

高月芳,肖冬冬,傅汝佳,等.面向作物表型分析的大豆植株叶片语义重建[J].华中农业大学学报,2023,42(3):177-186.
DOI:10.13300/j.cnki.hnlkxb.2023.03.021

面向作物表型分析的大豆植株叶片语义重建

高月芳¹,肖冬冬¹,傅汝佳²,冼楚华²,李桂清²,黄琼^{1,4},杨存义³

1. 华南农业大学数学与信息学院,广州 510642; 2. 华南理工大学计算机科学与工程学院,广州 510006;
3. 华南农业大学农学院,广州 510642; 4. 广州市智慧农业重点实验室,广州 510642

摘要 为解决三维扫描仪、多视图数据获取的三维点云因缺少语义信息导致难以从点云上判别植株器官部位问题,提出一种二维先验语义嵌入的大豆植株叶片三维语义建模方法:首先,基于Mask R-CNN模型对大豆叶片进行语义分割;然后,对分割结果和多视图数据进行立体重建融合学习,实现大豆植株叶片二维语义到三维叶片点云迁移,获得植株叶片点云语义信息,进而建立植株叶片三维语义模型。通过多组盆栽大豆植株试验对该模型进行验证,提取叶长和叶宽与人工实测数据进行对比分析,叶长和叶宽均方误差分别为2.53和1.52 mm,决定系数分别为0.97和0.89。结果表明,该方法能够便捷、精准地构建植株叶片三维语义模型。

关键词 语义分割;植株三维建模;深度网络模型;点云语义重建;作物表型分析

中图分类号 S126 **文献标识码** A **文章编号** 1000-2421(2023)03-0177-10

作物品种的生产潜力受多方面因素影响,其中作物生长发育过程中植株的外观形态和组织结构起重要作用^[1-3],因此,全面测量和定量分析其形态和结构对理想株型培育、产量和品质提升均具有重要意义。相较于耗时费力、效率低、误差大、且不适用于大批量处理的人工获取方法,近年来基于三维重建模型的高通量自动化作物形态特征获取研究发展迅速^[4-6]。Zheng等^[7]采用基于模板的动态跟踪算法从三维点云序列中重建开花过程中的花朵几何形变,并设计自适应约束优化策略对开花早期不可见部分建模。傅汝佳等^[8]提出一种基于骨架的大豆植株点云三维重建方法,该方法基于RGB-D深度相机采集大豆植株深度图像,使用基于自旋转平台的点云刚体粗配准方法提升点云数据配准,并在此基础上设计点云分层聚类方法对不同形态特征的大豆植株进行快速三维重建。类似地,Liu等^[9]提出的基于双目深度相机的三维建模方法通过感兴趣区域选择算法对点云噪声进行预处理,并根据点云特征的位置信息进行旋转和变换,进而构建花生植株完整三维模型。此外,Ando等^[10]提出一种对三维点云中的噪声和缺失点具有稳健性的叶片表面重建模型,该模型对叶片的形状和畸变分别建模表达,在减少点

云和缺失点影响的同时可维持叶片表面重建的准确性。

与上述基于三维扫描仪和深度相机模式建立植株三维模型不同,一些研究利用二维图像或视频进行作物形态结构三维建模^[11-13]。冯佳睿等^[14]基于大豆冠层图像同步采集平台构建三维大豆冠层结构形态模型,并借助其提供的深度信息计算个体和群体株高。Doi等^[15]则借助对极约束建立多视图分割实例之间的对应关系以实现大豆植株叶片三维建模。该方法将各视图中每个实例作为图中1个节点,将实例对应问题转化为图聚类问题实现多视图之间叶片对应关联,在此基础上结合MVS模式重建植株叶片实例。不过该方法仅适用于叶片在多视图中均可见的情况,当叶片存在一定遮挡则影响重建精度。此外,Wu等^[16-17]基于多视图数据构建玉米幼苗三维模型,并开发一个玉米幼苗表型分析平台以实现玉米采集、处理、表型性状自动提取及数据管理。Das-Choudhury等^[18]基于多视角图像构建玉米植株网格模型,用于株高、冠层地面投影面积、叶长等形态结构特征解析。

尽管目前三维植株建模研究取得了较好进展,但无论三维扫描仪或多视图方式获取的三维数据大

收稿日期:2022-11-19

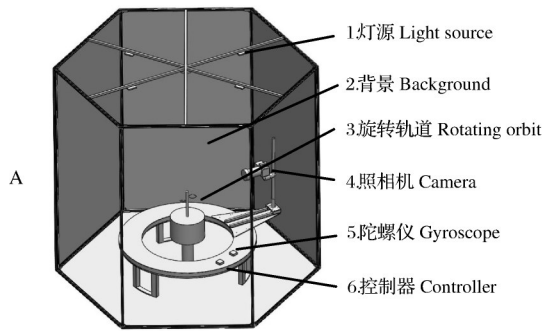
基金项目:广东省重点研发项目(2020B020220008);广州市科技计划项目(201902010081)

高月芳,E-mail:gaoyuefang@scau.edu.cn

通信作者:杨存义,E-mail:yey@scau.edu.cn

都只包含场景的几何属性,如形状、法向、曲率等,缺少高层语义类别属性。此外,三维点云的非结构化特性使得点云具有较强的主观性和不确定性,导致植株不同器官(如叶柄、叶片、茎干等)颜色相近或遮挡时,难以在点云上实现不同器官的准确分割,进而影响后续各器官形态结构解析。另一方面,植株二维图像蕴含丰富的语义信息和结构信息(如叶片和茎干结构、不同视角对应等),这些可为植株精准三维建模提供监督信息。

针对上述问题,本研究提出一种植株三维叶片语义建模方法,通过将植株二维叶片语义信息作为一种先验知识嵌入到现有植株多视图三维建模学习过程,以提升模型的语义表达能力。首先,基于Mask R-CNN模型对大豆植株叶片进行语义分割,获得不同视角大豆植株叶片语义信息,然后结合叶片语义信息和多视角重建模型进行二维到三维的语义迁移学习,实现植株叶片三维语义建模,旨在便捷、精准构建植株叶片三维语义模型,为后续植株生长发育、理想株型设计及株型解析等提供支持。

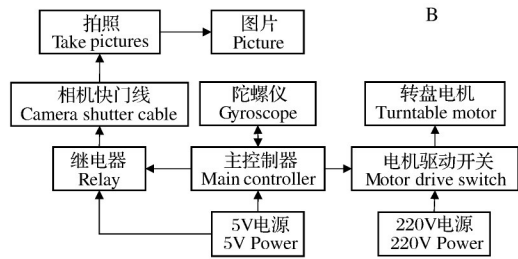


1 材料与方法

1.1 试验材料

选取盆栽大豆植株为试验材料。由于大豆植株的复杂性和特殊性,如叶片密集、非刚体且柔软、茎较细等,为获取高精度的植株图像数据,本研究搭建一个全自动密闭采集环境(图1),以保证采集数据不受外界干扰。该采集环境主要包含:(1)灯源。控制内部 24 h 为同一光源。(2)背景。选取 6 张 2 040 mm×1 200 mm 黑色遮光布,使拍摄环境与外界隔开。(3)旋转轨道。可以带动物体任意旋转。(4)照相机。拍摄时设定恒定的相机参数。(5)陀螺仪。自动控制相机到指定角度拍摄。(6)控制器。控制相机自动拍摄。

此外,进行植株叶片语义信息采集的植株叶片应满足以下要求:(1)拍摄角度按照预定义的角度旋转拍摄;(2)使用统一的相机设定曝光时间、焦距等;(3)植株尽可能摆放在扫描装置的中心位置;(4)尽可能保证植株每张叶片至少在某视图中无遮挡出现。



A:采集装置结构图; B:内部控制结构图。A:Structure of acquisition device; B:Control structure diagram.

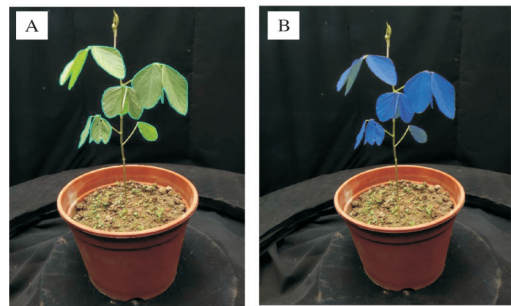
图 1 数据采集平台及结构图

Fig.1 Data acquisition platform and structure diagram

基于上述设备得到大豆植株图像,使用标注工具Labelme对叶片进行标注(图2A),用于叶片语义分割模型学习。由于叶片存在相互遮挡,标注原则是人眼能否辨别叶片基本轮廓,若叶片被严重遮挡且不具备基本形状时放弃标注。叶片标注完毕,生成对应的json文件,该文件通过坐标形式保存叶片的位置信息,解析 json 文件,则可获得叶片的掩码图(图2B)。

1.2 二维语义引导大豆植株叶片三维语义建模

为充分利用多视角图像蕴含的语义信息驱动现有基于图像的三维重建模型学习,以得到具有叶片



A:叶片标注图; B:叶片掩码图。A:The leaf annotation; B:The leaf mask.

图 2 大豆植株叶片标注

Fig.2 Samples of soybean plant with leaf label

语义信息的植株点云,首先,对采集得到的植株图像进行分割,获得叶片分割实例;然后,借助叶片分割结果结合多视角重建进行二维到三维的语义迁移学习,进而实现大豆植株叶片三维语义建模。建模流程如图 3 所示,首先,将采集的部分大豆植株数据作

为训练集输入 Mask R-CNN 分割模型进行训练;然后,将其余大豆植株图像输入训练好的模型,获得叶片语义分割结果。在此基础上对分割结果和多视角重建融合学习,通过二维到三维的叶片语义迁移实现植株叶片三维语义建模。

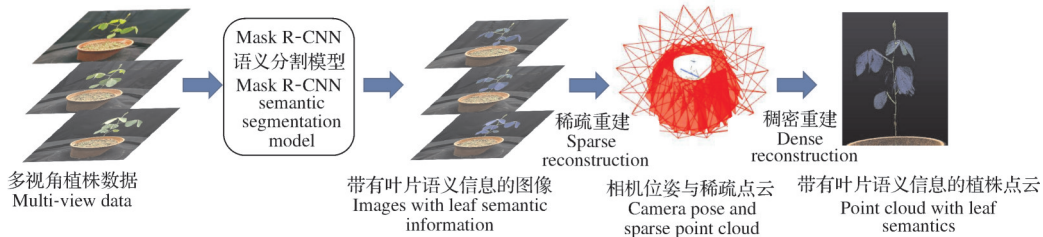


图 3 大豆植株叶片语义建模流程

Fig.3 Pipeline of semantic model of soybean plant leaf

1)大豆植株叶片语义分割。基于 Mask R-CNN 模型对大豆植株叶片进行语义分割,获得其语义信息。分割模型如图 4 所示,主要由卷积模块(convolutional neural network, CNN)、区域候选模块(region proposal network, RPN)、感兴趣区域对齐 RoIAlign、类别分类分支和边界框回归分支(class and Bbox)及语义分割分支组成。对输入的大豆植株图像,首先采用 CNN 模块提取其特征图,该特征图通过共享用于后续的 RPN 模块和全卷积层;然后将该特征图送入 RPN 生成叶片候选框,同时使用 RoIAlign 策略从上述叶片不同尺寸候选框中得到固定尺寸的特征图。进一步地,类别分类分支和边界框回归分支使用这些特征进行类别预测及检测框定位,获得每个叶片实例及其检测框精准位置。在此基础上,由全卷积网络(fully convolutional network, FCN)组成的语义分割分支通过像素级别的预测从上述检测框中得到检测区域的二值掩膜,进而得到大豆植株每个叶片的语义分割结果。

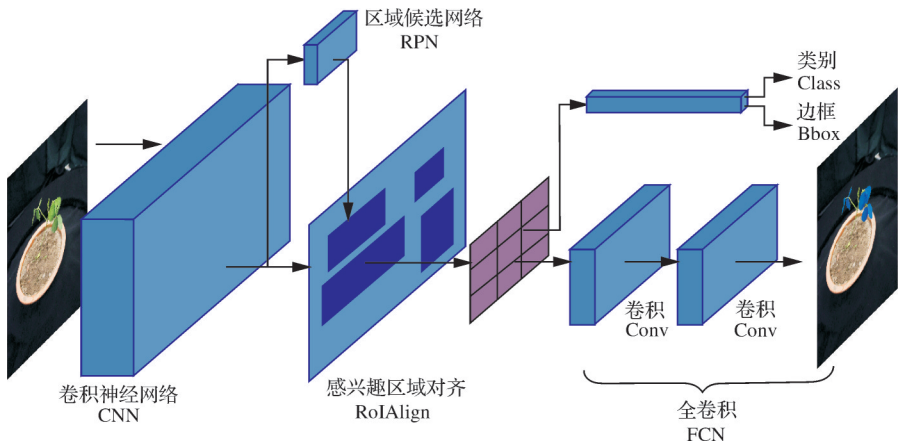


图 4 大豆植株叶片语义分割

Fig.4 Semantic segmentation of soybean plant leaf

2)大豆植株叶片语义建模。为获得包含叶片语义信息的大豆植株三维点云,将经由 Mask R-CNN 模型学习得到的二维叶片语义知识嵌入到基于图像的三维重建模型,自动利用这些高层语义信息驱动模型训练,实现二维到三维的叶片语义迁移学习,进而得到包含叶片语义的大豆植株三维模型。具体流程如图 3 所示,输入带有叶片语义信息的多视图大豆植株图片,首先通过 SfM 重建获得所有图片对应的

相机内外参数和带有叶片语义信息的稀疏点云,然后使用深度图融合的三维重建方法得到带有叶片语义信息的稠密点云。

①基于 SfM 的大豆植株叶片稀疏语义重建。为获得包含叶片语义信息的大豆植株稀疏三维点云,采用 SfM 方法对输入的一系列无序且带有叶片语义信息的大豆植株图像进行重建,其流程见图 5。首先,输入 1 组无序大豆植株图像,检测每帧图像中的

特征点,并对所有图像两两匹配。获取所有图像之间的匹配关系后,对匹配较好的图像进行连接,进而构建图像连接图。然后,选取2张图像构成初始图像对并进行重建,该图像对为三维重建中的1对匹配点多、基线长且满足homography匹配尽可能少的图像。根据初始2帧图像获取Tracks并进行滤波。由于Tracks存在一些外点且其精确性影响点云重建精

度,因此,需要三角量测计算初始三维点,并通过捆绑调整优化当前的相机位姿和三维点坐标。最后,基于当前的相机位姿和三维点坐标,利用PnP获取更多相机位姿,进而新建更多Tracks,并对Tracks进行滤波及Tracks重建,每添加4幅图像进行1次捆绑调整优化,得到包含叶片语义信息的大豆植株稀疏三维点云。

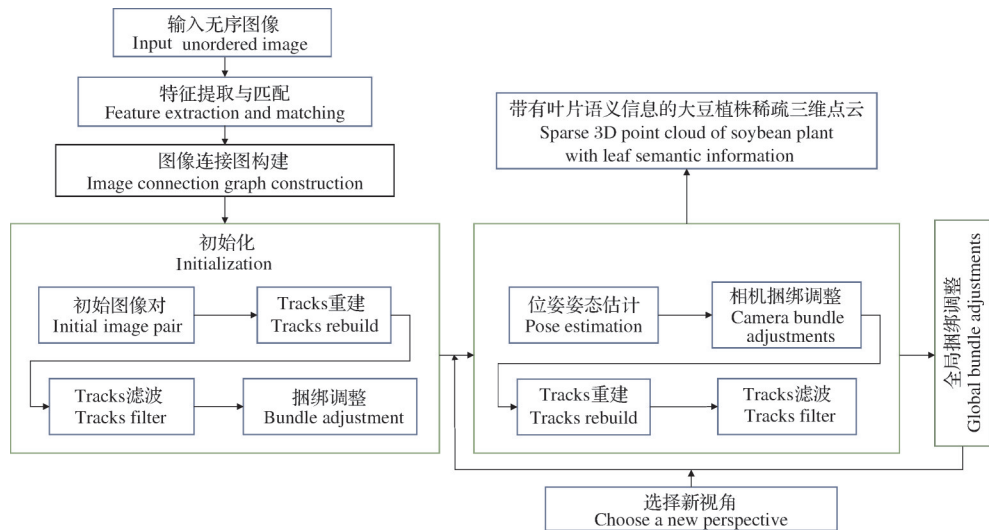


图5 包含叶片语义的大豆植株稀疏三维点云重建

Fig.5 Sparse reconstruction process of soybean plant with leaf semantics

②基于深度图融合的大豆植株叶片稠密语义重建。基于上述SfM得到的相机位姿和稀疏语义点云,进一步通过深度图融合方法对其进行稠密点云语义重建。如图6所示,首先,为每幅图像选择邻域图像构成立体图像对,根据对极约束,在对应视角基线上计算NCC(normalized cross correlation,归一化互相关性)值以寻找对应像素点,并通过相机位姿计算视差,同时将这些特征点记为种子点。其次,通过区域生长方法对这些种子点进行扩散,根据种子点

的NCC值创建优先级队列,每次扩散选取优先级高的点进行扩散,以保证种子点的深度可信赖。最后,对每个种子点均进行非线性深度优化,根据图像相似性原理,若优化过程中邻域像素缺失深度值,且种子点是周围NCC值最高的一个点,则将该种子点的patch法向量和深度初始值赋值给该邻域点。得到所有大豆植株深度图后,根据已知图像的相机位姿和深度图进行点云融合,进而得到包含叶片语义信息的大豆植株稠密点云。

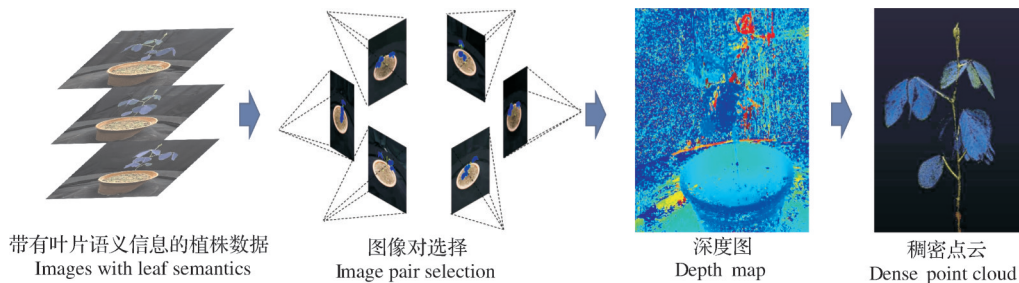


图6 包含叶片语义的大豆植株稠密点云重建流程

Fig.6 Flow chart of dense reconstruction of soybean plant with leaf semantics

3)叶片重建评价准则。由于重建过程会丢失尺度信息,为准确提取大豆植株叶片性状,我们基于大豆植株三维点云和实际植株获得其对应比例

$r = h_{\text{real}} / h_{\text{reconstructed}}$ 。其中, h_{real} 为花盆的真实直径, $h_{\text{reconstructed}}$ 为花盆重建模型的直径, r 为坐标比例。对应计算公式为 $\text{predicted} = r \times \text{reconstructed}$,其中

predicted 为带真实尺度的叶子点云测量值, reconstrued 为基于叶子点云自动获得的测量值。试验结果采用均方根误差(root mean square error, RMSE)和决定系数 R^2 (determination coefficient) 进行度量。RMSE 表示预测值和观测值之间差异的

样本标准偏差, R^2 则是实测值与预测值之间的模型拟合度, 取值范围为 $[0, 1]$ 。计算公式如下所示:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{actual}_i - \text{predicted}_i)^2} \quad (1)$$

$$R^2 = \frac{\sum_{i=1}^n (\text{actual}_i - \text{actual}_{\text{avg}}) \times (\text{predicted}_i - \text{predicted}_{\text{avg}})}{\left[\sum_{i=1}^n (\text{actual}_i - \text{actual}_{\text{avg}})^2 \right]^{0.5} \left[\sum_{i=1}^n (\text{predicted}_i - \text{predicted}_{\text{avg}})^2 \right]^{0.5}} \quad (2)$$

其中, predicted_i 和 actual_i 分别为第 i 个叶子的预测值和人工测量的实际值, $\text{predicted}_{\text{avg}}$ 和 $\text{actual}_{\text{avg}}$ 分别为预测值和实际值的平均值, n 为目标个数。

1.3 验证试验

为验证本研究提出的算法性能, 以盆栽大豆植株幼苗为材料, 进行如下几组试验: (1) 自然环境与自动密闭采集环境下三维重建性能分析; (2) 大豆植株图像语义分割性能分析; (3) 大豆植株叶片语义重建性能分析; (4) 不同生长期及视图数量的大豆植株重建性能分析。试验环境为 Ubuntu 18.04, Intel(R) Core(TM) i7-6700K CPU @ 4.00 GHz, NVIDIA GeForce GTX TITAN X, 开发语言为 Python 和 C++。详细试验数据和相关代码可通过如下链接获取: <https://github.com/xiaodongdong101/-git>。

2 结果与分析

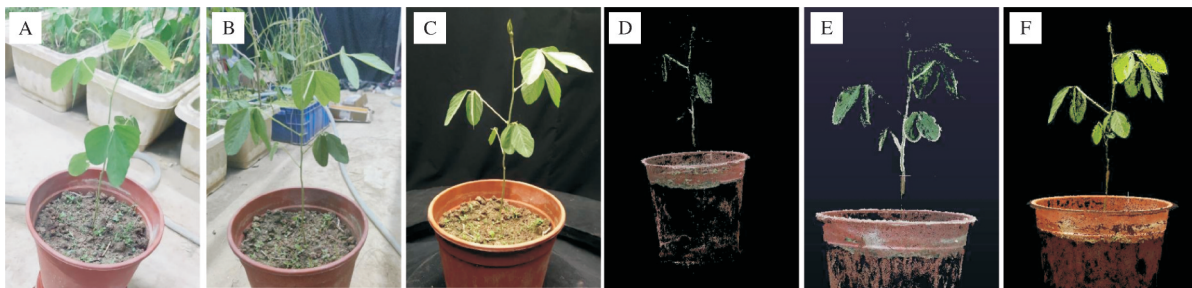
2.1 不同采集环境下的大豆植株三维重建

为验证本研究搭建的自动密闭环境采集的大豆植株三维重建效果, 分别对自然环境下采集的 20 张、40 张大豆植株图像和本研究搭建环境下采集的 20 张

大豆植株图像, 均采用 colmap 算法 (<https://github.com/colmap/colmap>) 进行多视图建模。重建效果如图 7 所示, 图 7A-C 为不同环境下采集的原始图像, 图 7D-F 为其对应的三维点云重建效果。由图 7D 和图 7E 可以看出, 随着用于重建的大豆植株图像数量增加, 三维重建效果逐渐变好。此外, 同等视图数量下, 如图 7A 和图 7C 均采用 20 张图像进行重建。图 7E 和图 7F 的重建效果表明, 尽管密闭环境下的视图数量比自然环境下减少一半, 但其三维重建效果仍优于后者。本研究后续试验数据均基于本文“1.1”搭建的密闭环境获取。

2.2 大豆植株叶片语义分割

采集不同大豆植株幼苗多株多视角图像 (2 800 多张叶片共 387 张图像) 进行叶片语义分割试验, 其中 80% 为训练集, 20% 为测试集。分割结果见图 8, 可以看出, 除了严重遮挡的部分叶片, 大部分叶片能较好地被蓝色掩膜覆盖, 这表明基于 Mask R-CNN 的分割模型可以获得较好的叶片分割结果, 它包含的二维语义信息可为大豆植株叶片三维语义重建提供语义迁移信息。



A 和 B 分别为自然场景采集的 20 张和 40 张图像随机样本, C 为密闭环境采集的 20 张图像随机样本; D、E、F 为其对应的重建效果。A and B are random samples of 20 and 40 images acquired in natural scenes, respectively, and C is a random sample of 20 images acquired in closed environment; D, E and F are their corresponding reconstruction point clouds.

图 7 不同采集环境大豆植株三维重建效果

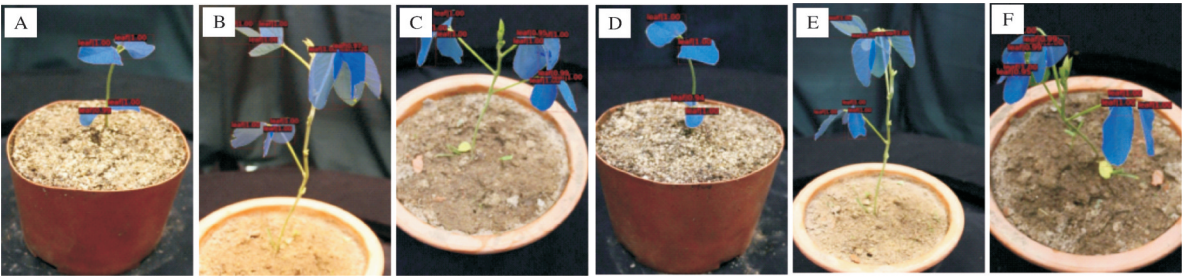
Fig.7 3D reconstruction of soybean plant under different scenes

2.3 大豆植株叶片语义重建效果

为验证大豆植株叶片三维语义建模效果, 进行以下 4 组试验: (1) 原始图像与带有掩膜图像的大豆植株重建; (2) 大豆植株叶片表型参数提取及误差分

析; (3) 不同大豆植株的叶片语义重建; (4) 同一大豆植株不同视图个数的大豆植株重建。

1) 基于原始图像与带掩膜图像的大豆植株重建。为验证基于原始图像与带有语义掩膜图像的大



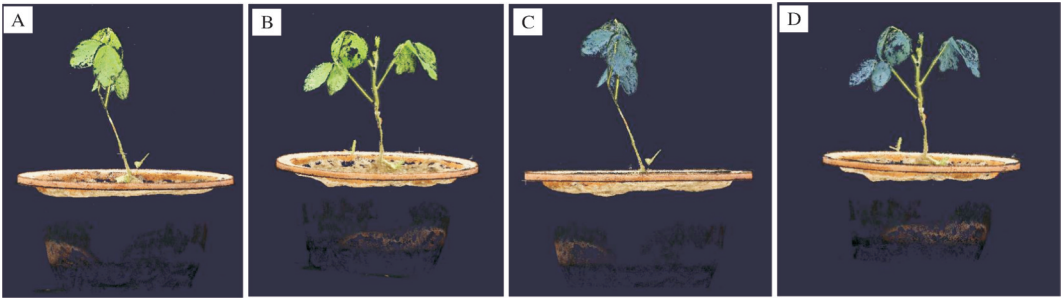
A、B 和 C 分别为正视图,D、E 和 F 分别为对应的侧视图。A, B, and C are front views;D, E, and F are their corresponding side views.

图 8 不同类型大豆植株幼苗叶片语义分割图

Fig.8 Leaf semantics segmentation of different soybean plant seedlings

豆植株重建效果,分别对 72 张原始大豆植株图像和 72 张带有叶片语义信息的大豆植株图像进行多视角重建。重建效果如图 9 所示,图 9A 和图 9B 为基于原始大豆植株图像生成的三维点云,图 9C 和图 9D 是基于带叶片语义信息大豆植株图像生成的三维点云图。图 9C 和 9D 所展示的点云正视图和侧视图中三

维叶片表面均基本被蓝色覆盖。此外,从图 9 可以看出,原始图像重建的三维点云与带有叶片语义信息重建的三维点云在叶片表面颜色上有明显区别,这表明在重建过程中二维图像中的叶片语义被较完整地迁移到三维叶片点云上,进而使得三维叶片点云包含语义信息。



A 和 B 分别为原始图像的大豆植株点云,C 和 D 分别为带叶片语义信息图像的大豆植株点云。A and B are the soybean plant point clouds of the original image, respectively; and C and D are the soybean plant point clouds with semantic information images of the leaves, respectively.

图 9 基于原始图像和带叶片语义信息图像的大豆植株点云图

Fig.9 Soybean plant point cloud based on with/without leaf semantic information

为进一步验证图像叶片语义信息是否影响大豆植株重建性能,表 1 为基于原始图像和带叶片语义信息图像的大豆植株重建计算时间及生成点云数量。从表 1 可以看出,两者在重建时间、原始点云数量及优化后的点云数量上差异较小,分别为 1.761 min、154 819 个和 2 498 个,这表明本研究提出的方法可在叶片语义迁移的同时较好地保证大豆植株重建性能。

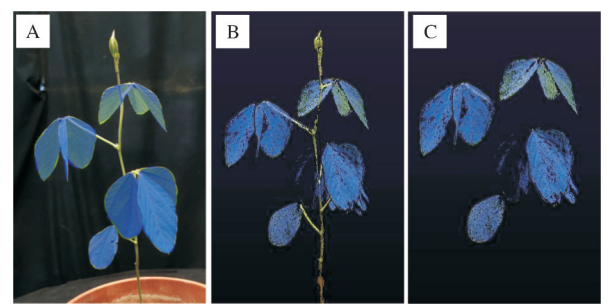
2)大豆植株叶片表型参数提取与分析。为验证大豆植株叶片三维语义的重建精度,对重建得到的大豆植株叶片语义点云进行叶长和叶宽测量,并与人工实测数据进行对比分析。首先,使用 20 张带语义信息的图像重建大豆植株;然后,基于叶片语义信息提取出叶片点云数据(图 10)。

计算图 10 大豆植株叶片点云中每个叶片的叶长(每片叶片的最长距离)和叶宽(垂直叶长的最大宽度),并基于公式(1)和公式(2)计算其与人工测量数

表 1 原始与带有语义信息的大豆植株图片重建开销对比
Table 1 Reconstruction cost of soybean plant with/without leaf semantic information

类别 Category	重建时间/min Reconstruction time	原始点云数 Number of original point cloud	优化后点云数 Number of optimized point cloud
原始数据 Original data	77.488	1 219 505	807 797
带语义信息数据 Data with semantic information	75.727	1 064 686	810 295
误差绝对值 Absolute value of error	1.761	154 819	2 498

据的误差(表 2)。大豆植株叶片点云中叶长和叶宽计算方式如下:首先对需测量的叶片进行噪声点滤波,然后计算叶片中任意 1 对点的欧式距离,距离最长两点为叶长;在此基础上,通过距离最长的 1 对点将叶片点云划分为上半区和下半区,并分别计算 2 个



A 为植株二维图像,B 为 A 对应的语义点云图,C 为 B 中提取的叶片点云图。A is a 2D image of the soybean plant, B is the semantic point cloud of corresponding to A, and C is the point cloud of the leaf extracted from B.

图 10 大豆叶片语义点云提取
Fig.10 Soybean leaf semantic point cloud extraction

表 2 大豆植株叶长、叶宽度量
Table 2 Phenotypic measurement of length and width of leaves mm

序号 Leaf	类型 Type	人工测量值 Manual measurement	点云测量值 Point cloud measurement	误差值 Error value
叶 1	长 Length	40	42	2
Leaf 1	宽 Width	25	26	1
叶 2	长 Length	46	46	0
Leaf 2	宽 Width	30	28	−2
叶 3	长 Length	37	38	1
Leaf 3	宽 Width	25	24	−1
叶 4	长 Length	60	64	4
Leaf4	宽 Width	38	37	−1
叶 5	长 Length	56	61	5
Leaf 5	宽 Width	35	39	4
叶 6	长 Length	60	62	2
Leaf 6	宽 Width	38	39	1
叶 7	长 Length	55	55	0
Leaf 7	宽 Width	32	34	2
叶 8	长 Length	57	59	2
Leaf 8	宽 Width	30	31	1
叶 9	长 Length	55	56	1
Leaf 9	宽 Width	30	32	2
叶 10	长 Length	46	49	3
Leaf 10	宽 Width	26	28	2

半区上距离该直线最长距离的两点,此两点距离之和为叶宽。从表 2 可以看出,与人工测量值相比,叶长和叶宽整体误差较小,其中叶长最大误差为 5 mm,最小误差为 0 mm, R^2 为 0.97,叶宽的最大误差和最小误差分别为 4 mm 和 0 mm, R^2 为 0.89,该结果表明本研究提出的方法具有较高的准确性,可较好地实现叶片语义信息的迁移学习。此外,叶长和叶

宽的均方误差分别为 2.53 和 1.92,叶长误差相对较大的原因可能来自点云测量值,叶片存在一定的弯曲、叶片遮挡等因素在基于重建点云测量时对叶长的影响较叶宽偏大。

由图 11 可以看出,叶长和叶宽整体的 $R^2=0.98$, $RMSE=2.25$ mm,研究得到的测量值和人工实测值具有较好的一致性,这表明本研究提出的语义建模方法在保证大豆植株三维建模精度的同时,可以较精准地实现其叶片三维语义建模。

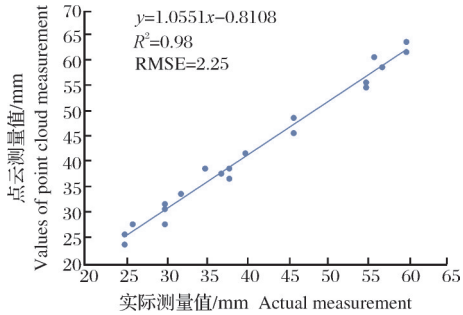
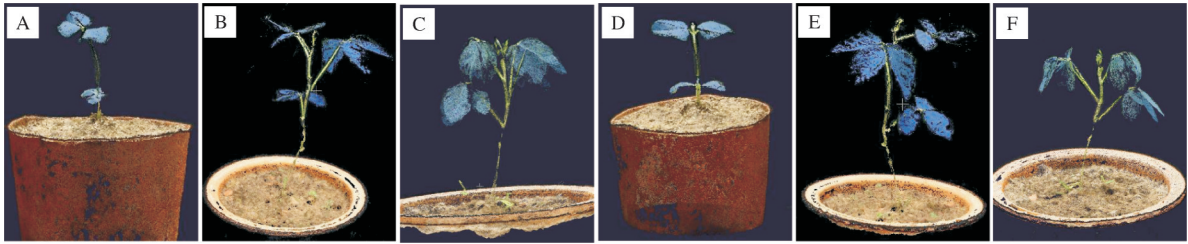


图 11 大豆植株叶片精度对比
Fig.11 Accuracy comparison of soybean leaf measurement

3)大豆植株叶片语义重建。为验证不同大小的植株叶片语义重建的效果,选择 3 组不同大小、每组带叶片语义信息的大豆植株 36 张图像进行试验,重建效果见图 12。从图 12 可以看出,图片数量相同情况下,不同大小的大豆植株,其语义重建效果存在一定差异。如图 12A 和 12D 所示,叶片较少且相互之间不遮挡或存在少量遮挡时,叶片点云语义重建效果较好。随着叶片密度增加,叶片语义重建的精度有所降低(图 12E 和 12F),叶片上存在一些空洞或边缘缺失,其原因在于叶片之间存在大量相互遮挡,导致其二维语义信息丢失,进而无法在重建过程中对其进行语义迁移学习。

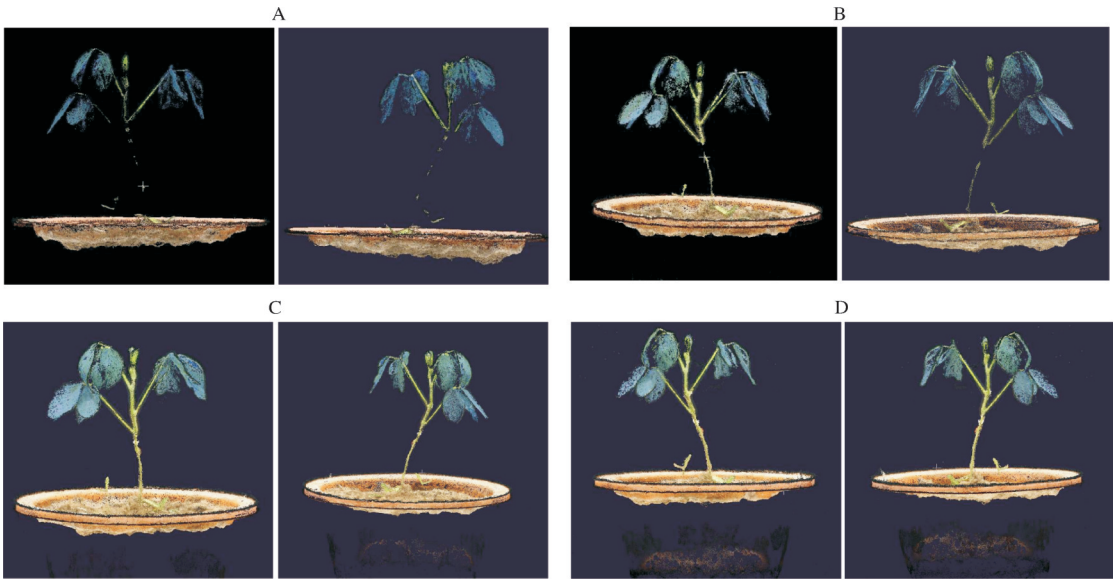
4)同株且不同数量视图的大豆植株语义重建。为验证不同视图数量的同一大豆植株的重建性能,选取包含不同视图数的同一植株叶片进行重建试验。该试验共分 4 组:第 1 组包含 18 张图像,图像间拍摄角度间隔为 20° ;第 2 组为 36 张图像,图像间拍摄间隔为 10° ;第 3 组共 54 张图像,其中 36 张图像拍摄间隔为 10° ,另外 18 张图像拍摄角度通过增加 5° 偏移后得到;第 4 组共 72 张图像,图像间拍摄间隔 5° 。图 13 和表 3 分别给出上述 4 组大豆植株数据的重建效果和重建性能对比。



A、B 和 C 分别为不同大小大豆植株点云正视图,D、E 和 F 为其对应点云侧视图。A, B and C are point cloud front views of soybean plants with different sizes, and D, E, and F are their corresponding point cloud side views, respectively.

图 12 不同大小大豆植株叶片语义点云图

Fig.12 Semantic point cloud of soybean plants with different sizes



A: 18 张带叶片语义信息的重建点云图 Point cloud based on 18 images with leaf semantic information; B: 36 张带叶片语义信息的重建点云图 Point cloud based on 36 images with leaf semantic information; C: 54 张带叶片语义信息的重建点云图 Point cloud based on 54 images with leaf semantic information; D: 72 张带叶片语义信息的重建点云图 Point cloud based on 72 images with leaf semantic information.

图 13 同株不同视图数量的大豆植株叶片语义重建

Fig.13 Semantic reconstruction of soybean plant leaves from the same plant with different images

从图 13A 和表 3 可以看出,植株点云中的茎秆和叶片等器官均出现断层,这是因为用于重建的图像数量较少,仅 18 张,导致获取信息不足,重建得到原始点云的数量和优化后点云的数量均较少,分别仅为 240 073 和 239 670。随着用于重建的图像数量增加,如图 13B 和图 13C,植株整体重建精度和叶片语义重建精度均有较大提升,原始点云数分别为 461 193 和 767 094,与图 13A 相比,增加了 92.1% 和 219%。不过,当用于重建的图像超过一定数量时,其重建精度并没有太大提升,如图 13C 和 13D,大豆植株重建精度近似,但后者重建开销却增加约 25%。这表明重建过程中在保证重建精度的前提下,可以通过递增方式对图像数量进行选择,以协调重建开支和重建精度的平衡。

表 3 带有语义信息的图像重建开销对比
Table 3 Reconstruction cost of soybean plant with leaf semantic information

图像重建数量 Number of images	重建时间/min Reconstruction time	原始点云数 Number of original point cloud	优化后点云数 Number of point cloud after optimization
18	17.968	240 073	239 670
36	37.731	461 193	451 157
54	56.369	767 094	668 552
72	75.727	1 064 686	810 295

3 讨 论

当前常用的植株三维建模手段(三维扫描仪、多视图建模等)缺少对器官分割和关键部位定位的语义信息,导致当重建的植株点云中叶片和茎秆等颜

色相近或遮挡时,较难从点云上判别所属器官,进而影响后续的株型解析。本研究提出一种二维先验语义引导的大豆植株叶片三维语义建模方法,可以较准确地将植株二维图像中的叶片语义信息迁移到三维叶片点云上(图9C和9D)。大豆植株的叶长和叶宽状测量结果(表2和图11)表明,本研究设计的方法在保证语义学习的同时,还可以确保植株叶片语义重建的准确性。基于该研究结果,可以借助不同语义信息对植株器官(如叶片、茎干等)进行自动提取和表型解析。此外,本研究的方法具有一定的通用性,可以应用到更广泛的作物三维语义建模。

本研究结果表明,利用深度学习技术对植株图像进行准确语义分割处理,可以实现基于二维图像重建的不同语义信息由二维到三维的迁移,进而得到包含不同语义信息的作物点云语义模型(图9),三维叶片表面基本被蓝色覆盖,这表明二维叶片语义信息较好地迁移到三维叶片点云上。此外,试验结果显示不同大小大豆植株的叶片语义重建效果有一定差异(图12),其原因在于叶片之间遮挡严重导致无法获得其二维语义信息。图像语义分割试验(图8)表明,叶片较少且相互之间少量遮挡时,可以较好地对大豆植株幼苗图像进行语义分割,但当图像中各器官遮挡严重时,会导致分割结果出现偏差,进而影响植株叶片点云建模性能。因此,选择并训练鲁棒性能好的图像语义分割算法对作物点云语义模型学习非常重要。

本研究结果还表明,相同重建技术下,相比开放环境下采集的图像数据,基于全自动密闭环境下采集相同数量图像可更高精度重建大豆植株三维点云(图7),其原因在于自动密闭环境下获得的大豆植株图像更干净,较少受目标之外的物体或者风等干扰。因此,为构建精准植株三维模型,需提供尽可能少外物干扰的环境。此外,不同数量的二维图像生成的三维点云模型的精度和时间有所不同,如表3所示,重建时间与图像数为正比关系,但随着图像数达到一定规模,重建精度并不会线性提升(图13C和13D),其原因在于图13C采用54张大豆植株图像进行重建,这些多视角图像可对大豆植株视角基本覆盖,此种情况下,进一步增加视图数据对重建精度提升有限,但视图数量的增多会增加大豆植株三维重建时间开销,如表3所示,用于重建的大豆植株视图数量从54增加到72,重建精度近似(图13C和13D),但重建时间却从56.369 min增加到75.727 min,增加

了34.3%。因此,在保证重建精度前提下,如何选择适当数量的二维图像及如何评价语义重建性能与图像数量之间的关系等有待进一步研究。

参考文献 References

- [1] WANG L, MING L, LIAO K Y, et al. Bract suppression regulated by the miR156 / 529-SPLs -NL1 -PLA1 module is required for the transition from vegetative to reproductive branching in rice [J]. *Molecular plant*, 2021, 14(7): 1168-1184.
- [2] LIU J, HUANG J, GUO H, et al. The conserved and unique genetic architecture of kernel size and weight in maize and rice [J]. *Plant physiology*, 2017, 175(2): 774-785.
- [3] PAN Q C, XU Y C, LI K, et al. The genetic basis of plant architecture in 10 maize recombinant inbred line populations [J]. *Plant physiology*, 2017, 175(2): 858-873.
- [4] 朱荣胜, 李帅, 孙永哲, 等. 作物三维重构技术研究现状及前景展望 [J]. *智慧农业*, 2021, 3(3): 94-115. ZHU R S, LI S, SUN Y Z, et al. Research advances and prospects of crop 3D reconstruction technology [J]. *Smart agriculture*, 2021, 3(3): 94-115 (in Chinese with English abstract).
- [5] 张慧春, 周宏平, 郑加强, 等. 植物表型平台与图像分析技术研究进展与展望 [J]. *农业机械学报*, 2020, 51(3): 1-17. ZHANG H C, ZHOU H P, ZHENG J Q, et al. Research progress and prospect in plant phenotyping platform and image analysis technology [J]. *Transactions of the CSAM*, 2020, 51(3): 1-17 (in Chinese with English abstract).
- [6] 徐凌翔, 陈佳伟, 丁国辉, 等. 室内植物表型平台及性状鉴定研究进展和展望 [J]. *智慧农业*, 2020, 2(1): 23-42. XU L X, CHEN J W, DING G H, et al. Indoor phenotyping platforms and associated trait measurement: progress and prospects [J]. *Smart agriculture*, 2020, 2(1): 23-42 (in Chinese with English abstract).
- [7] ZHENG Q, FAN X C, GONG M L, et al. 4D reconstruction of blooming flowers [C]. *Computer graphics forum*, 2017, 36(6): 405-417.
- [8] 傅汝佳, 冼楚华, 李桂清, 等. 面向表型精确鉴定的豆株快速三维重建 [J]. *浙江大学学报(理学版)*, 2021, 48(5): 531-539. FU R J, XIAN C H, LI G Q, et al. Rapid 3D reconstruction of bean plant for accurate phenotype identification [J]. *Journal of Zhejiang University (science edition)*, 2021, 48(5): 531-539 (in Chinese with English abstract).
- [9] LIU Y D, YUAN H B, ZHAO X, et al. Fast reconstruction method of three-dimension model based on dual RGB-D cameras for peanut plant [J/OL]. *Plant methods*, 2023, 19: 17 [2022-11-19]. <https://doi.org/10.1186/s13007-023-00998-z>.
- [10] ANDO R, OZASA Y, GUO W. Robust surface reconstruction of plant leaves from 3D point clouds [J/OL]. *Plant phenomics*, 2021, 2: 3184185 [2022-11-19]. <https://spj.science.org/doi/10.34133/2021/3184185>.
- [11] ZERMAS D, MORELLAS V, MULLA D, et al. 3D model processing for high throughput phenotype extraction - the case of corn [J/OL]. *Computers and electronics in agriculture*, 2020, 172: 105047 [2022-11-19]. <https://doi.org/10.1016/j.compag.2019.105047>.

- [12] LI Y C, LIU J Y, ZHANG B, et al. Three-dimensional reconstruction and phenotype measurement of maize seedlings based on multi-view image sequences [J/OL]. *Frontiers in plant science*, 2022, 13: 974339 [2022-11-19]. <https://doi.org/10.3389/fpls.2022.974339>.
- [13] 孙凯. 基于三维重构的普兰种业主栽大豆种植密度的优化研究 [D]. 哈尔滨: 东北农业大学, 2019. SUN K. Study on optimization of soybean planting density based on three-dimensional reconstruction of Pulan plantation [D]. Harbin: Northeast Agricultural University, 2019 (in Chinese with English abstract).
- [14] 冯佳睿, 马晓丹, 关海鸥, 等. 基于深度信息的大豆株高计算方法 [J]. *光学学报*, 2019, 39(5): 258-268. FENG J R, MA X D, GUAN H O, et al. Calculation method of soybean plant height based on depth information [J]. *Acta optica sinica*, 2019, 39(5): 258-268 (in Chinese with English abstract).
- [15] DOI T, OKURA F, NAGAHARA T, et al. Descriptor-free multi-view region matching for instance-wise 3D reconstruction [C]//Computer Vision-ACCV 2020: 15th Asian Conference on Computer Vision, November 30-December 4, 2020, Kyoto, Japan. Kyoto: Springer International Publishing, 2021: 581-599.
- [16] WU S, WEN W L, WANG Y J, et al. MVS-Pheno: a portable and low-cost phenotyping platform for maize shoots using multi-view stereo 3D reconstruction [J/OL]. *Plant phenomics*, 2020, 12: 1848437 [2022-11-19]. <https://spj.science.org/doi/10.34133/2020/1848437>.
- [17] WU S, WEN W L, GOU W B, et al. A miniaturized phenotyping platform for individual plants using multi-view stereo 3D reconstruction [J/OL]. *Frontiers in plant science*, 2022, 13: 897746 [2022-11-19]. <https://doi.org/10.3389/fpls.2022.897746>.
- [18] DAS-CHOUDHURY S, MATURU S, SAMAL A, et al. Leveraging image analysis to compute 3D plant phenotypes based on voxel-grid plant reconstruction [J/OL]. *Frontiers in plant science*, 2020, 11: 521431 [2022-11-19]. <https://doi.org/10.3389/fpls.2020.521431>.

Soybean leaf 3D semantic reconstruction for plant phenotype analysis

GAO Yuefang¹, XIAO Dongdong¹, FU Rujia², XIAN Chuhua²,
LI Guiqing², HUANG Qiong^{1,4}, YANG Cunyi³

1. College of Mathematics and Informatics, South China Agricultural University,
Guangzhou 510642, China;

2. School of Computer Science & Engineering, South China University of Technology,
Guangzhou 510006, China;

3. College of Agriculture, South China Agricultural University, Guangzhou 510642, China;

4. Guangzhou Key Laboratory of Intelligent Agriculture, Guangzhou 510642, China

Abstract The 3D point clouds obtained from 3D scanner and multi-view data lack semantic information, leading to difficulties in discriminating the plant organ parts from the point clouds when the number of plant point clouds is large, or when different organs of the plant are similarly colored or obscured. To deal with the problem, this article proposes a three-dimensional semantic modeling method for soybean leaves embedded with a two-dimensional semantic prior. The semantic segmentation of soybean leaves based on Mask R-CNN was conducted. The three-dimensional reconstruction, fusion and learning of the segmentation results and multi-view data were performed to transfer the semantic information of leave from 2D semantics to 3D point clouds and obtain the point cloud semantic information of plant leaf. The 3D semantic model of plant leaves was established. The model was validated through multiple sets of potted soybean plant experiments. The length and width of leaf were extracted and compared with the manual measurement data. Results showed that the mean square error of the length and width of leaf was 2.53 and 1.52 mm, with the determination coefficients of 0.97 and 0.89, respectively. It is indicated that the proposed method can conveniently and accurately construct the 3D semantic model of plant leaves.

Keywords semantic segmentation; three-dimensional modeling of plant; deep network model; semantic reconstruction of point cloud; phenotype analysis of crop

(责任编辑: 张志钰)