

基于多视角图像的作物果实三维表型重建

杨会君^{1,4}, 王瑞萍¹, 王增莹³, 王 昕²

(1.西北农林科技大学信息学院, 陕西 杨凌 712100)

(2.西北农林科技大学外语系, 陕西 杨凌 712100)

(3.三只松鼠南京研发创新中心, 江苏 南京 210019)

(4.农业农村部物联网重点实验室, 陕西 杨凌 712100)

[摘要] 针对基于激光扫描设备获取点云存在操作复杂、成本高、难以被普及等问题, 本文研究了基于普通图像的复杂背景中作物果实三维表型重建。我们建立了集 SFM 算法、PMVS 算法以及半自动化去噪方法的优势为一体的三维重建架构。以一组多视角目标作物果实二维图片为输入源, 首先基于 SIFT 算子的比例和旋转不变性参数, 提取多幅二维图像特征信息。其次, 结合 FLANN 算法实现不同角度的数据匹配, 并提出了基于二维图像关键点和相机参数等信息的稀疏点云快速生成方法。然后, 基于 PMVS 初始特征匹配的种子面片提取、扩散获取密集面片, 进一步利用可见性约束过滤不正确匹配导致的错误面片, 以实现复杂果实点云模型生成。最后, 我们提出了交互式选择和滤波器相结合的、半自动化的果实表型离群点去除方法, 解决了作物果实模型的准确重建问题。结果表明, 本文的方法能有效解决复杂实验环境中果实表型数据的低成本、准确、方便快捷获取问题。

[关键词] 三维果实表型, 稀疏点云算法, 稠密点云算法, 去噪

[中图分类号] O4-39, TP37 **[文献标志码]** A **[文章编号]** 1001-4616(2021)02-0092-12

Three-Dimensional Phenotypic Reconstruction of Crop Fruit Based on Multi-View Image

Yang Huijun^{1,4}, Wang Ruiping¹, Wang Zengying³, Wang Xin²

(1.College of Information Engineering, Northwest A&F University, Yangling 712100, China)

(2.Department of Foreign Languages, Northwest A&F University, Yangling 712100, China)

(3.Three Squirrels Nanjing R&D Innovation Center, Nanjing 210019, China)

(4.Key Laboratory of Agricultural Internet of Things, Ministry of Agriculture and Rural Affairs, Yangling 712100, China)

Abstract: Obtaining point cloud based on laser scanning equipment has many problems such as complicated operation, high cost, and difficulty in popularization. Therefore, this paper studies the three-dimensional phenotypic reconstruction of crop fruit in complex background based on common images. We have built a 3D reconstruction architecture integrating the advantages of SFM algorithm, PMVS algorithm and semi-automatic denoising method. Taking a set of multi-view two-dimensional images of target crop fruits as input sources, we first extracted the feature information of multiple two-dimensional images based on scale of the SIFT operator and rotation invariance parameters. After that, we combined the FLANN algorithm to achieve data matching from different angles. Furthermore, we proposed a fast method of generating sparse point clouds based on the information of key points of 2D pictures and camera parameters. Then, the dense patches are obtained by seed patch extraction and diffusion based on the initial feature matching of PMVS algorithm. We further used visibility constraints to filter out the wrong patches caused by incorrect matching, and realized the generation of complex point cloud model. Finally, we proposed a semi-automatic method to remove outliers from fruit phenotypic point cloud by combining interactive selection and filter, which solved the problem of accurate reconstruction of crop fruit model. The results show that this method in this paper can effectively solve the problem of low-cost, accurate, convenient and fast acquisition of fruit phenotype data in complex experimental environment.

Key words: three-dimensional fruit phenotype, sparse point cloud algorithm, dense point cloud algorithm, denoising

收稿日期: 2020-09-10.

基金项目: 陕西省重点研发计划项目(2021NY-179, 2019ZDLNY07-02-01, 2020NY-205)、大学生创新创业训练计划项目(S202010712238、S202010712063、X202010712373)。

通讯作者: 王增莹, 工程师, 研究方向: 计算机图形学. E-mail: 976104383@qq.com

计算机三维重建技术在农业上的应用是目前国内外的一个热点研究课题. 通过建立三维果实模型为进行农作物果实生长规律分析和测量提供了必要手段.

三维数字化技术的出现极大地促进了作物果实三维表型的精确建模研究. 但传统三维数字化技术在获取数据的过程中会接触并扰动作物,且获取时间较长,使非接触式的作物三维重建技术的大力开发与广泛应用成为果实表型精准数字化领域亟待解决的问题.

研究者们通常采用两种三维重建方法. 其中一类为通过传感器直接获取作物果实的三维点云^[1]. 此类方法能够获得高精度的果实点云,但存在所需仪器设备复杂昂贵、数据量大、重建效率低、使用中受环境如光照限制较大等问题. 另一类为通过二维图像恢复作物果实的三维表型点云^[2]. 此类方法操作简单、价格低廉,生成的点云模型虽不复杂但往往包含很多噪点,且存在使用中受相机定标的精确程度、环境光影影响较大等问题^[3]. 大量研究表明基于多视角拍摄二维图像序列的、能够自动进行相机标定的、从运动恢复结构(structure from motion, SFM)的重建方法具有很大的优势. 但对于作物果实、植株等结构复杂的目标物体,基于 SFM 方法获取的稀疏点云所包含的三维信息较少. 为了精确表示果实、植株的表型特征,得到真实的重建效果,需要获取包含更多三维信息的稠密果实点云,因此,在稀疏获取作物果实点云的基础上进行稠密解析是三维重建过程的重要步骤.

本文通过将 SFM 方法与多视角立体视觉(multiple view stereo, MVS)方法相结合,对稀疏点云模型进行面片生成、扩散与过滤,从而生成了包含较多三维信息的稠密果实点云. 然而基于多视角图像生成的点云模型通常包含大量噪点,对后续的曲面重构阶段造成很大的麻烦^[4]. 因此,本文提出了一种针对不同噪点类型的交互式去噪与滤波去噪相结合的半自动化去噪方法. 首先去除模型中数量大、排列紧密、肉眼清晰可见的噪点,之后自动去除附着于点云模型表面的噪点以及排列稀疏的噪点,使模型更加光滑. 针对点云获取设备操作复杂、成本高、受环境影响较大等问题,本文提出了集 SFM 算法、PMVS(patch based multi-view stereo)算法、半自动化去噪方法为一体的果实表型重建策略,对从不同角度拍摄的苹果、菠萝、榴莲、桃子等果实的图像序列进行处理,完成了复杂实验背景中作物果实的三维表型重建. 主要贡献有:

(1)在获取作物果实表型稀疏点云的过程中,融合 SIFT 算子、旋转不变性参数、FLANN 算法和相机位姿,实现二维图像特征信息的提取和匹配,解决稀疏点云的快速生成问题;

(2)对稀疏点云进行稠密解析时,引入基于面片的多视图立体视觉算法 PMVS,以生成包含较多三维信息的稠密果实点云模型,解决重建效果真实性的问题;

(3)在作物果实表型重建过程中,融合交互式去噪与滤波去噪方法,实现对不同类型噪点的去除,解决果实点云曲面重构的准确性问题.

1 国内外研究现状

计算机图形技术在农业生产中具有广泛的应用,是目前农业领域中最热门的研究课题之一. 国内外学者已做了大量的研究,其中已有部分成果应用于实际的生产. 图形技术在农业上的应用研究主要包括基于模型的虚拟植物建模^[5]和基于扫描数据集的三维建模. 由于现实世界中作物果实生长环境复杂、果实形态存在随机性和多样性,因此很难通过现有的固定模型表达果实的外观. 基于扫描数据集的三维建模则可能受到扫描精度和数据全面性的影响,存在诸多的不确定性. 因此,由于各种因素计算机图形技术在精准农业果实表型重建方面的应用存在很大的局限性.

随着三维数字摄影技术和三维扫描技术的发展,设备精度的提高,使得具有丰富细节的植物果实表面高密度数据点的获取成为可能^[6],传感设备及扫描设备如激光扫描仪、三维数字化仪等在农业领域的应用也更加广泛. 三维激光影像扫描技术是 20 世纪 90 年代中期开始出现的一项高新技术,是继 GPS 空间定位系统之后又一项测绘技术新突破,是一种崭新的革命性的测量工具^[7]. 地面三维激光影像扫描仪是一种集成了多种高新技术的新型测绘仪器,采用非接触式高速激光测量方式,以点云的形式获取地形及复杂物体三维表面的阵列式几何图形数据. 作为一种直接获得物体表面三维信息的手段,激光成像测量技术可以快速生成地物的三维高精度点云模型^[8]. 通过在点云模型引入控制点坐标,可以让点云模型的可靠性与全站仪获得的点位数据相媲美. 但由于三维激光影像扫描设备价格高昂、操作复杂且采集的数据量太大,导致作物果实表型重建效率低下,因此在精准农业领域无法进行全面推广.

在微软的 Kinect 问世之后,不少学者使用 Kinect 进行了快速甚至实时重建. Henry^[9]通过 Kinect 对室内场景进行了三维重建,虽然效果和速度一般,但是结果是成功的. 2011 年起,微软的英国剑桥研究院发起了 KinectFusion 项目,利用 Kinect 的深度数据实时地跟踪传感器的三维姿态,融合不同帧得到的数据从而重建出整体的模型^[10-11]. KinectFusion 虽然可以达成实时重建,但是速度不够理想,对于动态物体不敏感^[12]. Kinect 的出现也为广大农业科技研究人员提供了新的思路与方向. Sasaki Yutaka 为了结合作物生长条件及环境分析劳作者的行为轨迹,再挖掘作业过程中的一些隐性信息,利用 Kinect 设计了一个自动农业劳作轨迹记录系统为农业生产系统化技术终端的计算机决策提供建议,但是 Kinect 受光照影响较大,所以只能在清晨或者傍晚进行采集^[13].

基于图像进行三维重建研究一直是计算机视觉领域的热门课题,仅通过二维图像的变化就可得到具有真实感的三维模型^[14]. 利用图像进行三维建模的技术从获取多视图图像开始,逐步通过相应的算法恢复场景中物体的三维结构^[15-16]. 基于图像的三维重建技术一般分为三个步骤:稀疏重建、稠密重建和表面重建,能够根据使用情况的不同对生成的三维模型的结构进行调整,也可以只针对某一特定的部分进行优化完善,解决了传统三维建模方法的不足. 然而由于存在以下挑战使直接基于图像序列建模依然无法应用于农业领域的作物果实三维表型重建:一方面,对于特征复杂的作物果实,拍摄过程易受环境光影响进而出现遮挡、空洞等问题^[17];另一方面,重建过程复杂程度高,相关技术不成熟. 本文拟将运动恢复结构(SFM)、基于面片的多视角视觉(PMVS)和半自动化去噪方法引入精准农业果实表型的三维重建过程,以解决作物果实表型点云的快速获取和精准、高效的三维重建问题.

综上,三维重建技术目前的三维重建方法有:(1)基于生长规则的建模方法,如 L-system^[18]、过程式参数方程等. 它们仅需要少量的参数控制某些基于植物学或者数学的模型生长规则就可以获得植物模型,该方法真实还原度低.(2)基于图像的三维重建,有基于单幅图像和基于图像序列的方法,该方法复杂度高.(3)基于激光点云的重建方法. 利用 FastSCAN、Focus3D、HandSCAN 等激光扫描仪获取三维点云数据,再进行模型重建. 该方法技术简易,但是扫描设备价格高昂^[19]. 到目前为止,尽管国内外在作物三维重建领域取得了丰硕的研究成果,但仍存在重建效果真实感低、重建方案复杂度高、扫描设备价格高昂等问题.

探索一种低成本、便捷、易操作、准确高效的作物果实表型三维重建方法成为精准农业领域急需解决的问题.

2 直接基于图像序列的果实点云获取

通过图像序列数据采集要求、基于 SFM 算法稀疏果实点云模型重建和基于 PMVS 算法密集点云模型生成,给出了果实表型精准重建架构.

2.1 硬件与图像序列数据采集

本文通过普通手机或数码相机拍照的方式,获取目标作物果实的一系列图像. 为提高匹配效果和效率,使用普通拍照方式获取目标作物果实不同视角的二维图片素材,相邻 2 张素材的角度变化幅度应不大于 10 度为最佳取材参数. 基于序列图像进行三维点云重建的核心是匹配图片之间的特征点,因此在拍摄照片时尽量保证同一面能从 2 个以上的角度进行拍摄.

2.2 稀疏点云获取

在采集到作物果实的序列图像后,首先要解决的是从图像进行稀疏点云生成问题. 本文首先利用 SIFT 检测并提取图像的特征点,然后使用 FLANN 算法对不同维度的特征点进行匹配,根据匹配的结果,利用射影定理计算相机相对位姿与三维点的坐标等场景信息,最终得到稀疏点云模型.

(1)基于 SIFT 算子的图像特征点提取. 进行稀疏重建的第一步是特征提取,采用具有尺度和旋转不变性特点的 SIFT 算子检测图像的特征信息. 该算法的实质是在不同的尺度空间上查找关键点(特征点),并计算出关键点的方向. 用 SIFT 描述子提取的特征点包含四个部分:位置,尺寸,方向,特征描述. 采用不同尺寸的高斯滤波器 DOG 计算出关键点的位置,并通过式(1)、(2)采集其所在高斯金字塔图像 3σ (尺度)领域窗口内像素的梯度和方向分布特征. 在完成关键点的梯度计算后,使用直方图统计邻域内像素的梯度和方向,并以峰值方向作为关键点的主方向. 至此,所有关键点都拥有三个信息:位置、尺度和方向.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (1)$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (2)$$

式中, $m(x, y)$ 为梯度的模值, $\theta(x, y)$ 为方向, L 为关键点所在的尺度空间值。

接下来为每个关键点建立一个不随光照、视角变化而改变的描述符。将关键点附近的区域划分 4×4 个方格子区域, 每个子区域内的梯度值分配到 8 个方向上并计算其权值, 由此生成一个 $4 \times 4 \times 8 = 128$ 维的梯度信息即特征点的特征向量, 以便直接通过特征向量比较其两点信息, 而不需要考虑其他参数。如此, 同样的特征点在理论上会有相同的描述向量, 故在不同的尺寸, 方向, 光照下都能被检测到。

(2) 基于快速近似最近邻算法 (FLANN) 的图片特征匹配。提取特征点之后需要对两图像之间进行特征匹配, 本文以对任何维度的数据进行匹配的快速近似最近邻算法 FLANN 为基础实现图片特征匹配。首先选取一个特征点并建立 kd-tree 数据结构进行邻域搜索, 给定一个阈值作为邻域点的范围取值, 当搜索到的邻域点距离小于这个阈值时就认为匹配成功。然而, 这样得到的匹配对的误匹配比较多, 故采取两种方法对误匹配进行修正: (1) 如果最近距离与次近距离的比值大于某个阈值, 则剔除。 (2) 对匹配点采用采样一致性算法 RANSAC 的八点法计算基础矩阵, 剔除不满足基础矩阵的匹配对。其次, 在匹配关系建立后, 通过将多个图像中都出现的共同特征匹配点进行组合, 生成同名点的集合, 即 track 列表。如, 假设图像 2 的 22 号点和图像 5 的 3 号点及图像 9 的 110 号点是匹配点, 则 (2, 22)、(5, 3)、(9, 110) 属于同一个 track, 多个 track 可以合并为一个 tracks 集合。同时为提高图片特征匹配的准确性, 需要在生成 track 时过滤无用匹配: 如果一个 track 多次包含同一幅图, 即同一幅图的多个特征点都匹配了同一个点, 则认定这个匹配关系不准确而将其过滤; 如果 track 太少, 也应该剔除, 因为如果仅有两幅图有同一个点, 那么三维重建的信息过少, 很容易产生误匹配。

(3) 基于相机位姿与三维坐标的稀疏点云提取。既特征匹配之后, 通过运动恢复结构 (struct-from-motion) 从图像中提取稀疏点云。首先, 寻找相机基线最大的像对作为初始化像对, 其中对图像采用 RANSAC 算法四点法计算单应矩阵, 满足单应矩阵的匹配点称为内点, 不满足单应矩阵的称为外点, 找到一个内点占比最小的像对就是初始化像对, 同时需要通过匹配点个数保证它是可重建的。然后对初始化像对进行相对定向, 根据 RANSAC 八点法计算本征矩阵, 通过对本征矩阵 SVD 分解得到第二张图像的相机位置旋转矩阵 R 和平移向量 T , 然后根据 R 、 T 和矫正后的像点坐标三角反向计算三维点坐标。其次, 通过调整空间变换和三维点信息使反向投影差最小, 进行一次 Bundle Adjustment 非线性优化使重建误差降到最小。之后通过 Bundle Adjustment 迭代过程进行优化, 在第一次迭代基础上, 将所有三维点反向投影到各自图片的像素坐标并分别与初始坐标比对, 如果大于某个阈值, 则将其从 track 中删除, 如果 track 中的匹配对已经小于 2 个了, 则整个 track 也去掉, 一直优化到没有点可去为止。之后加入更多图像, 直到没有合适的图像加入时, BA 结束。最后得到相机的相对位姿和三维点坐标即为果实的稀疏点云模型。综合以上内容, 作物果实稀疏点云获取算法的伪代码由 Algorithm 1 给出。

Algorithm 1: rare point calculation method

```

Input: Collection of image sequences  $C$ 
Output: Target result practical rare point  $P = \{X_i\}$ 
Output: Camera internal and external parameters  $K_l, R_l, T_l$ , image  $I \in C$ 
1 for Image  $I \in C$  do
2   Extract the feature point set  $F(I)$  of the image  $I$ 
3 end
4 for Image pair  $I, J \in C$  do
5   for Each feature point  $f \in F(I)$  do
6     Find the nearest neighbor  $f_{nn}$  of  $f$  in  $F(J) : f_{nn} \leftarrow \underset{f' \in F(J)}{\operatorname{argmin}} \|f - f'\|_2$ 
7   end
8   Basic rectangle  $F \leftarrow \{(f, f_{nn}) \mid f \in F(I), f_{nn} \in F(J)\}$  eight-point arithmetic
9   for  $f \in F(I)$  do
10    if Characteristic point formula  $|f_{nn}^x - Ff| > \text{sword } t$  then
    
```

```

11 Delete outliers that do not satisfy  $F: F(I) \leftarrow F(I)/f, F(J) \leftarrow F(J)/f_{nn}$ 
12 end
13 end
14 end
15 for Image  $I \in C$  do
16 for  $f \in F(I)$  do
17  $C \leftarrow C/I$ 
18 Combine the common feature points of all image pairs into a feature point trajectory:  $\text{Track}(f) \leftarrow \{f_{nn} | f_{nn} = \underset{f \in F(J)}{\operatorname{argmin}} \|f_d - f_d'\|_2, J \in C\}$ 
19 end
20 end
21 for Image pair  $I, J \in C$  do
22 Homography matrix of image pairs:  $H \leftarrow \{(f, f_{nn}) | f \in F(I), f_{nn} \in F(J)\}$  four-point algorithm
23 if Characteristic point formula  $\|f_{nn}^x - Hf\| > \text{sword } t$  then
24 External points not satisfying  $H$ : outlier ++
25 else
26 Interior points satisfying  $H$ : inlier ++
27 end
28 Proportion of interior points:  $\text{score}(I, J) \leftarrow \text{inlier}/(\text{inlier} + \text{outlier})$ 
29 end
30 select Image pair  $I, J \in C$  s.t minimization  $\text{score}(I, J)$  and matched interior points  $> 100$ 
31 Essential matrix  $E \leftarrow \{(f, f_{nn}) | f \in F(I), f_{nn} \in F(J)\}$  five-point algorithm
32 Camera external parameters:  $R_I, R_J, T_I, T_J \leftarrow \text{decompose the essential matrix } E$ 
33 Triangulation, get the characteristic point trajectory coordinate  $X_i$ 
34 Beam adjustment method to optimize camera parameters:  $K_I, K_J, R_I, R_J, T_I, T_J, X_i$ 
35 repeat
36 Add new image  $M$  s.t  $M$  has the most  $X_i$  observed
37 DLT to solve the internal and external parameters of new image  $M$ :  $K_M, R_M, T_M$ 
38 Triangulation, get  $X_M$ 
39 Beam adjustment method to optimize camera parameters:  $K_M, R_M, T_M, X_M$ 
40 until No image

```

2.3 密集重建

本文通过基于面片的多视图立体视觉算法 PMVS 进行密集重建,校准多视图立体视觉,输出一组密集的矩形斑块,覆盖在输入图像的可见表面. 通过特征提取和匹配来获得果实及背景的三维稀疏点,并借助矢量矩形面片模型进行描述,然后通过面片扩散形成致密的三维点. 算法的过程如图 1 所示.

首先将 n 张果实图片中的每个图像 $I_i (i = 1, 2, \dots, n)$ 划分成 $\beta \times \beta$ 像素大小的图像单元 $C_i(x, y)$, 并尝试在每个图像单元中重建至少一个面片 patch. 主要通过初始特征匹配、面片扩展和面片过滤三步进行稠密重建.

(1) 初始化特征匹配. 利用 Harris 和 Dog (Difference-of-Gaussians) 操作符在多张图片上检测特征点并进行匹配,生成一组与突出图像区域相关的稀疏面片 patch 集;

(2) 面片扩展是算法最核心的一部分. 将初始种子面片扩展至其邻近像素,得到一组密集的面片,以确保在每个图像单元 $M_i(x, y)$ 中重建至少一个面片. 选择一个种子面片 p , 首先获得一个满足如式(3)条件的邻域图像块集合 $C(p)$.

$$C(p) = \{C_i(x', y') | p \in Q_i(x, y), |x - x'| + |y - y'| = 1\} \quad (3)$$

式中, x, y 是图像块的下标, $Q_i(x, y)$ 用于记录投射到图像 $V(p)$ 中每个图像块 $M_i(x, y)$ 的所有点云片的集合. 以此为基础重建一个新的面片 p' , 初始化新面片的中心点 $r(p')$ 为射线穿过图像单元 $M_i(x, y)$ 中心和种子面片 p 所在平面的交点, 面片 p' 对应的参考图像 $R(p')$ 、图像 $V(p')$ 和单位法向量 $n(p')$ 由面片 p 的相应值初始化. 通过 $V(p')$ 可得到面片 p' 的可视化图像集 $V^*(p')$, 进而通过最小化光度一致性分数 $g^*(p')$, 实现 $r(p')$ 和 $n(p')$ 的优化, 如式(4)、(6). 优化完成后, 将深度图测试判定的一系列可视图像添加到图像

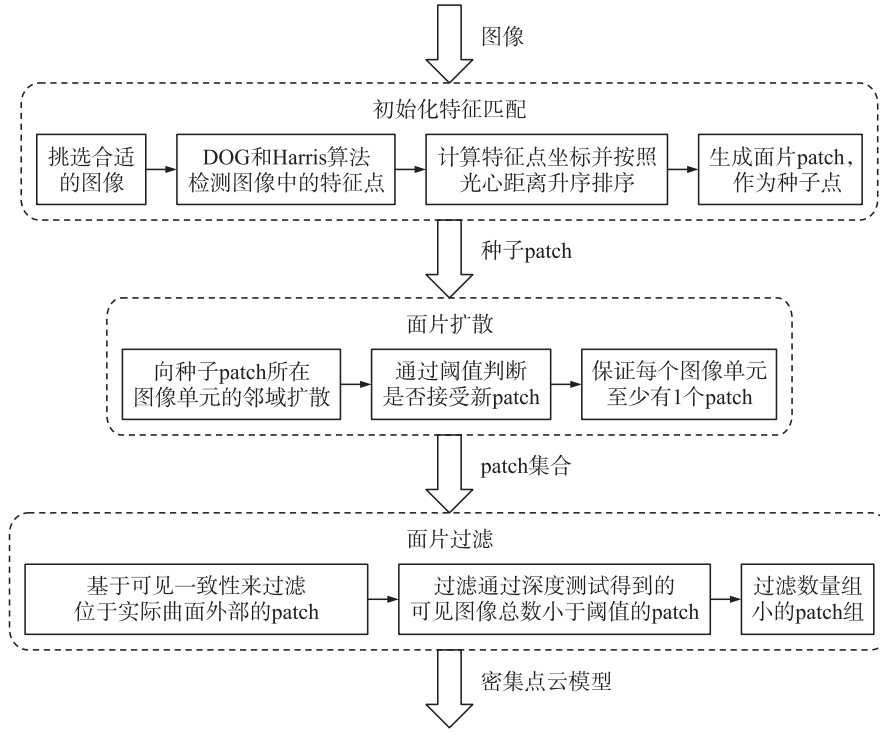


图1 稠密重建算法流程

Fig. 1 Dense reconstruction algorithm flow

集合 $V(p')$ 中,并据此更新 $V^*(p')$ 的值,如式(4)、(5). 最后,如果 $V^*(p')$ 大于一个阈值则接受新面片作为成功面片.

$$V^*(p) = \{I_i | I_i \in V(p), h(p, I_i, R(p)) \leq \alpha\} \quad (4)$$

$$V(p) = \{I_i | I_i \in I, d_{I_i}(p) \leq d_{I_i}(x, y) + \rho_1\} \quad (5)$$

$$g^*(p) = \frac{\sum_{I_i \in (V^*(p) - R(p))} h(p, I_i, R(p))}{|V^*(p) - R(p)|} \quad (6)$$

式中, $V^*(p) - R(p)$ 是指除了 $R(p)$ 的 $V^*(p)$ 的其他元素; $h(p, I_i, R(p))$ 表示图像 I_i 和 $R(p)$ 之间的成对光度差函数.

(3) 过滤: 使用可见性约束来消除不正确的匹配. 在进行面片扩散时,可能会生成一些不准确的面片,本文结合三个过滤器来过滤错误的面片.

首先基于可见一致性实现第一个过滤器,通过可见一致来移除位于实际曲面外部的面片. 设 $U(p)$ 为当前可见信息不连续的面片集合. 对于 $U(p)$ 中的面片 p ,如果满足式(7),则将该面片过滤掉.

$$|V^*(p)| (1 - g^*(p)) < \sum_{p_i \in U(p)} 1 - g^*(p) \quad (7)$$

从公式可以看出,如果面片 p 位于实际曲面外部,那么 $1 - g^*(p)$ 和 $|V^*(p)|$ 都会比较小,这样 p 一般都会被过滤掉.

通过更严格的可见一致性实现第二个过滤器. 将图像单元中与深度最小的 patch 距离大、法向量夹角大的 patch 过滤. 对于图像单元中的每个面片 p ,通过计算深度数据得到其可见图像的总数,如果这个总数小于给定阈值则将其过滤掉.

最后,通过引入弱规则来实现第三个过滤器,对于每个面片 p ,找出其对应的图像 $V(p)$ 中映射到面片 p 所在的图像块以及其相邻的图像块中的面片集,若 p 邻域内的面片数量与收集到的面片数量之比小于 0.25,则过滤掉 p .

面片扩展和过滤步骤重复多次,以使面片密集化并删除错误的面片. 作物果实密集点云重建算法的伪代码由 Algorithm 2 给出.

Algorithm 2: Dense point cloud reconstruction algorithm

Input: Collection of image sequences C
Input: Camera internal and external parameters K_i, R_i, T_i
Output: Dense point cloud of target fruit $P = \{X_i\}$

- 1 for Image $I \in C$ do
- 2 Extract the feature points of the image to the set $F(I)$, divide image blocks $C_i(x, y)$
- 3 end
- 4 for Image $I \in C$ do
- 5 for Feature point $f \in F(I)$ do
- 6 Find the corresponding feature point f_{nn} in $F(J)$: $f_{nn} = \underset{f' \in F(J)}{\operatorname{argmin}} \|f_d - f'_d\|_2$, image $J \in C/I$
- 7 end
- 8 Feature point set $F \leftarrow \{f, f_{nn}, \dots\}$
- 9 for Feature point $f' \in F/f$ do
- 10 Patch center point $c(p) \leftarrow \text{triangulate } f \text{ and } f', p \text{ is the candidate patch}$
- 11 $\text{dis} \leftarrow \text{distance from optical center } O(I) \text{ of image } I \text{ to } c(p)$
- 12 end
- 13 sort dis
- 14 for dis increase do
- 15 Construct a candidate patch p , initialize its $c(p), n(p), R(p)$
- 16 Central point $c(p) \leftarrow \text{triangulate } f \text{ and } f'$
- 17 Unit normal vector $n(p) \leftarrow \frac{\overrightarrow{c(p)O(I)} \times \overrightarrow{c(p)O(I)}}{\|\overrightarrow{c(p)O(I)}\| \times \|\overrightarrow{c(p)O(I)}\|}$
- 18 Reference image $R(p) \leftarrow I_i$
- 19 The image set $V(p)$ which can observe the patch p : $V(p) \leftarrow \{I | n(p) \cdot \overrightarrow{c(p)O(I)} / \|\overrightarrow{c(p)O(I)}\| > \cos(\theta)\}$
- 20 Remove the image that does not satisfy the grayscale consistency function and get $V^*(p): V^*(p) \leftarrow \{I | I \in V(p), h(p, I, R(p)) \leq \alpha\}$
- 21 Optimize $c(p), n(p)$, recalculate $V(p)$ and $V^*(p)$
- 22 if The visual image set of patch p satisfies: $|V^*(p)| \geq \gamma$ then
- 23 Accept the reconstruction result of patch p and store $p: Q_i(x, y) \leftarrow p$ exit
- 24 end
- 25 end
- 26 end
- 27 repeat
- 28 for Image $I \in C$ do
- 29 for Each image block $C_i(x, y) \in I$ do
- 30 for Each patch $p \in \text{patch set } Q_i(x, y) \text{ projected on image block } C_i(x, y)$ do
- 31 Get the image block neighborhood $C(p)$ of patch p : $C(p) \leftarrow \{C_i(x', y') | p \in Q_i(x, y), |x - x'| + |y - y'| = 1\}$
- 32 for Each image block $C_i(x', y') \in C(p)$ do
- 33 Expand to get a new patch p'
- 34 Initialize the relevant values of patch $p': n(p'), R(p'), V(p') \leftarrow n(p), R(p), V(p)$
- 35 Optimize $n(p'), R(p'), V(p')$
- 36 Recalculate $V(p')$ and $V^*(p')$
- 37 if The visual image set of patch p' satisfies: $V^*(p') \geq \gamma$ then
- 38 p' is a renaoie expansion paccn and store it: $Q_i(x, y) \leftarrow p'$
- 39 exit
- 40 end
- 41 end
- 42 end
- 43 end
- 44 end
- 45 for Image $I \in C$ do

```

46 for Each image block  $C_i(x,y) \in I$  do
47 for Each patch  $p \in Q_i(x,y)$  do
48  $U(p)$  is the set of patches whose visibility is inconsistent with that of patch  $p$ 
49 if Patch  $p$  satisfies the following condition:  $V^*(p) \mid (1-g^*(p)) < \sum_{p_i \in U(p)} (1-g^*(p_i))$  then
50 delete  $p$ , patch filter 1
51 else if Number of visual images  $\{V^*(p)\} < \gamma$  then
52 delete  $p$ , patch filter 2
53 else if Number of  $p$  neighbor patches of  $C(p)/total$  number of patches in  $C(p) < 0.25$  then
54 delete  $p$ , patch filter 3
55 end
56 end
57 end
58 end
59 until Each image block  $C_i(x,y)$  reconstructs a patch
    
```

3 点云去噪

在拍摄目标果实时,由于受背景和拍摄环境的影响,拍摄结果图像中不可避免地含有除果实以外的枝杈、叶子等背景噪声.进而在采用 SFM 算法生成点云时,会生成排列紧凑,数量较多的噪点,使其无法通过现有的去噪算法删除.针对该问题,本文提出了一种交互式去噪和滤波相结合的半自动化去噪方法.

3.1 交互式去噪

交互式去噪算法通过交互式界面将点云选择与点云删除相结合实现去噪.在点云选择过程中,本研究面向采用框选的点集选择算法,作为一种主要的点云选择策略.故可将点云选择可看作是三维点云的拾取.由于三维点云的点是一种没有大小且不具备面状特征的图形对象,所以设计了对点集的拾取,而非对单个点的拾取.本研究所采用的点云框选算法是面向多点云数据的拾取算法,通过将点云模型映射到点云可视化的查看器中,使用鼠标确定一个矩形,检测各点的位置坐标与此矩形的边界相交关系,以判断其是否位于矩形内部,即为拾取点.

在点云框选去噪中,首先利用鼠标在场景模型中交互式确定一个矩形,再计算此矩形的大小与空间位置并得到一个特殊的视景体,然后基于此视景体对各点进行相交检测,与视景体有交集的点即为被选择的点云.进而获取视景体中被选择的点云数据索引,根据索引获得确定的点云数据,从而实现点云选择.为避免用户操作失误,在框选设计中,不同视角的点云数据可以被重复选择.通过将被框选的点颜色设置为红色,以标记该点的信息已发生变化,同时通过迭代点云模型将框选后的点云图像与原点云的图像进行增量比较,增量信息改变的点云不再保留与绘制,从而实现对选中的点云进行删除.

3.2 滤波去噪

在生成果实点云模型时由于误差通常会产生密度不均匀的点云数据集以及稀疏的离群点,这样会使生成果实点云模型效果大打折扣,也会对后续点云模型的表面重建造成影响.据此本文主要基于 K 邻近点距离统计算法设计滤波器,进行该类离群点的去除.所需优化的参数主要包括:邻域内需计算的点个数 N ;标准范围与标准差的关系,即离群点阈值 k 等.去噪过程中,首先随机选择一个中心点 s ,并计算出该点到 N 个邻近点的欧氏距离 $d_i (i=1,2,\dots,N)$,随后求取所有邻近点距离的均值 μ 和标准方差 σ ,如式(8)、式(9).

$$\mu = \left(\frac{1}{N} \right) \sum_{i=1}^N d_i \quad (8)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (d_i - \mu)^2} \quad (9)$$

以阈值 k 作为判断条件,判断范围内某一临近点 s' 到 s 的距离 d' (10) 是否满足条件(11);如果满足条件,则保留该点,继续计算下一个临近点.

$$d' = \sqrt{(x'-x)^2 + (y'-y)^2 + (z'-z)^2} \quad (10)$$

$$d' \leqslant k \cdot \sigma^2$$

(11)

如果不满足,即该点到中心点的距离超出 k 个标准差以上,那么该点将被标记为离群点,同时在点云模型中移除。

4 实验结果及结论

本文研究了作物果实表型的稀疏点云获取、密集点云生成、噪点多模式去除等方法,并开发一套对导入图片、特征提取、特征匹配、建立稀疏点云模型、建立密集点云模型、格式转换、查看 pcd 点云文件、交互式去噪、滤波去噪等操作进行管理和验证的系统,如图 2 所示。用户可选择拍照方式或者通过三维设计工具获取目标果实作物在不同视角且变化幅度 10 度左右的二维图片素材组,导入开发的程序;对于输入的二维图像组,进行信息提取与图片匹配,根据匹配结果,得到关键点、相机位置等场景信息进行稀疏点云重建;然后在稀疏点云的基础上进行稠密的点重建;进一步的,可以进行点云文件 ply 至 pcd 格式转换和 pcd 格式与 txt 格式的转换,并实现 pcd 点云文件加载、显示和编辑;之后可以通过交互式选择和基于滤波器的方法进行离群点去除。结果生成的三维点云模型有较强的完整性和真实性。具体测试过程如下:

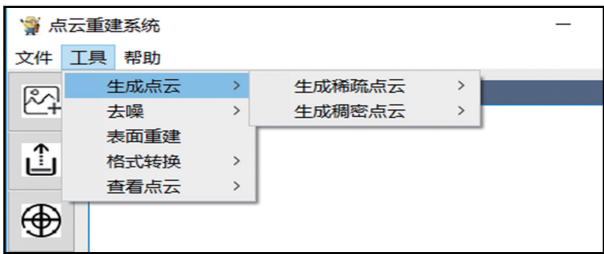


图 2 系统界面
Fig. 2 System interface



图 3 部分苹果序列图像
Fig. 3 Partial apple sequence images

(1)以复杂实验背景中的果实为研究对象,以 Iphone 6s 数码相机作为拍照设备,拍摄时的镜头焦距固定为 5 mm,采用自动曝光模式,图像分辨率 4 032×3 024 像素。围绕果实每转 10°~20°拍 1 张,确保形成连续的果实图像序列。本研究采集了 25 张苹果序列图像,如图 3 所示。利用数据文件提取模块和图片特征匹配模块,对导入的多幅作物果实二维图像进行数据和特征提取,并使用 FLANN 算法对不同维度的数据进行匹配;生成果实的稀疏点云模型,根据匹配结果,基于 SFM 算法利用射影定理计算得到相机位置等场景信息进行稀疏重建(sparse reconstruction)。其中,以苹果为目标物,得到稀疏点云模型如图 4 所示。

(2)生成果实的密集点云模型。通过 PMVS 算法,利用所得到的场景信息与原始照片,将面片扩展与筛选步骤重复三次,从而扩散已有数据点进行密集重建(dense reconstruction)。其中,在进行面片扩展时,本研究采用的阈值为 3。以苹果为目标物,密集点云重建过程如图 5 所示。其中,图 5 左侧为密集点云重建过程的运行状态信息,右侧为苹果稀疏点云模型文件。

(3)提供点云文件格式转换。实现了 ply 与 pcd 格式、pcd 与 txt 格式的转换、pcd 点云文件加载、显示与编辑,使不同平台上果实点云无缝对接。图 6 为 pcd 文件格式存储的苹果稠密点云模型。

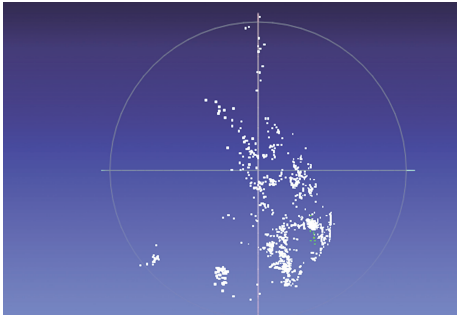


图 4 苹果稀疏点云模型
Fig. 4 Apple sparse point cloud model

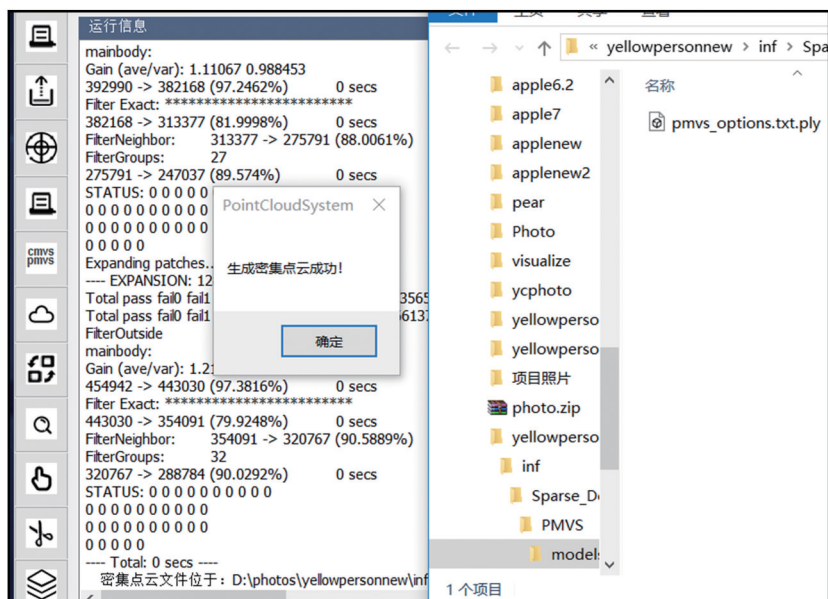


图 5 密集点云重建过程

Fig. 5 Dense point cloud reconstruction process

(4)果实密集点云模型的去噪预处理. 对于一些滤波去噪无法去除的明显大尺度背景可手动去除,该方法将点云框选与点云删除进行关联,支持用户通过鼠标手动选中噪点然后进行删除,该方法能够实现连续框选,并删除被框选的点云,同时也能够保存删除噪点之后的点云图像. 对图 6 所示的苹果稠密点云模型进行交互式去噪,过程如图 7 所示. 图 7 左侧为选择的去噪区域,右侧为交互式去噪后的苹果点云模型.

(5)果实密集点云模型的自动去噪处理. 通过统计滤波器处理离群点(本研究取 $N=1\ 000$, $k=1.0$),避免由这些点引起的局部结构复杂化,从而保证三维点云模型的完整性和真实性,如图 8 所示为滤波去噪后的果实点云模型.

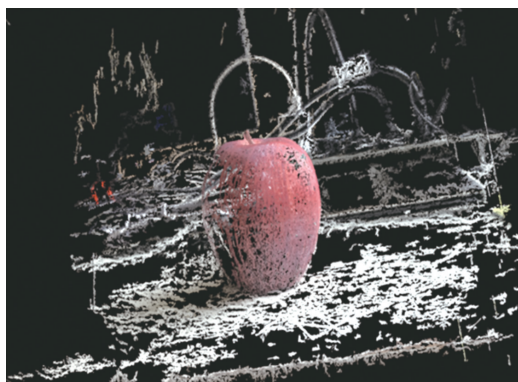


图 6 pcd 格式密集点云模型

Fig. 6 Dense point cloud model in pcd format



图 7 苹果点云交互式去噪

Fig. 7 Apple point cloud interactive denoising



图 8 苹果点云滤波去噪

Fig. 8 Apple point cloud filtering denoising

(6)算法泛化能力测试.

本研究还以榴莲等其他果实为对象,进行了复杂背景中作物果实的三维重建测试. 其中,采集的 35 张榴莲序列图像如图 9 所示. 使用本文方法处理榴莲序列图像,得到的稀疏点云模型如图 10 所示,最终得到的稠密重建效果如图 11 所示. 结果显示,本文的方法对复杂背景中不同作物果实进行处理时,具有较

好的鲁棒性. 各阶段的测试数据如表 1.

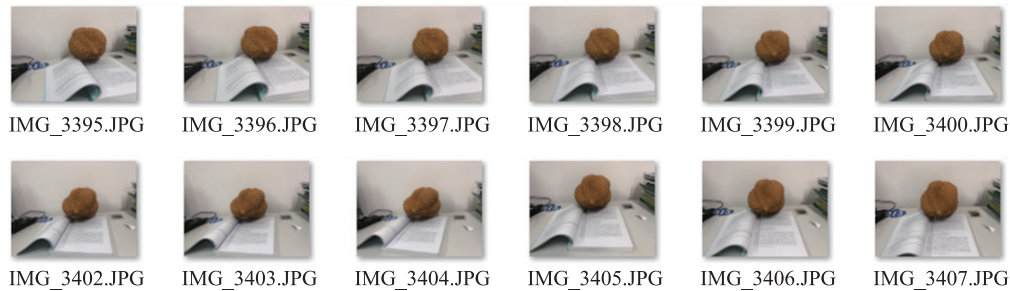


图 9 部分榴莲序列图像
Fig. 9 Partial durian sequence images



图 10 榴莲稀疏点云模型
Fig. 10 Durian sparse point cloud model

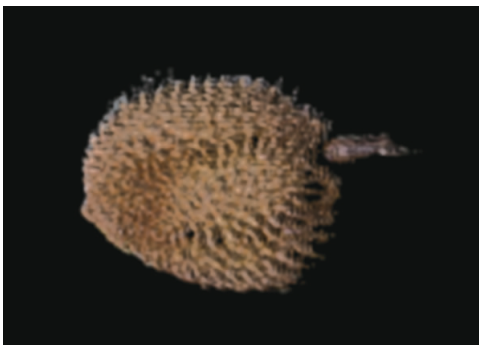


图 11 榴莲稠密重建效果
Fig. 11 Dense reconstruction effect of durian

为验证本文所提出方法的通用性,拍摄了一组室外建筑序列图像作为研究对象,如图 12 所示. 经过处理得到的图书馆稀疏点云模型如图 13 所示,稠密重建效果如图 14 所示. 结果证明,本文提出的方法在基于多视角图像的三维重建方面具有一定的普适性.

表 1 测试数据
Table 1 Test data

果实类型	图像数量	稀疏点数	稠密点数	去噪点数
苹果	25	10 808	3 743 041	1 438 080
榴莲	35	13 585	4 655 093	1 833 733



图 12 图书馆序列图像
Fig. 12 Library sequence images

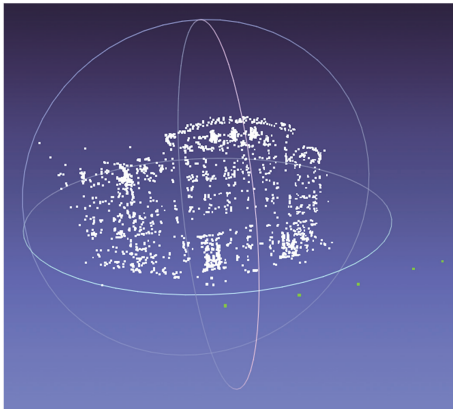


图 13 图书馆稀疏点云模型
Fig. 13 Library sparse point cloud model

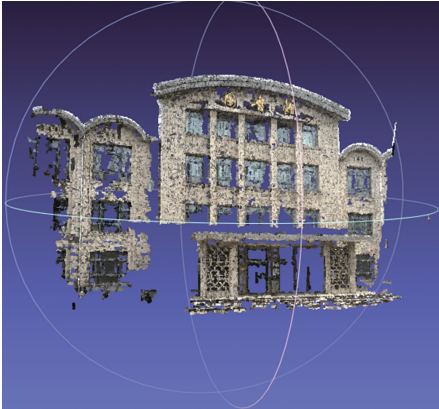


图 14 图书馆稠密重建效果
Fig. 14 The dense reconstruction effect of the library

[参考文献]

- [1] 王孟博. 基于点云的作物植株三维重建技术研究[J]. 鄂州大学学报, 2018, 25(1): 104–106, 109.
- [2] 史维, 张吴平, 郝雅洁, 等. 基于视觉三维重建的作物表型分析[J]. 湖北农业科学, 2019(16): 125–128.
- [3] 周静静, 郭新宇, 吴升, 等. 基于多视角图像的植物三维重建研究进展[J]. 中国农业科技导报, 2019, 21(2): 9–18.
- [4] 黄雄. 三维重建中点云数据配准算法的研究[D]. 秦皇岛: 燕山大学, 2016.
- [5] 贾鹤鸣, 孟羽佳, 邢致恺, 等. 基于点云拼接的植物三维模型重建[J]. 应用科技, 2019, 46(1): 23–28.
- [6] 杨会君. 基于点云的果实三维重建技术研究[D]. 杨凌: 西北农林科技大学, 2014.
- [7] 娄吕. 三维激光扫描点云数据精简算法研究[D]. 昆明: 昆明理工大学, 2017.
- [8] 皮志荣. 地面三维激光扫描技术在工程测量中的实践[J]. 江西建材, 2017(24): 225, 230.
- [9] HENRY P, KRAININ M, HERBST E, et al. RGB-D mapping: using depth cameras for dense 3D modeling of indoor environments[C]. Berlin, Heidelberg: Springer, 2010.
- [10] 马银中. 基于 Kinect 的三维场景实时重建及相关技术研究[D]. 合肥: 安徽大学, 2018.
- [11] 任谦. 基于 Kinect 传感器的三维稠密地图构建系统设计[J]. 电工技术, 2019(6): 144–146, 149.
- [12] 杨红庄. 全自动深度相机三维扫描系统[D]. 合肥: 中国科学技术大学, 2016.
- [13] 余秀丽, 王丹丹, 牛磊磊, 等. Kinect 在现代农业信息领域中的应用与研究进展[J]. 农机化研究, 2015(11): 216–221.
- [14] 董鹏辉, 柯良军. 基于图像的三维重建技术综述[J]. 无线电通信技术, 2019, 45(2): 4–8.
- [15] XU B, YANG L F, LEI B, et al. A study of the accuracy improvement strategy for multi-view 3D reconstruction instrument based on automatic calibration technology[J]. Acta geoscientica sinica, 2017, 38(2): 243–248.
- [16] RUPNIK E, PIERROT D M, DELORME A. 3D reconstruction from multi-view VHR-satellite images in MicMac[J]. ISPRS journal of photogrammetry and remote sensing, 2018, 139: 201–211.
- [17] 周静静, 郭新宇, 吴升, 等. 基于多视角图像的植物三维重建研究进展[J]. 中国农业科技导报, 2019, 21(2): 9–18.
- [18] GANG C, BIN C, YUXIN L, et al. 3D virtual plant photosynthesis simulation model based on L-system[J]. Transactions of the chinese society for agricultural machinery, 2018, 49(10): 275–283.
- [19] 全思博, 李伟光, 郑少华. 基于 RGB-D 传感器的车间三维场景建模[J]. 华南理工大学学报(自然科学版), 2015, 43(6): 42–47.

[责任编辑: 陆炳新]