

郭俊, 杨达成, 莫振杰, 等. 基于 SLAM 与神经辐射场的柑橘幼苗三维重建方法 [J]. 华南农业大学学报, 2025, 46(3): 429-438.
GUO Jun, YANG Dacheng, MO Zhenjie, et al. 3D reconstruction of citrus seedlings based on SLAM and NeRF[J]. Journal of South China Agricultural University, 2025, 46(3): 429-438.

基于 SLAM 与神经辐射场的柑橘幼苗 三维重建方法

郭俊¹, 杨达成¹, 莫振杰¹, 兰玉彬², 张亚莉^{1,2}

(1 华南农业大学 工程学院, 广东 广州 510642; 2 国家精准农业航空施药技术
国际联合研究中心, 广东 广州 510642)

摘要:【目的】针对现有三维重建技术难以获取柑橘幼苗精细三维点云及其三维表型参数来表征幼苗状态的问题, 提出一种基于同时定位与地图构建算法 (Simultaneous localization and mapping, SLAM) 与神经辐射场 (Neural radiance fields, NeRF) 的柑橘幼苗三维重建方法。【方法】以 1 年期的柑橘幼苗为研究对象, 首先, 利用深度传感器获取柑橘幼苗的 RGB 和深度图; 其次, 使用 SLAM 计算深度传感器在每一帧图像的位姿; 然后训练幼苗 NeRF 网络, 将附有位姿的多视角图像输入多层感知机 (Multilayer perceptron, MLP); 最后, 通过体积渲染监督训练, 重建高精度的三维实景点云模型。【结果】本文方法重建的柑橘幼苗三维模型, 在色彩与纹理方面极具真实性, 模型轮廓清晰、层次分明, 具有实景级的精度。该模型可有效提取柑橘幼苗的三维表型参数, 株高、宽幅长、宽幅宽和胸径的准确率分别为 97.94%、93.95%、94.11% 和 97.62%。【结论】研究有助于加快优良柑橘幼苗的选育进程, 可为柑橘产业的可持续性发展提供技术支持。

关键词: 柑橘幼苗; 植物三维表型; 三维重建; 神经辐射场; SLAM; 深度学习

中图分类号: S24; S666

文献标志码: A

文章编号: 1001-411X(2025)03-0429-10

3D reconstruction of citrus seedlings based on SLAM and NeRF

GUO Jun¹, YANG Dacheng¹, MO Zhenjie¹, LAN Yubin², ZHANG Yali^{1,2}

(1 College of Engineering, South China Agricultural University, Guangzhou 510642, China; 2 National Center for International Collaboration Research on Precision Agricultural Aviation Pesticide Spraying Technology, Guangzhou 510642, China)

Abstract: 【Objective】Aiming at the problem that it is difficult to obtain the accurate 3D point cloud of citrus seedlings and their 3D phenotypic parameters to characterize the state of seedlings with the existing 3D reconstruction techniques, this paper proposes a method based on the simultaneous localization and mapping (SLAM) and neural radiance fields (NeRF) for 3D reconstruction of citrus seedlings. 【Method】One-year old citrus seedlings were taken as the research object. Firstly, a depth sensor was used to capture the RGB map and depth map of the citrus seedling. Secondly, SLAM was employed to obtain the poses of the depth sensor in each frame of the image. Then, NeRF was trained for citrus seedlings, and the multi-view images with attached

收稿日期: 2024-05-20 网络首发时间: 2025-02-24 09:26:39

首发网址: <https://link.cnki.net/urlid/44.1110.S.20250221.1647.003>

作者简介: 郭俊, E-mail: junguo@stu.scau.edu.cn; 通信作者: 张亚莉, 主要从事农业航空传感器和农产品产地环境监测等研究, E-mail: ylzhang@scau.edu.cn

基金项目: 岭南现代农业实验室项目 (NT2021009); 高等学校学科创新引智计划 (D18019); 广东省重点研发计划 (2019B02022101); 广东省科技计划 (2018A050506073)

positional pose were fed into the multilayer perceptron (MLP). Finally, through supervised training with volume rendering, a high-precision 3D realistic point cloud model of citrus seedlings was reconstructed. 【Result】 The 3D model of citrus seedlings reconstructed by this method was highly realistic in terms of color and texture, with clear contours and distinct layers, and had real-world level accuracy. Based on this model, the 3D phenotypic parameters of citrus seedlings could be effectively extracted with the accuracy of 97.94% for plant height, 93.95% for breadth length, 94.11% for breadth width and 97.62% for stem thickness. 【Conclusion】 This study helps to accelerate the selection and nursery process of excellent citrus seedlings and provides a technical support for the sustainable development of the citrus industry.

Key words: Citrus seedling; Plant 3D phenotype; 3D reconstruction; Neural radiance fields (NeRF); Simultaneous localization and mapping (SLAM); Deep learning

广东是全国传统的柑橘优势产区之一, 柑橘种植历史悠久, 年产量逐年上升。优质的种苗是确保柑橘产业持续发展的基石^[1-2], 而准确获取柑橘幼苗三维表型参数成为实现高质量育苗的关键^[3-5]。传统的人工测量方式不仅效率低下, 而且主观性较强^[6], 难以准确反映柑橘幼苗的真实生长状态。基于三维点云获取植物三维表型参数^[7-9], 能够实现对柑橘幼苗的精准、快速、无损的三维表型检测^[10], 从而为后续育种和栽培提供科学依据, 推动柑橘产业的可持续性发展。

前人对苹果^[11]、玉米^[12]、大豆^[13]以及橡胶^[14]等多种农作物进行了视觉三维重建, 且多以运动恢复结构-多视图立体视觉 (Structure from motion-multiple view stereo, SFM-MVS) 方法为主。该方法受环境光线和背景复杂程度影响较大, 同时由于需要采集和处理庞大体量的图片, 重建速度缓慢; 其次精度不够, 局限于植株的整体效果, 不能上升至更为精细的器官等层面。

神经辐射场 (Neural radiance fields, NeRF) 是一种三维场景建模和合成的深度学习方法, 近年来在视觉三维重建领域备受关注。凭借操作简单、重建迅速、精度高等优势, NeRF 在一定程度上克服了 SFM-MVS 方法的不足。Mildenhall 等^[15]在 2021 年首次提出 NeRF, 该方法利用稀疏二维图像重建高质量的三维场景, 保证了三维点云的精度。宫金良等^[16]使用 NeRF 技术对果树进行重建, 获得了果树高精度的三维点云模型, 精度达到了 cm 级。Kang 等^[17]研究了 NeRF 合成二维图像和提取三维几何的性能, 并提供了一个全新的植物表型数据集。NeRF 适用于枝叶繁多、空间结构复杂的柑橘幼苗三维重建, 能够基于高精度的三维点云提取三维表型参数, 从器官尺度来表征柑橘幼苗的三维形态。

针对现有三维重建技术难以获取柑橘幼苗精

细三维点云及其三维表型参数来表征幼苗状态的问题, 本研究创新性地提出了一种基于同时定位与地图构建算法 (Simultaneous localization and mapping, SLAM) 与 NeRF 的柑橘幼苗三维重建方法。本研究旨在解决柑橘幼苗选育过程中的关键技术问题, 同时加快柑橘苗期生长状态检测进程, 为柑橘产业的可持续性发展提供可行的技术手段。

1 材料与方法

1.1 数据采集及处理

试验于 2023 年 9 月 22 日上午 10:00 至 11:00 在广东省肇庆市四会市沙糖橘幼苗培育基地 (23°36'N, 112°68'E) 进行, 试验对象为 10 株形态各异的 1 年期柑橘幼苗植株。首先, 使用图像采集设备环绕柑橘幼苗采集 360°视频并将视频每一帧都导出像素一致的 RGB 和深度图; 其次, 针对制作好的柑橘幼苗图像数据集, 使用 SLAM 技术对幼苗进行稀疏重建获得深度传感器在每一帧图像的位置; 最后, 通过图像位姿数据训练柑橘幼苗植株的神经辐射场, 获取高精度的点云模型。

为了提高数据的采集效率以及拍摄稳定性, 本文采用可自动导航的机器人底盘系统搭载 OAK-D-PRO 深度传感器采集柑橘幼苗的环绕视频数据。图像数据的采集在温室大棚中进行, 对间隔 30 cm、成行排列的盆栽柑橘幼苗进行环绕采集数据。该款深度传感器具备出色的深度和彩色图像捕捉能力。在深度性能方面, 支持 0.7~12.0 m 的测量范围, 覆盖近距离到中远距离的场景。在深度为 4.0 m 时测量误差 < 2%, 确保了高精度的测量效果。深度视场角 (Field of view, FOV) 为对角线 82°、水平 72°、垂直 50°, 能够提供较宽广的感知范围。深度图的分辨率为 1280×800, 清晰且细节丰富。在彩色图像方面, 彩色 FOV 为对角线 81°、水平 69°、垂直 55°, 能够

捕捉到宽广的彩色视野。彩色图分辨率高达 4032×3040 ,生成的彩色图像细节丰富,适用于对图像质量要求较高的应用场景。

本次试验采集的柑橘幼苗图像数据按照文献[18]的方法,相邻2张图像的画面重合率需 $\geq 80\%$,将采集的视频抽帧成1245张多视角图像。获取图像后,利用直尺(精度1 mm)和游标卡尺(精度0.02 mm)对柑橘幼苗的三维表型参数进行测量,统计每一棵幼苗的实际三维形态,测量参数主要包括株高、宽幅和胸径。

1.2 基于SLAM的幼苗稀疏重建与位姿恢复

在数据采集阶段,我们从多个视角获取柑橘幼苗的图像,利用SLAM计算相机在每一帧图像中的位姿。基于SLAM技术的稀疏重建与位姿恢复主要包含相机标定和位姿估计2个主要步骤。为确保计算机能准确识别和处理从RGB-D相机得到的图像信息,在使用前需进行标定。本研究采用广泛认可的张友正标定法^[19-20]进行相机标定。位姿估计涉及从世界坐标系到相机坐标系的转换,包括旋转和平移2个部分。具体的计算流程可以分为以下几个步骤:首先,进行特征点的提取和描述子的计算;接着进行初始的定位;然后,跟踪与重定位,闭环检测,以确认并修正漂移误差;最后,进行地图的更新与优化,以提高地图的精度和可靠性。

对输入的图像序列进行特征点提取,首先是FAST角点检测,其核心思想是如果1个像素与周围的像素差异比较大,则认为是1个角点;在图像中选择1个像素 P 令其像素值为 I_p ;以 p 为中心,以3个像素值为半径画圆,以对称的方式选择圆上16个像素;选择一个阈值 t ,如果16个像素中有连续 n (n 一般为12)个像素的值都大于 $I_p + t$ 或者小于 $I_p - t$,则可以认为 p 是角点。为了简化流程,本研究使用一种快速测试的方法:先测试点1和点9,如果这2个点与周围的像素差都小于 t ,则 p 不可能是角点,否则继续测试点5和点13。只有当1、5、9、13这4个点中至少有3个点与 p 的像素差大于阈值 t 才进行描述子的计算。ORB-SLAM3使用BRIEF描述子,以二进制向量的形式表达,大小为32个字节。先对图像进行高斯模糊,去除高频噪声对特征的干扰;然后选取一个以选择关键点为圆心,边长为31个像素的正反方内接圆的圆形范围,在该范围内随机挑选点对,如果点对里的第1个像素值大于第2个点的像素值,则描述子对应位的值为1,否则为0;最后计算描述子。

初始定位是通过分析2个连续帧之间的特征

点匹配来实现的。在这个过程中,采用随机一致性采样(RANDom sample consensus, RANSAC)算法来筛选并优化大量的特征点对,进而估计出最佳的基本矩阵,为获取相机的初始位姿提供关键依据。一旦获得初始位姿,可以结合三角化技术,通过最小化重投影误差来进一步细化和校准相机的精确位置和方向。

在跟踪与重定位过程中,依赖当前帧与上一帧之间的特征点匹配,并通过PnP解算精确求解当前帧的相机位姿。为了提高跟踪的鲁棒性,SLAM系统运用光流技术估计相机的运动,并整合IMU数据以补偿潜在的误差。一旦跟踪失败,ORB-SLAM3会迅速启动重新定位机制。这一机制首先计算当前帧的词袋模型向量(Bag of words, BOW),从而识别出大规模地图中潜在的重定位关键帧。接着,通过BOW匹配进一步筛选这些潜在的关键帧,最终精确求解并验证相机的位姿。

闭环检测专注于识别相机是否回到了先前访问过的地点,从而构建闭环约束。在执行闭环检测时,首先在全球地图中搜索与当前帧相似的关键帧。随后,利用BOW模型和词袋匹配算法,精确计算帧之间的相似性。一旦检测到闭环,ORB-SLAM3会立即执行回环关键帧的优化,并闭合地图,以消除累积误差,提高定位和建图的准确性。

地图更新与优化是SLAM系统的核心环节。基于当前帧的位姿信息和深度数据,新的三维点云被精确地添加到全局地图中,不断丰富和完善环境的几何表示。为提升SLAM系统的性能和稳定性,本文持续优化相机轨迹和地图点云。利用非线性优化方法,最小化重投影误差,精确调整相机位姿和地图点云的位置,确保地图的准确性和一致性。

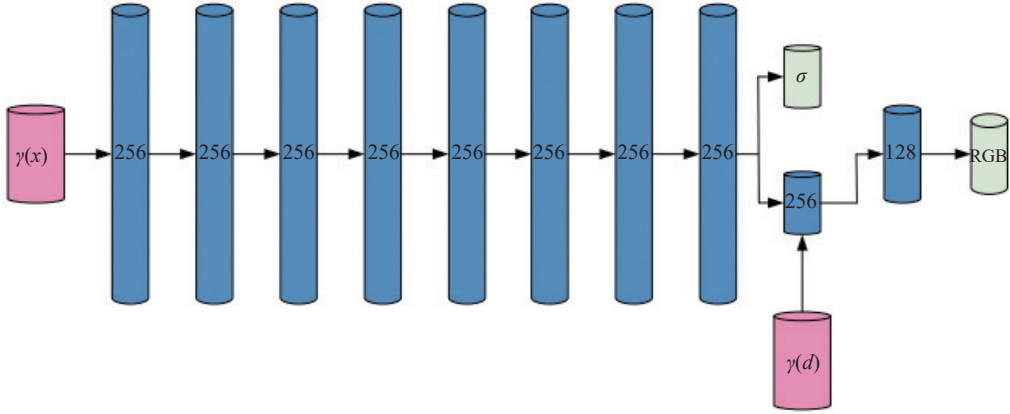
1.3 基于NeRF的幼苗稠密重建与点云获取

为解决传统立体视觉三维重建技术在准确表征柑橘幼苗多尺度复杂表型细节方面存在的问题,利用NeRF技术进行三维重建。相较于传统方法,NeRF能够更精细地捕捉柑橘幼苗表型的多尺度细节,提高表征的准确性和全面性。神经辐射场的建模过程需要借助真实世界中的图像和相应的相机位姿,通过神经网络技术模拟光的特性,以实现场景的高质量三维重建。在本研究中,试验环境的光照条件为10000~20000 lx,湿度为50%。

NeRF使用多层全连接网络(Multi-layer perceptron, MLP)表示场景中每个点的辐射特性,并通过优化隐式连续的体积函数来拟合场景。这个过程主要分为训练阶段和体渲染阶段。

在训练阶段,首先通过相机中心对图像的每个像素使用投影法,生成穿过三维实体的射线。通过对这些射线采样,获取一系列采样点的信息,其中包括采样点的位置数据 (x,y,z) 和射线的方向 (θ,φ) ,用三维单位向量 d 表示射线方向,作为神经网络的输入,并优化神经网络的权重,输出该采样点的颜色 RGB 和体积密度 σ ,使用符号 $F_\theta:(x,d)\rightarrow(c,\sigma)$ 表示网络训练过程。具体网络架构如图 1 所示。由于

神经网络更倾向于学习低频函数,为了更好地拟合高频数据,使用映射函数 $\gamma:R\rightarrow R^{2L}$, R 为维度空间,将输入的 5 维变量通过映射函数得到 $2L$ 维度的数据。随后将编码后的采样点位置信息输入到 8 层 MLP 网络里面,该网络输出该点的体积密度 σ 和 1 个 254 维的向量,将此向量与编码后的采样点方向信息结合,输入到另一层 MLP 网络,输出该点的 RGB 颜色信息。



γ 为映射函数, x 为点的位置信息, d 为视角方向, σ 为体素不透明度,256、128表示所在卷积层输入的图像尺寸。

γ is the mapping function, x is the position information of the point, and d is the perspective direction, σ is the voxel opacity, 256 and 128 indicate the size of image input in the located convolution layer.

图 1 NeRF 网络架构示意图

Fig. 1 Illustration of the NeRF network architecture

在体渲染阶段,对训练阶段相机每条射线上的各采样点的颜色和密度进行加权求和,并运用光线渲染公式(1)进行体素合成,计算得到每条相机射线 $r(t)=0+t_d$ 在 $t_n\sim t_f$ 范围内的颜色和密度,最终将合成的体素颜色和密度转换为图像像素值。

$$\begin{cases} C(r) = \int_{t_n}^{t_f} T(t) \sigma[r(t)] c[r(t), d] dt \\ T(t) = e^{-\int_{t_n}^t \sigma[r(s)] ds} \end{cases}, \quad (1)$$

式中, $C(r)$ 代表某条光线在平面上形成的像素值, $T(t)$ 代表沿相机射线方向在 $t_n\sim t$ 范围的累计透射率,即在射线向前传播过程中,不与其他粒子发生碰撞的概率。这个透射率用于权衡颜色和密度的累计贡献,从而影响最终的体素合成结果。 t_n 代表光线的起点, t_f 代表光线的终点, $\sigma[r(t)]c[r(t),d]$ 表示粒子密度和颜色的乘积, $\sigma[r(s)]$ 表示粒子的体积密度。

上述公式涉及连续积分,在实际应用中并不合理。为了更贴合实际,对其进行离散化处理。使用分层抽样方法,将相机射线在 $t_n\sim t_f$ 范围均匀分为 N 等份,从每一份中抽取 1 个样本。分层抽样的方法允许在不同位置对神经网络进行评估,从而确保对三维场景的描述是连续的。采用 Max 思想,用公式(2)来表示公式(1)。

$$\begin{cases} C'_c(r) = \sum_{i=1}^{N_c} T_i (1 - e^{-\sigma_i \delta_i}) C_i \\ T_i = e^{-\sum_{j=1}^{i-1} \sigma_j \delta_j} \end{cases}, \quad (2)$$

式中, $C'_c(r)$ 表示粗体积预测的 RGB 颜色信息, N_c 表示点的数量, T_i 表示第 i 等份的透射率, δ_i 表示相邻样本之间的距离。

然而,上述公式的计算效果受点的数量和位置影响。为了应对这一问题,我们引入粗细采样方法将公式(2)作为粗采样的基础,按公式(3)计算得到的 ω_i 则被用作细采样的概率密度函数。

$$\omega_i = T_i (1 - e^{-\sigma_i \delta_i}). \quad (3)$$

通过逆采样变换,我们得到了更精细化的采样点,进而通过细采样计算得到颜色。通过使用粗细采样方法获得的颜色与图像像素颜色构建残差,遍历整张图像像素,并将其作为损失函数(Loss)。

$$\text{Loss} = \sum_{r \in R} [\|C'_c(r) - C(r)\|_2^2 + \|C'_f(r) - C(r)\|_2^2], \quad (4)$$

式中, $C(r)$ 表示射线 r 的基本事实; $C'_f(r)$ 表示精体积预测的 RGB 颜色; R 为每批射线的集合,以对神经网络进行反向优化。

为了加速神经网络训练过程,采用网格密度与

哈希表来替代之前的位置编码, 如图 2 所示。首先, 将三维场景压缩至 0~1 的空间范围, 然后将空间划分为多尺度的网格。每个网格使用 8 个顶点作为可优化特征, 并通过哈希表高效地存储每个网格的位置, 以提高搜索效率。在采样阶段, 当采样点位于网格密度中时, 通过采样点的位置信息 (x, y, z) 在哈希表中找到对应的网格, L 表示维度空

间的数量; 使用该网格的 8 个特征进行线性插值, 得到采样点的特征向量; 将多尺度的特征向量连接在一起, 作为神经网络的输入; 神经网络处理输入、输出采样点的 RGB 和体积密度 σ 。为了更有效地处理密度信息, 使用另一个网络来存储每个网格存在物体的概率。当相机射线穿过密度网格时, 仅在含有物体的网格上进行取样, 优化采样点的采样策略。

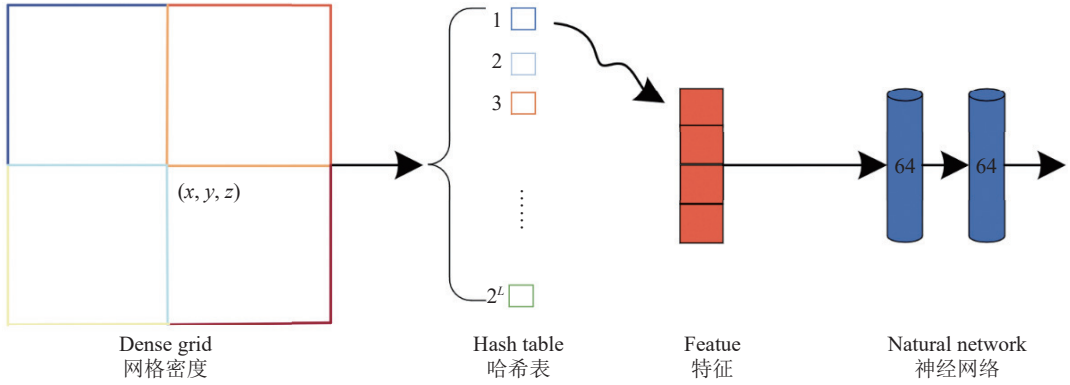


图 2 改进位置编码示意图
Fig. 2 Encoding illustration of improved position

2 结果与分析

2.1 位姿恢复效果分析

使用标定工具箱对深度相机进行内参标定, 确定相机的内部参数, 包括焦距以及畸变系数。这些参数可将图像坐标转换为实际的物理坐标, 并校正图像的畸变, 从而提高图像的精度和真实性。标定试验在 Ubuntu18.04 上进行, 按照以下步骤进行操作: 首先, 准备棋盘格作为标定板, 以提供足够的特征点; 随后, 手持相机分别沿 X、Y、Z 轴移动与绕坐标轴旋转拍摄包含标定板的图像, 使之收敛至 0.5 个像素误差。

通过标定获得深度相机的内参矩阵为:

$$\begin{bmatrix} 809.981 & 0 & 636.069 \\ 0 & 801.108 & 403.936 \\ 0 & 0 & 1 \end{bmatrix}。$$

畸变参数矩阵为:

$$[-4.298 \quad 37.726 \quad -0.000 \ 22 \quad -0.000 \ 23]。$$

使用 SLAM 的方法恢复每一帧图像的位姿。通过 ORB 进行特征点提取和特征匹配, 图 3 分别展示了图像筛选与前、后的特征点匹配结果, 共获得 5997 个特征点以及 205 个正确的特征点对。

以“1.1”中采集到的柑橘幼苗的图像数据为例, 所有图像经过 SLAM 稀疏重建后, 获得的图像位姿与每一帧的图像一一对应, 结果表明, SLAM 对每一帧图像的位姿都恢复成功。图像位姿如图 4 所示, 根据数据采集方案, 计算出的图像位姿与拍摄路径保持一致, 基本可以看出是绕幼苗 1 周的椭圆形范围, 并形成闭环。在试验中, 根据前期设置的深度传感器的参数, 果树的 RGB 图像分辨率为 1920×1080, 帧率为 30 帧/s, 深度图像素也与 RGB

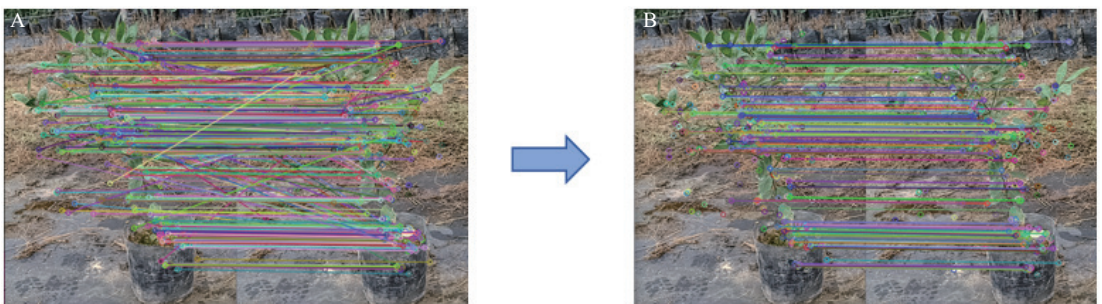
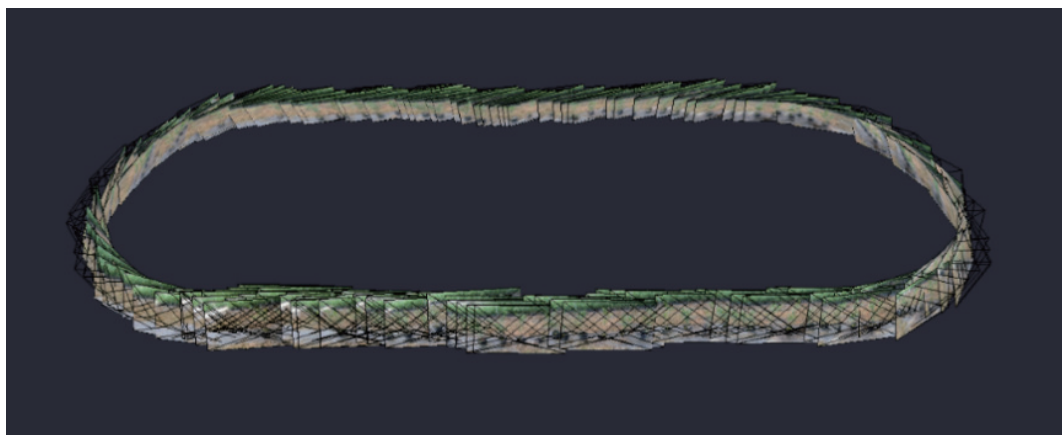


图 3 筛选前 (A)、后 (B) 特征点匹配示意图
Fig. 3 Matching illustration of unfiltered (A) and filtered (B) feature points

图像保持一致, 满足后期三维重建的要求。试验结果表明, 由 SLAM 方法所计算出来的每一帧图像的

位姿均具有较高的可靠度和准确性, 同时稳定性也较好, 图像具有高特征密度和良好的特征点分布。



图中的黑色四棱锥表示相机位姿, 即相机拍摄方向。

The black quadrangular pyramid in the figure indicates the camera position, that is, the camera shooting direction.

图 4 图像位姿示意图

Fig. 4 Illustration of image pose

2.2 柑橘幼苗的 NeRF 表征效果

将“2.1”通过 SLAM 计算得到的图像位姿以及柑橘幼苗的 RGB 图像制作成数据集, 导入 NeRF 网络进行训练。本次训练网络为 Nerfacto, 通过训练后获得柑橘幼苗植株的三维点云模型, 预处理后点云图如图 5 所示。同时, 柑橘幼苗原图、NeRF 场景图和 NeRF 深度图的对比如图 6 所示。为进一步验证本文提出方法的优越性, 应用 SFM-MVS 来重建柑橘幼苗的点云, 结果如图 7 所示。从预处理的柑橘幼苗三维点云图 (图 5) 可以看出, 用 NeRF 重建出来的三维模型, 在色彩与纹理方面极

具真实性, 模型轮廓清晰、层次分明; 从对比图 (图 6) 也可以看出重建的柑橘幼苗具有实景级的精度, 在 NeRF 深度图中也是枝叶分明, 本研究重建的柑橘幼苗模型完整清晰无任何缺失, 高精细的质量可以反映出 NeRF 在三维重建领域的优越性以及柑橘幼苗三维尺度的卓越表征能力。从图 7 可以看出 SFM-MVS 所生成的柑橘幼苗点云模型重建的柑橘幼苗点云稀疏、缺失, 总体三维轮廓不清晰, 在色彩与纹理方面也与实景相差较远。结果说明本文方法所重建的柑橘幼苗三维点云模型更具优越性, 更能反映柑橘幼苗真实的三维形态。

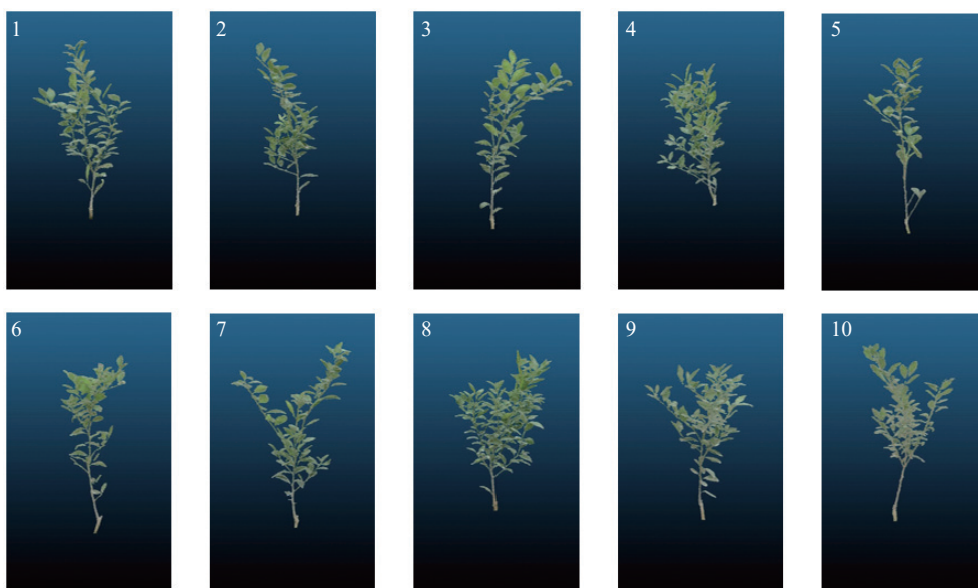


图 5 NeRF 重建柑橘幼苗三维点云图

Fig. 5 Citrus seedling 3D point cloud maps reconstructed by NeRF

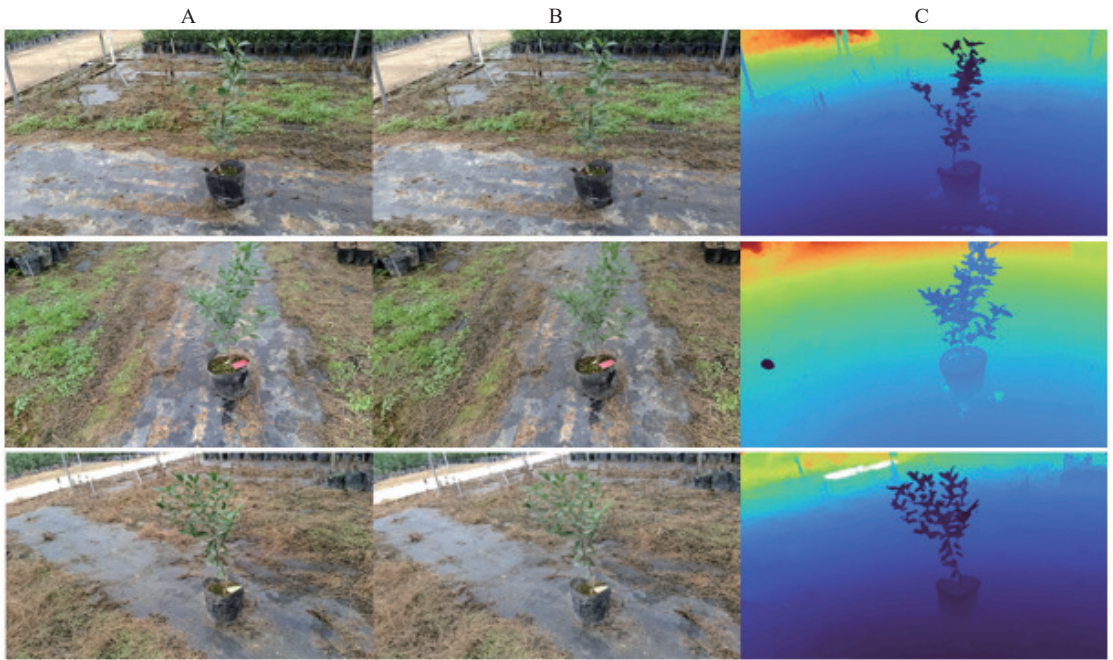


图 6 柑橘幼苗原图 (A)、NeRF 场景 (B) 和 NeRF 深度 (C) 对比

Fig. 6 Comparison of citrus seedling (A), NeRF scene (B) and NeRF depth (C) images

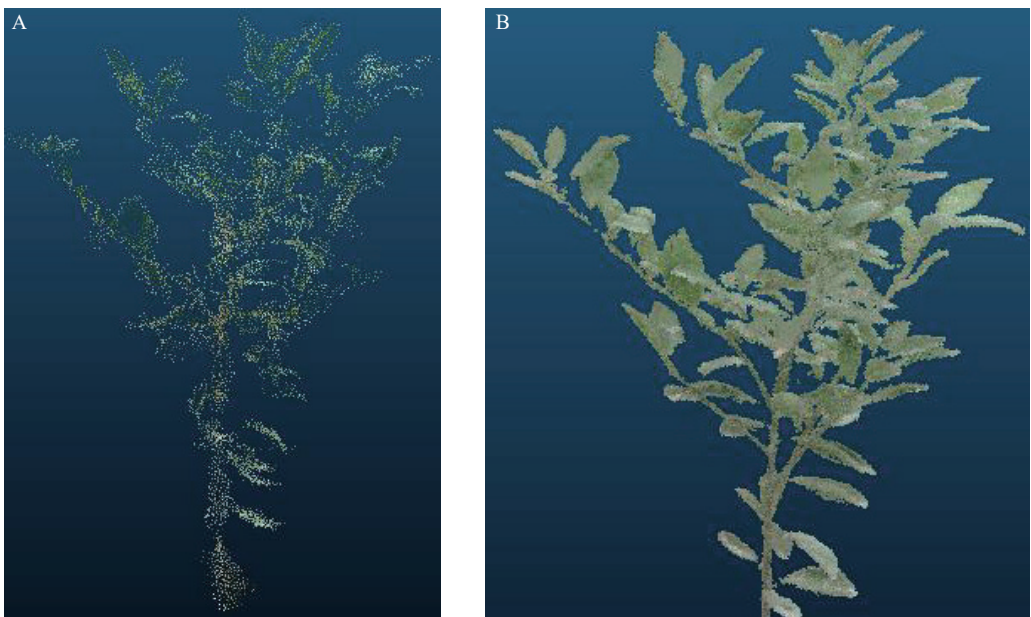


图 7 SFM-MVS(A) 与 NeRF(B) 重建柑橘幼苗三维点云图对比

Fig. 7 Comparison of citrus seedling 3D point cloud maps reconstructed by SFM-MVS (A) and NeRF (B)

2.3 幼苗株型参数的获取

为了精确、快速、无损地对柑橘幼苗进行高通量检测获取其三维表型信息, 从而选育出具有优良性状的柑橘幼苗, 本文通过 NeRF 渲染得到的柑橘幼苗三维点云模型来提取柑橘幼苗的三维表型参数, 包括株高、宽幅和胸径; 并对 10 株形态各异的柑橘幼苗的三维表型参数提取精度进行定量分析与评估; 采用线性回归分析评价柑橘幼苗的株型参数的人工实测值与提取值的相关性, 以决定系数

(Determination coefficient, R^2)、均方根误差 (Root mean square percentage error, RMSE) 以及平均绝对百分比误差 (Mean absolute percentage error, MAPE) 作为评价指标。相关指标计算公式如下:

$$R^2(X, Y) = \frac{\text{Cov}^2(X, Y)}{\text{Var}[X] \text{Var}[Y]}, \quad (5)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ai} - x_{mi})^2}, \quad (6)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|x_{ai} - x_{mi}|}{x_{mi}} \times 100\%, \quad (7)$$

式中, $\text{Var}[X]$ 表示 X 的方差, $\text{Var}[Y]$ 表示 Y 的方差, $\text{Cov}^2(X, Y)$ 表示 X 与 Y 的协方差, n 表示样本的数量, x_{ai} 表示提取值, x_{mi} 表示人工实测值。

将 NeRF 网络训练得到的柑橘幼苗点云进行点云预处理, 得到较为精细的点云模型, 基于该模型提取柑橘幼苗的三维表型参数。点云预处理主要包括目标区域分割、点云降噪、地面点云去除、单株幼苗分割、高程归一化以及幼苗与盆栽分割 6 个步骤。

首先, 由于深度相机的测量深度最高达 12 m, 所以重建出来的原始点云还包括了大棚、杂草和其他幼苗等非目标区域点云 (图 8A)。为了分割出目标点云区域, 本研究使用 CloudCompare 软件中的点云剪切功能对原始点云进行裁剪, 分离出只包含 10 株柑橘幼苗点云的区域; 其次, 受到相机精度、

外界环境等因素的影响, 原始点云数据不可避免地产生了一些噪声点和局外点, 对后续表型提取产生不良影响, 因此采用高适应性、高稳定性、高效性的统计滤波算法来清除点云中的噪声 (图 8B)。接着, 在进行点云分割前对点云数据进行地面滤波, 消除地面点云对分割精度的不良影响, 本文中使用了布料模拟滤波算法 (Cloth simulation filter, CSF) 实现地面曲面的拟合, 去除地面点云 (图 8C)。然后, 采用条件欧式聚类算法分割出单株的盆栽幼苗植株 (图 8D), 以便提取三维表型参数。为了避免盆栽对幼苗三维表型参数提取的影响, 以盆栽的真实高度为基准, 使用直通滤波将完整的柑橘幼苗点云分割出来。最后, 为了保证提取株高的精度, 本文使用高程归一化算法, 将所有点云的 Z 轴方向的值统一到共同的起点值, 有效解决由于地势变化引起的柑橘幼苗起始高度不一致的问题。

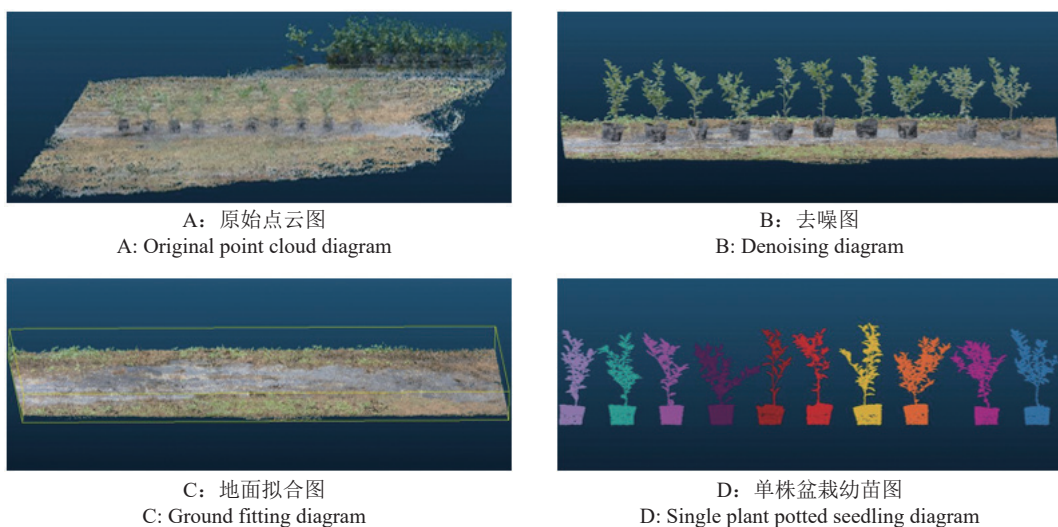


图 8 点云预处理示意图

Fig. 8 Illustration of point cloud preprocessing

在本研究中, 将柑橘幼苗点云图在 Z 轴方向的最大值作为单株柑橘幼苗的株高; X 和 Y 方向的最大值分别作为单株柑橘幼苗的宽幅长和宽。将预处理得到的柑橘幼苗三维点云图导入 Cloudcompare 软件, 遍历模型中的所有点, 找到点云最低点和最高点的 Z 坐标, 其差值的绝对值即为该株幼苗的株高; 找到点云最低点和最高点的 X 、 Y 坐标, 其差值的绝对值即为该株幼苗的宽幅。

参照文献 [21] 的定义方法, 选取幼苗主茎固定的某一高度进行测量, 将水平方向坐标的最大值与最小值差值的绝对值定义为柑橘幼苗的胸径。在本研究中, 选取柑橘幼苗土壤平面往上 5 cm 的位置, 同株高的获取办法一致, 利用预处理后的柑橘幼苗

点云, 在 Cloudcompare 软件中遍历该定义位置的点云坐标, 找到水平面内坐标的最大值和最小值差值的绝对值 (胸径)。

本研究基于 SLAM 与 NeRF 的方法对柑橘幼苗植株进行三维重建, 同时提取幼苗的三维表型参数, 包括株高、宽幅和胸径, 并将这些参数与人工实际测量值进行对比 (图 9), 以评价本研究重建方法的精度, 从而为柑橘幼苗的生长状态检测提供保障。重建的柑橘幼苗三维模型可有效提取柑橘幼苗的三维表型参数, 株高、宽幅长、宽幅宽和胸径的准确率分别为 97.94%、93.95%、94.11% 和 97.62%。

综上, 本文方法对于柑橘幼苗三维表型参数的

提取值与人工的实测值具有较好的一致性,由此反映出本文方法重建的柑橘幼苗的三维点云模型具

有较高的精度,可以用来表征柑橘幼苗的三维形态特征。

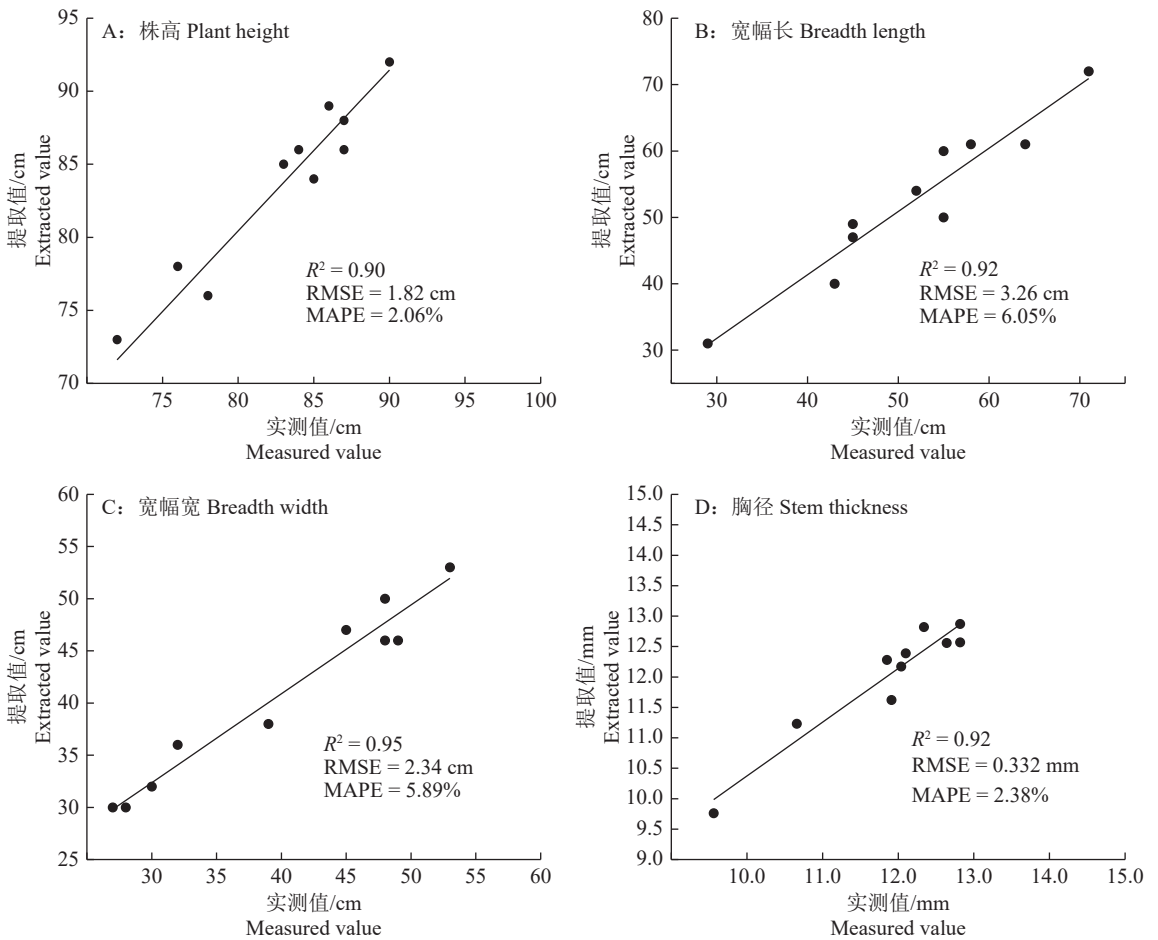


图9 柑橘幼苗三维表型参数人工实测值与提取值对比

Fig. 9 Comparison of manually measured and extracted values for 3D phenotypic parameters of citrus seedlings

3 结论

本研究针对目前传统三维重建技术难以精细表征植株的三维尺度信息、精准获取柑橘幼苗立体点云模型的问题,以10株形态各异的柑橘幼苗植株为研究对象,提出了一种基于SLAM与NeRF的柑橘幼苗三维重建方法并提取了其三维表型参数,得出以下结论:

1) 基于SLAM与NeRF的三维重建方法克服了传统的三维重建技术设备昂贵、计算难度大、操作烦琐以及精度低等一系列的弊端。通过该方法重建出来的柑橘幼苗三维点云模型,无论在色彩还是纹理方面都比传统方法更具真实性。所重建模型的高精度可以反映出NeRF在三维重建领域的优越性以及三维结构的卓越表征能力。

2) 基于本研究重建出的柑橘幼苗三维点云模型,对株高、宽幅长、宽幅宽和胸径三维表型参数进行提取,提取的准确率分别为97.94%、93.95%、

94.11%和97.62%。表明本文所提出的方法具有较高的优越性,点云提取值与人工测量值具有较好的一致性,对获得柑橘等植株幼苗的高精细的点云模型具有重要意义,有助于加快优良柑橘幼苗的选育进程,为农业新质生产力的发展提供技术支撑。

本研究中基于SLAM恢复相机位姿,但实际结果不尽如人意。计算的位姿不够准确,这对三维重建的精度造成了一定的影响,主要有以下几个方面的原因:深度传感器获取的数据存在一定的噪声和误差,并且误差随着时间累积导致位姿估计的不准确;本研究仅使用单一深度传感器,SLAM的运行中没有来自不同传感器的数据关联,特征点的检测可能会出错,从而影响位姿估计;此外,还有现实环境的影响,SLAM技术通常假设环境是静止且没有发生变化的,现实中的变化导致地图的更新和位姿估计变得困难。在后续的研究中,可以着重解决以上问题,设计位姿优化策略,进一步提升位姿估计

的精度,增强该算法的鲁棒性和可靠性,从而实现更高质量的植株三维重建,尽可能降低植株三维表型参数提取的误差,更精确地分析植株的形态特征、生长状态和生理变化,为植物表型研究提供更准确的数据支持。同时,三维点云重建在实际应用中仍面临一些局限性。点云数据往往存在噪声,尤其是在复杂场景中,对三维重建算法提出了更高要求,不仅需要算法稳定性方面的提升,更要考虑对数据集的优化,提升数据的采集精度,有效减少光照变化、环境复杂度及遮挡等因素对重建效果的影响,实现更高水平的三维重建,助力农业科技的进一步发展,使三维点云重建在精准农业领域获得更广阔的应用。

参考文献:

- [1] 柴宏红, 邵科, 于超, 等. 基于三维点云的甜菜根表型参数提取与根型判别[J]. *农业工程学报*, 2020, 36(10): 181-188.
- [2] PAULUS S, SCHUMANN H, KUHLMANN H, et al. High-precision laser scanning system for capturing 3D plant architecture and analysing growth of cereal plants[J]. *Biosystems Engineering*, 2014, 121: 1-11.
- [3] LADANIYA M S, MARATHE R A, DAS A K, et al. High density planting studies in acid lime (*Citrus aurantifolia* Swingle)[J]. *Scientia Horticulturae*, 2020, 261: 108935.
- [4] CONTI G, XOCONOSTLE-CÁZARES B, MARCELINO-PÉREZ G, et al. *Citrus* genetic transformation: An overview of the current strategies and insights on the new emerging technologies[J]. *Frontiers in Plant Science*, 2021, 12: 768197.
- [5] BISI R B, ALBRECHT U, BOWMAN K D. Seed and seedling nursery characteristics for 10 USDA *Citrus* rootstocks[J]. *HortScience*, 2020, 55(4): 528-532.
- [6] SHAKOOR N, LEE S, MOCKLER T C. High throughput phenotyping to accelerate crop breeding and monitoring of diseases in the field[J]. *Current Opinion in Plant Biology*, 2017, 38: 184-192.
- [7] 黄成龙, 李曜辰, 骆树康, 等. 基于结构光三维点云的棉花幼苗叶片性状解析方法[J]. *农业机械学报*, 2019, 50(8): 243-248.
- [8] 何海清, 严椰丽, 凌梦云, 等. 结合三维密集点云的无人机影像大豆覆盖度提取[J]. *农业工程学报*, 2022, 38(2): 201-209.
- [9] MA H, ZHU X, JI J, et al. Rapid estimation of apple phenotypic parameters based on 3D reconstruction[J]. *International Journal of Agricultural and Biological Engineering*, 2021, 14(5): 180-188.
- [10] CHAWADE A, VAN HAM J, BLOQUIST H, et al. High-throughput field-phenotyping tools for plant breeding and precision agriculture[J]. *Agronomy*, 2019, 9(5): 258.
- [11] 杨会君, 王瑞萍, 王增莹, 等. 基于多视角图像的作物果实三维表型重建[J]. *南京师范大学学报(自然科学版)*, 2021, 44(2): 92-103.
- [12] 梁秀英, 周风燃, 陈欢, 等. 基于运动恢复结构的玉米植株三维重建与性状提取[J]. *农业机械学报*, 2020, 51(6): 209-219.
- [13] HE W, YE Z, LI M, et al. Extraction of soybean plant trait parameters based on SfM-MVS algorithm combined with GRNN[J]. *Frontiers in Plant Science*, 2023, 14: 1181322.
- [14] LEI S, LIU L, XIE Y, et al. 3D visualization technology for rubber tree forests based on a terrestrial photogrammetry system[J]. *Frontiers in Forests and Global Change*, 2023, 6: 1206450.
- [15] MILDENHALL B, SRINIVASAN P P, TANCİK M, et al. NeRF: Representing scenes as neural radiance fields for view synthesis[J]. *Communications of the ACM*, 2021, 65(1): 99-106.
- [16] 宫金良, 刘镇霄, 魏鹏, 等. 基于相机位姿恢复与神经辐射场理论的果树三维重建方法[J]. *农业工程学报*, 2023, 39(22): 157-165.
- [17] KANG H, HU K, YING W, et al. High-fidelity 3d reconstruction of plants using neural radiance field[J]. *Computer and Electronics in Agriculture*, 2024, 220: 108848.
- [18] IGLHAUT J, CABO C, PULITI S, et al. Structure from motion photogrammetry in forestry: A review[J]. *Current Forestry Reports*, 2019, 5(3): 155-168.
- [19] ZHANG Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 22(11): 1330-1334.
- [20] ZHAO C, FAN C, ZHAO Z. A binocular camera calibration method based on circle detection[J]. *Heliyon*, 2024, 10(19): e38347. doi: 10.1016/j.heliyon.2024.e38347.
- [21] 彭程, 苗艳龙, 汪刘洋, 等. 基于三维点云的田间香蕉吸芽形态参数获取[J]. *农业工程学报*, 2022, 38(S1): 193-200.

【责任编辑 霍 欢】