[11] T. Müller, A. Evans, C. Schied, and A. Keller, "Instant neural graphics primitives with a multiresolution hash encoding," ACM Trans. Graph., vol. 41, no. 4, pp. 1–15, Jul. 2022, doi: 10.1145/3528223.3530127.

[12] X. Li, J. Park, C. Reberg-Horton, S. Mirsky, E. Lobaton, and L. Xiang, "Photorealistic Arm Robot Simulation for 3D Plant Reconstruction and Automatic Annotation using Unreal Engine 5".

[13] A. Pinkus, "Approximation theory of the MLP model in neural networks," Acta Numer, vol. 8, pp. 143–195, Jan. 1999, doi: 10.1017/S0962492900002919.

[14] M. Ester, H.-P. Kriegel, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise".

[15] Q. Wu, D. Bauer, M. J. Doyle, and K.-L. Ma, "Interactive Volume Visualization Via Multi-Resolution Hash Encoding Based Neural Representation," IEEE Trans. Visual. Comput. Graphics, pp. 1–14, 2024, doi: 10.1109/TVCG.2023.3293121.