

# 采摘机器人全果园视觉感知及自主作业综述

陈明猷, 罗陆锋\*, 刘威, 韦慧玲, 王金海, 卢清华, 骆少明

(佛山大学机电工程与自动化学院, 广东佛山 528231, 中国)

**摘要:** [目的/意义] 采摘机器人是智慧农业的重要组成部分, 其感知、规划、控制相关基础方法理论目前已有系统化研究。然而, 构建具备全果园“感知-移动-采摘”一体化作业能力的实用型采摘系统仍面临诸多挑战。针对该问题, 本文调研并报道了本领域近期案例, 将全果园自主作业的关键技术划分为局部目标感知、全局地图构建和自主作业行为规划三个子问题并进行综述。[进展] 首先回顾了近距离、局部范围内水果目标的精细视觉感知方法, 包括基于低级特征融合、高级特征学习、RGB-D信息融合, 以及多视角信息融合的4种方法; 介绍与分析了全局尺度下的果园地图构建与大规模场景视觉感知案例; 在感知的基础上, 调研分析采摘机器人自主作业行为规划方法, 包括底盘移动路径规划、机械臂视点规划与避障路径规划等方面的最新研究; 最后对采摘机器人自主作业系统构建案例进行报道与分析。[结论/展望] 感知、移动、采摘模块的高效协同是实现采摘机器人从基础功能样机进一步迈向实用型机器的关键, 已有的视觉感知、规划与控制算法的鲁棒性与稳定性均需增强, 协同程度需进一步提高。此外, 提及了采摘机器人应用的几个开放性研究问题, 并描述了其未来发展趋势。

**关键词:** 采摘机器人; 自主作业; 局部感知; 全局建图; 行为规划

中图分类号: TP391.41

文献标志码: A

文章编号: SA202405022

引用格式: 陈明猷, 罗陆锋, 刘威, 韦慧玲, 王金海, 卢清华, 骆少明. 采摘机器人全果园视觉感知及自主作业综述[J]. 智慧农业(中英文), 2024, 6(5): 20-39. DOI: 10.12133/j.smartag.SA202405022

CHEN Mingyou, LUO Lufeng, LIU Wei, WEI Huiling, WANG Jinhai, LU Qinghua, LUO Shaoming. Orchard-Wide Visual Perception and Autonomous Operation of Fruit Picking Robots: A Review[J]. Smart Agriculture, 2024, 6(5): 20-39. DOI: 10.12133/j.smartag.SA202405022 (in Chinese with English abstract)

## 0 引言

采摘机器人具有机械化、自动化和智能化的特点, 是智能农机装备和智慧农业系统的重要组成部分, 也是实现果园无人化采摘的关键。

构建接近人类作业能力的强自主型水果收获系统, 实现在全局非结构化果园下的“感知-移动-采摘”一体化无人作业, 是采摘机器人领域长期目标之一。近年来, 机器视觉与人工智能技术取得长足发展并成功应用于水果采摘领域<sup>[1,2]</sup>, 采摘机器人在手眼协同<sup>[3,4]</sup>、水果三维重构<sup>[5-7]</sup>、果园道路识别与导航<sup>[8]</sup>、避障路径规划<sup>[9,10]</sup>、末端抓取<sup>[11,12]</sup>等方面均取得显著进展。经过长期迭代与优化, 采摘机器人的视觉系统性能已有显著提高, 与早期采摘机器人相比能更好地适应果园中的混乱背景<sup>[13-17]</sup>、低

色彩对比度和动态光照<sup>[18]</sup>等复杂因素, 并有效响应细小目标<sup>[19]</sup>、遮挡目标和表型复杂目标<sup>[20-25]</sup>。

尽管如此, 采摘机器人在真实果园中的应用远未成熟<sup>[26]</sup>, 这主要由两方面因素决定: 一方面, 采摘机器人感知系统与人类相比仍有较大差距, 无法完全适应动态多变的非结构化果园环境, 导致采摘成功率偏低(平均66%), 耗时较长(平均33 s)<sup>[27]</sup>。另一方面, 已有研究多聚焦于单一问题, 如水果目标识别、作物产量估计、道路识别与导航、采摘路径规划等, 而对于如何协同上述多个基础功能模块, 在应用层面构建性能完备的自主作业系统的研究仍然缺乏。实际上, 具备类人水平的“感知-移动-采摘”一体化全果园作业能力, 是采摘机器人从基础功能样机进一步迈向智能与自主的关键, 相

收稿日期: 2024-05-28

基金项目: 国家自然科学基金项目(32301704, 32171909); 广东省自然科学基金项目(2024A1515010199, 2023A1515011255)

作者简介: 陈明猷, 研究方向为智能采摘机器人技术研究。E-mail: chenmingyou1@126.com

\*通信作者: 罗陆锋, 博士, 教授, 研究方向为智能采摘机器人技术研究。E-mail: luolufeng617@163.com

copyright©2024 by the authors

应的“眼-身-手”高效行为协同是制约采摘机器人实用的瓶颈因素。

目前为止，无论在基础理论还是工程实现层面，构建类人自主作业能力的采摘机器人都是一个复杂的系统问题，对算法、硬件、策略和系统架构设计提出了严苛要求。为了实现上述目标，采摘机器人需要具备三个关键能力：近距离感知水果与障碍物空间位姿信息、全方位感知整个果园场景的结构与几何信息、基于感知进行自主和高效的作业行为决策。

围绕上述三个关键问题，本文系统调研了采摘机器人的近期相关研究成果。首先，回顾了果园局部区域内的水果、障碍物精细视觉感知方法。然

后，介绍了全局尺度下的视觉感知与果园地图构建方法。接下来，调研了基于视觉的自主移动及采摘作业行为规划方法。最后，在上述关键方法调研的基础上，进一步对采摘系统的构建及其部署案例进行报道，分析已有采摘系统在“眼-身-手”协同应用层面的效能与局限性，并给出技术性结论与建议。此外，提出了与采摘机器人应用领域相关的几个开放性研究问题并分析了其未来发展趋势，为构建高度智能、自主与实用的采摘机器人，以及实现其大规模应用提供新的理论与技术参考。采摘机器人自主作业系统及其关键技术的内容组织框架如图1所示。

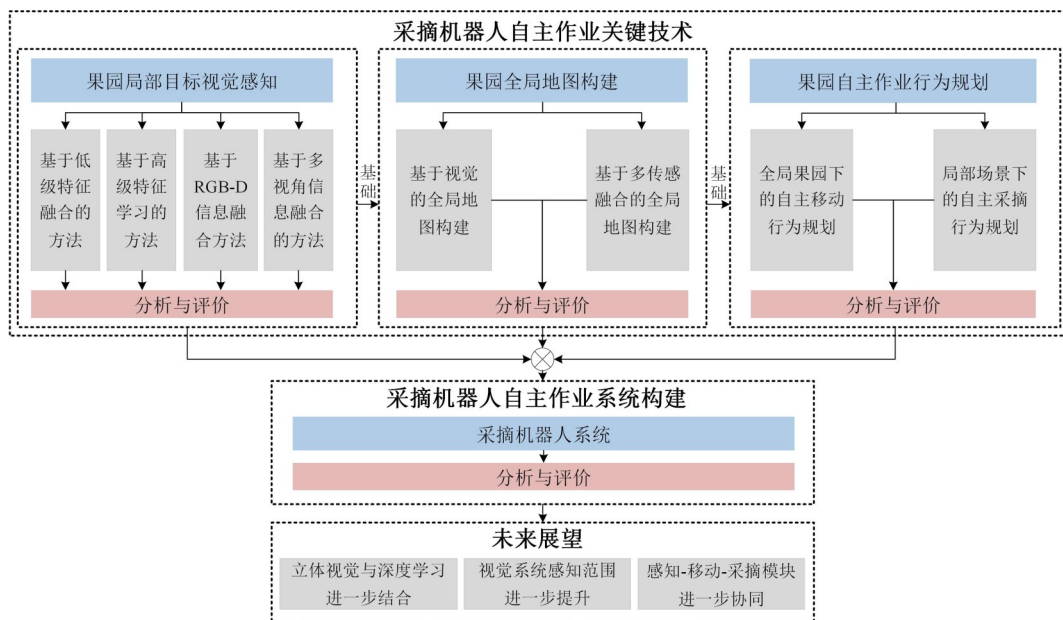


图1 采摘机器人自主作业系统及其关键技术内容组织框架图

Fig. 1 Organizational framework of autonomous operation systems and key technologies for fruit picking robot

## 1 果园局部目标视觉感知

局部目标视觉感知特指机器人处于果园中特定位点时，通过视觉方法感知其近距离视野内的精细目标信息。局部目标视觉感知是机器人实现后续全果园建图和自主作业行为规划的基础。

针对果园内颜色混杂、目标遮挡和光照变化等诸多复杂因素<sup>[28,29]</sup>，早期常见基于阈值分割、形状分类、颜色或传统机器学习等方法，其缺点是依赖人工特征和阈值，无法满足机器人对感知稳定性、灵活性和通用性的需求。随着研究的深入，目前已被证明有效的局部视觉感知方法主要有：基于低级特征融合、基于高级特征学习、基于RGB-D信息融合，以及基于多视角信息融合的方法。

### 1.1 基于低级特征融合的方法

图像的低级特征多表现为直观的物理属性，其优势是易于理解且包含丰富的细节信息。水果采摘环境中常见的低级图像特征有色相、饱和度、形状与法向量等<sup>[30,31]</sup>。基于多种低级特征融合的视觉算法在户外非线性光<sup>[32]</sup>和局部遮挡<sup>[33,34]</sup>等复杂情形下的性能可显著优于基于单一特征的方法。

水果植株的生长对称性是另一重要的低级特征，在采样不佳的情况下可辅助立体视觉系统复原完整的3D目标信息<sup>[35]</sup>。图像形态学与纹理特征则描述了目标的整体和内部拓扑结构，可为低级特征融合算法提供重要几何学约束，提高感知精度。此

外,果实与果梗的相对位姿也是采摘任务中的重要低级特征,可辅助视觉系统从复杂背景与混叠果实场景中优化末端采摘位姿<sup>[36]</sup>。

由于低级特征具有较强直观性,其提取过程一般在数学上可解析,在经典机器学习框架下将多种特征量化后作为输入可实现高效的特征融合,构建轻量、灵活且具备高度可解释性的局部视觉感知算法。基于低级特征融合方法的局部目标视觉感知效果如图2所示。

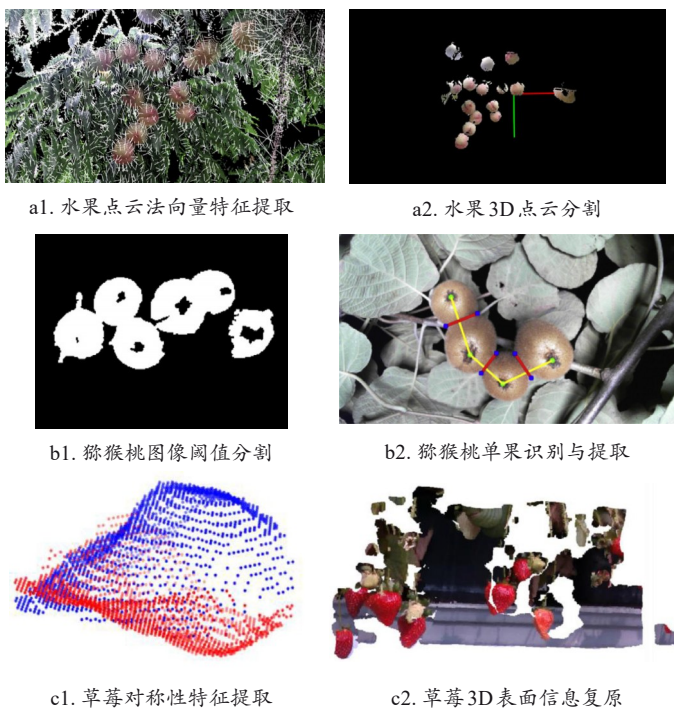


图2 基于低级特征融合方法的局部目标视觉感知效果示意图

Fig. 2 Schematic diagram of the effects of local target visual perception based on low-level feature fusion method

然而与早期基于单一特征的方法类似,基于低级特征融合的方法中,特征选择的过程仍依赖于开发者的经验与主观判断,因此其效果存在不稳定性。同时,有限个低级特征的组合可能无法充分覆盖特定任务所需的特征分布。因此,该方法虽然一定程度上有效,但仍难以完全满足强自主型采摘机器人对感知稳定性与可靠性的严苛要求。

## 1.2 基于高级特征学习的方法

高级特征指语义、状态、类别和属性等抽象特征,与低级特征相比更接近人的感知形式。深度神经网络是提取高级特征的重要技术手段。深度神经网络可实现端对端感知,即以原始图像作为输入,通过多个特征提取层自动获取目标抽象特征,最终

直接输出所需水果和障碍物高级特征信息的过程,其可避免复杂的人工特征设计。因此,深度神经网络对复杂多变的果园环境具有较强适配性,是未来构建强鲁棒、强智能视觉感知系统的重要技术。近年来,高性能并行计算硬件的飞速发展大大推动了大型深度神经网络在采摘机器人上的部署与应用,显著提高了采摘机器人的综合感知能力,在水果识别、空间定位和复杂表型重构等方面取得了喜人的成果。

针对水果目标图像识别问题,基于高级特征的深度学习算法比色差法、颜色空间法和聚类等传统方法具有更强的鲁棒性<sup>[37]</sup>和场景尺度自适应性<sup>[38]</sup>,对枝叶混叠和动态光照等复杂因素的响应良好<sup>[39]</sup>,识别成功率和效率总体高于传统方法。受益于神经网络的端对端特性,基于高级特征学习的方法可有效简化水果识别算法的设计逻辑与实现流程<sup>[40]</sup>,提高算法部署效率。此外,基于水果自身的生长特性,可针对性地优化深度学习网络结构与构造新型损失函数,通过注意力机制进一步提高识别算法的精度、实时性和降低计算复杂度<sup>[41,42]</sup>。

针对水果采摘位姿计算问题,部署端对端深度检测网络是提取安全采摘区域,避免抓取破损<sup>[43]</sup>或刚性碰撞<sup>[44]</sup>的有效手段。通过语义分割、实例分割等像素级处理模型提取复杂背景下采摘点的位置及序列信息,可辅助机器人完成精准采摘和高效序列规划<sup>[45,46]</sup>。在高级语义的基础上计算果实和果茎间的生长距离、角度和偏移等几何信息,可为采摘末端的结构优化设计提供重要参考<sup>[47]</sup>。

针对水果目标三维重建问题,以高级特征取代低级几何特征来表征复杂果实表型,可显著提高点云配准与重构算法的效率和稳定性<sup>[48]</sup>。对于表型信息缺失或不完整的水果点云,通过深度生成对抗网络可将信息残缺的表面复原为封闭和完整的曲面,提高重构的全面性和精度<sup>[21]</sup>。

深度神经网络的训练依赖于大量样本。对于复杂农业环境,优质训练集应具备广泛分布的视角、光照、尺度、果形或遮挡等条件信息<sup>[49]</sup>,并可通过图像、点云、体素、时间序列等多模态形式呈现。为确保网络性能满足果园高级特征提取需求,除了评估网络精度和实时性等基础性能指标外,还需在真实环境下进行系统与广泛地部署与测试,对稳定性和场景适用性等实际问题进行考量<sup>[50]</sup>。

基于高级特征学习方法的果园局部目标视觉感知效果如图3所示,具体研究情况见表1。



表1 近年基于高级特征学习的方法在果园局部目标视觉感知中的研究

Table 1 Recent researches on high-level feature-based methods used in local target visual perception of orchard

编号	作者	年份	所部署深度神经网络	特征提取方式	针对任务
1	Olatunji 等 <sup>[20]</sup>	2019	CGAN	主动生成	水果目标三维重建
2	Xie 等 <sup>[37]</sup>	2019	YOLOV3	目标检测	水果目标图像识别
3	Qiu 等 <sup>[38]</sup>	2022	YOLOV4	目标检测	水果目标图像识别
4	肖张娜等 <sup>[39]</sup>	2023	CBAM+YOLOV4	目标检测	水果目标图像识别
5	Majeed 等 <sup>[40]</sup>	2018	SegNet	语义分割	水果目标图像识别
6	Peng 等 <sup>[41]</sup>	2024	RetinaNet	目标检测	水果目标图像识别
7	Tao 等 <sup>[42]</sup>	2024	BiFormer+YOLOv5s	目标检测	水果目标图像识别
8	Guo 等 <sup>[43]</sup>	2024	MDETR	夹持检测	水果目标夹持检测
9	Ge 等 <sup>[44]</sup>	2019	Mask-RCNN	实例分割	水果采摘位姿计算
10	Zhang 等 <sup>[45]</sup>	2020	Mask-RCNN	实例分割	水果采摘位姿计算
11	Fang 等 <sup>[46]</sup>	2024	AZNet	语义分割	水果目标图像识别
12	Barth 等 <sup>[47]</sup>	2019	DeepLabV2	语义分割	水果采摘位姿计算
13	Gené 等 <sup>[48]</sup>	2020	Mask R-CNN	实例分割	水果目标三维重建
14	Altaheri 等 <sup>[49]</sup>	2019	不限于特定类型网络	不限于特定提取方式	水果及其环境数据集构建
15	Salim 等 <sup>[50]</sup>	2023	DenseNet-201, Xception, MobileNetV3-Small, ResNet-50	目标分类	多种预训练模型性能评估

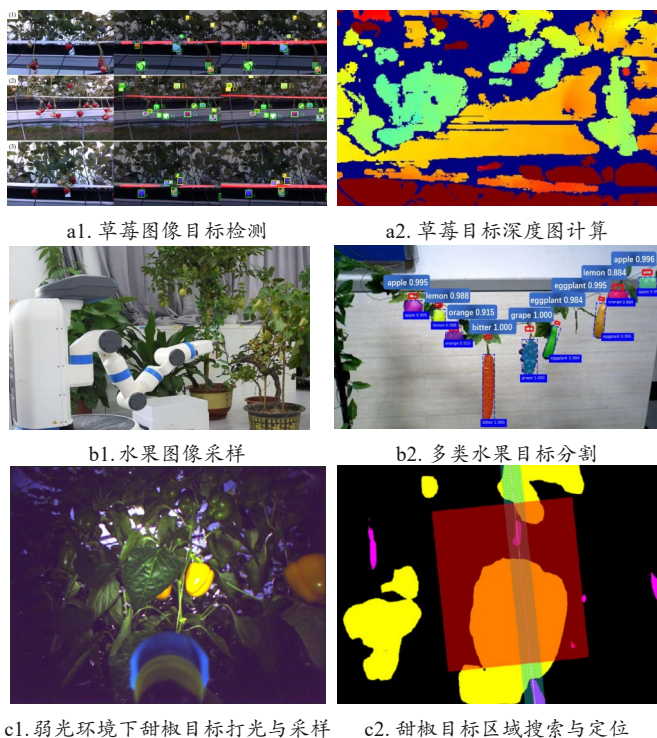


图3 基于高级特征学习方法的局部目标视觉感知效果示意图  
Fig. 3 Schematic diagram of the effects of local target visual perception based on high-level feature learning method

已有研究案例中，深度神经网络所提取的高级图像特征可有效适应非结构化果园中的枝叶遮挡、动态光照、颜色混叠和自然扰动等复杂因素，具有较强的鲁棒性、稳定性和通用性。大规模深度神经

网络的引进有效解决了农业视觉感知领域中的诸多瓶颈问题，在采摘机器人领域展现出了较强的应用潜力。随着鲁棒、稳定和通用的高级特征提取机制日益成熟，构建高度智能和自主的感知系统也逐渐成为可能<sup>[51]</sup>。

然而需要指出的是，高级特征的提取需依赖大量训练数据以及充足硬件算力支持，而户外果园的采样条件和硬件部署条件均受限明显。为了解决该问题，需进一步构造轻量化网络，需降低网络模型对训练样本和算力的要求。此外，已有的样本扩充技术需进一步优化，提高特征学习和网络参数化建模效率，实现户外环境下基于小样本的高质量深度神经网络训练。

### 1.3 基于RGB-D信息融合的方法

RGB-D信息通过深度相机捕获，是原始彩色图像信息与各像素对应的深度信息的合称。深度相机借助结构光、激光或红外传感器等辅助测量器件计算深度值，进而基于相机模型计算目标的三维位置信息，目前被广泛应用于果园感知任务中。

单目相机和双目相机等传统光学相机的成像过程基于被动视觉框架，通过捕获场景中已有的自然光或射线进行目标采样和图像处理。相比之下，深度相机多基于主动视觉框架，通过发射激光、红外线、结构光等光源来获取环境信息，有效避免了复杂的多视图立体匹配过程，大大提高了视觉测量效

率,对具有低像素占比的细小果梗识别尤为适用<sup>[19]</sup>。同时,深度信息可为颜色信息提供重要辅助特征,在果实与背景颜色对比度较低的情形下提高视觉识别算法的稳定性<sup>[52]</sup>。除了与颜色特征融合外,深度特征也可与形状、轮廓、法线等低级特征融合,实现对复杂场景目标的多维度特征描述,构建强鲁棒视觉感知算法<sup>[53,54]</sup>。基于RGB-D信息融合方法的局部目标视觉感知效果如图4所示。

然而,目前常见的深度相机对复杂果园采样环境的适应性欠佳。例如,基于结构光的KinectV1、

Xtion PRO Live、RealSense SR300等相机的投射光斑信息受实时环境扰动的影响较大,测量精度受制于外部光照条件;基于飞行时间(Time of Flight, ToF)法的KinectV2相机对时间测量模块的精度要求较高,在计算相位偏移时需要消耗大量的计算资源。同时,该相机需要外接电源,体积较大,能耗较高,在持续作业工况下的耐用性和实用性欠佳。基于主动红外立体成像(Active Infrared Stereo, AIRS)原理的RealSense D435i, R200等型号相机在投射信号的计算和优化上需要耗费大量计算资源,实时性较低<sup>[28]</sup>。目前市面上常见消费级深度相机的优缺点如表2所示。

除了上述针对特定类型相机的局限性外,深度相机采集RGB-D信息过程也存在共性问题<sup>[55,56]</sup>。例如,深度相机的感知质量显著受限于深度传感器的性能,其理论有效感知距离以及视场角总体上小于被动式双目相机。同时,深度相机依赖精密的深度测量模块,总体成本往往高于被动式视觉系统。此外,深度相机的成像过程与测量原理相对复杂,使用者往往不具备对深度信息生成过程的控制权限,导致视觉软件系统设计与集成的灵活性欠佳。

综上所述,虽然RGB-D信息融合方法在三维信息获取方面具有一定优势,但受限于深度相机自身的传感机制和硬件特点,其在真实采摘环境下的稳定性与适用性仍有待提高。

#### 1.4 基于多视角信息融合的方法

该方法通过部署多组相机或通过单一相机从不同角度采样,实现对目标表面完整三维信息的获取,具有感知范围广和测量细节丰富的特点,对果园中的枝叶遮挡、多果混叠与光照不均等因素具有较强的适应性。

相比于基于单视角采样的方法,基于多视图几

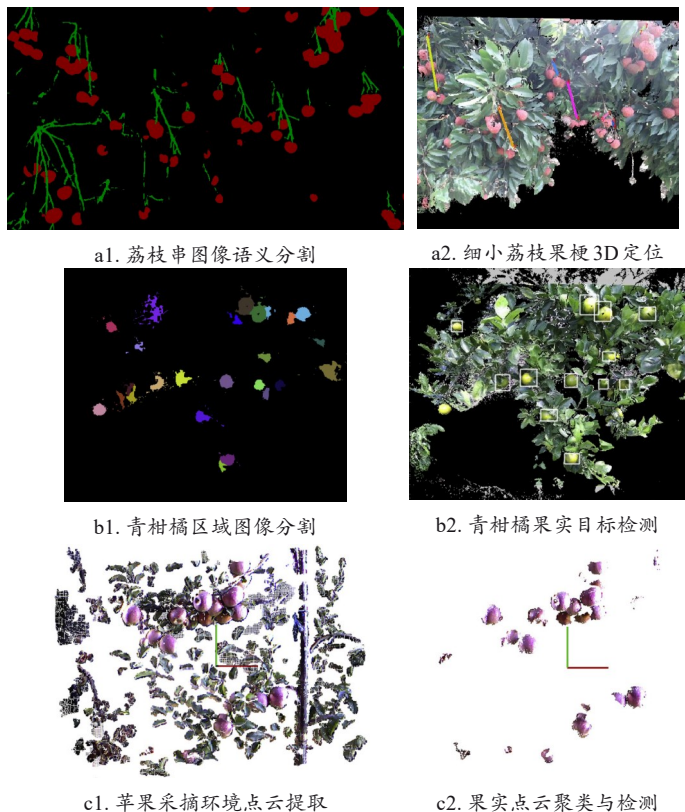


图4 基于RGB-D信息融合方法的局部目标视觉感知效果示意图

Fig. 4 Schematic diagram of the effects of local target visual perception based on RGB-D information fusion method

表2 主流消费级深度相机优缺点及其售价

Table 2 The advantages, disadvantages and prices of consumer-grade RGB-D cameras

相机型号	主要优点	主要缺点	售价/元
KinectV2	精度高	测距小于4.5 m、开发生态不佳、体积大、深度图分辨率像素仅有512×424	约1 500
RealSense SR300	精度高	受光照影响严重、测量距离小于2 m、深度图分辨率像素仅有640×480	约2 000
RealSense D435i	体积小、测距远	远距离信息失真严重、计算量大	约2 400
RealSense ZR300	体积小、功耗低	测量距离小于2.8 m、深度图分辨率像素仅有628×468、计算量大	约2 700
CamBoard Pico Flex	精度高	测量距离小于4 m、深度图分辨率像素仅有224×171	约2 600
Xtion PRO Live	精度高	测量距离小于3.5 m、受光照影响严重、深度图分辨率像素仅有640×480	约1 300

注:部分相机由于停产等原因,官方价格的指导意义不大。因此,表中价格数据均通过随机查阅主流购物平台中8个不同商家的价格取中位数,再四舍五入到百位得到,其更接近用户购买时的实际售价。



何构建的空间约束具有更强的抗扰性与稳定性，在水果表型参数测量、全表面重构和定位任务中应用较为广泛。近年来，多视图几何框架下的极线约束理论<sup>[57]</sup>、视差图优化策略<sup>[58]</sup>、视点规划<sup>[6]</sup>，以及多序列点云配准<sup>[7]</sup>等基础方法研究取得了一定进展，使得多视角信息融合方案更为灵活多变。在果园环境中，可通过多次移动同一个双目视觉系统，结合运动恢复结构（Structure from Motion, SfM）理论来模拟多目视觉系统的感知模式，扩大视觉感知范围并提升水果点云密度<sup>[59]</sup>；在多个相机的交汇视场中放置公共标靶可实现相机坐标的精确关联，构建高精度多视角信息融合系统<sup>[5]</sup>；此外，朝向同一视点的多个相机可形成多目汇聚式相机阵列，是实现植物3D模型高精度重建与关键表型参数提取的重要方案之一<sup>[60]</sup>。基于多视角信息融合方法的局部目标视觉感知效果如图5所示。

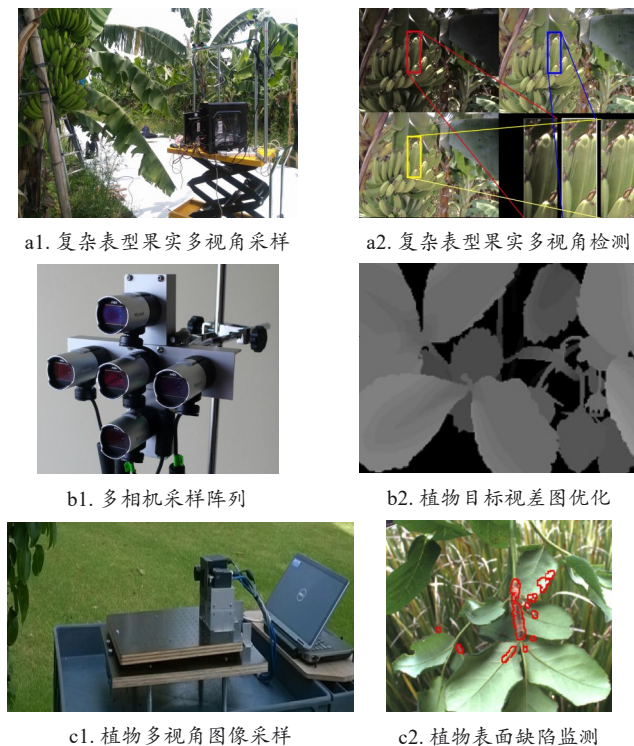


图5 基于多视角信息融合方法的局部目标视觉感知效果示意图  
Fig. 5 Schematic diagram of the effects of local target visual perception based on multi-view information fusion method

全方位、高精度和强鲁棒的多视角信息融合系统是未来构建强自主型采摘机器人重要基础。其中，基于多组立体相机构建的多视角采样系统需占据较大的采样空间或依赖复杂的多视图几何约束，其在某些狭小的果园位点上难以实现快速部署。相比之下，通过多自由度机械臂搭载单个立体相机，

可以便捷地实现小工作范围内的灵活采样，是未来多视角采样系统的重要发展方向。高质量臂载相机采样依赖于完善的采样策略以及实时环境反馈机制，其涉及作业过程的自主决策与规划问题，这将在第3.2节中详细介绍。

## 1.5 分析与评价

果园局部位点下的水果感知问题得到了长期和广泛关注，目前在视觉定位精度和实时性方面已取得较大进展，未来需要进一步解决对复杂场景和目标特征提取稳定性方面的问题与挑战。深度神经网络具有强大的特征提取与学习能力，在构建鲁棒、稳定和场景自适应的农业视觉系统方面具有广阔的应用前景。目标表面信息的完整感知是提高视觉系统性能的关键，目前最直接有效的手段是通过构建手眼视觉系统对目标进行全方位采样，并通过多视角信息融合方法来构建完整视觉信息。为了提高视觉系统综合性能，有必要继续研究深度神经网络与手眼视觉系统的深层耦合机制，形成具有更强鲁棒性与目标自适应性的端对端局部视觉感知方法体系。

## 2 果园全局地图构建

地图构建的目的是获取果园中的道路、种植行、果枝主干等大规模结构，实现机器人对果园宏观环境的理解，是后续机器人自主移动及采摘规划的基础。同时定位与建图（Simultaneous Localization and Mapping, SLAM）是实现全局地图构建的重要技术手段，其核心思想是通过机器人的移动实现对环境的持续采样，形成具备时间序列的连续数据流，然后结合图论、概率统计与知识推理等方法构建时间序列数据的关联，进而实现机器人对大规模连续物理环境的感知<sup>[61-63]</sup>。SLAM技术流程包括地图构建和传感器自定位两个部分，其同时运行并相互更新信息，因此SLAM方法具有较强的灵活性与自组织性。其中，地图构建部分对移动载体的形式没有严格要求，除了常见的地面移动平台外，还可以部署飞行载体或多个协同作业的载体<sup>[12]</sup>来实现建图。

视觉传感器是地图构建任务中最常用的传感器之一。视觉方法获取的信息量较大，其不仅包含三维表型和距离信息，还包含目标纹理和颜色等细节信息。通过对视觉序列图像进行一系列后处理可得到丰富的果园场景语义，构建包含水果、枝叶、道路和障碍物等标签的语义地图，有效指导机器人在

复杂果园中的移动、避障与采摘等工作。因此,视觉方法对果园地图构建任务具有较强的适用性。

## 2.1 基于视觉的果园建图

在视觉建图框架下,采摘机器人需全面采集果园场景的时间序列图像,基于立体视觉方法计算每一帧序列图像对应的三维点云,并通过合适的匹配策略得到序列点云间的空间关联,进而拼接得到全局点云地图。在复杂非结构化因素干扰下,如何准确可靠地从图像中提取各帧三维点云中的特征信息,并基于这些信息估计全局点云拼接位姿,是实现高质量果园地图构建的关键。

针对果园中的复杂纹理与背景干扰,可将单帧序列图像的处理问题视为第1节所述的局部目标感知问题,采取第1.2节中所述基于高级特征学习的方法,通过部署深度神经网络提取单帧图像的高级鲁棒特征。基于高级特征可实现对场景、道路、果实、枝叶等对象的目标检测<sup>[64]</sup>和语义分割<sup>[65]</sup>,构建具备感兴趣类别标签的全局语义点云地图<sup>[66]</sup>。

针对序列图像采样问题,通过在机器人上安装多个相机或以机械臂搭载单个相机进行多视点采样,可显著提升机器人对果园环境的采样覆盖率,确保特征采集的多样性,并缓解因遮挡和视觉盲区带来的信息丢失问题,提高全局点云位姿估计与帧间关联的稳定性<sup>[3,67]</sup>。

对于规模较大的果园场景,其采样所得的时间序列图像数据量往往较大,可考虑使用稀疏特征代替稠密特征来提高稠密建图效率,但相关稀疏特征需具备较强的旋转、平移或尺度不变性<sup>[68]</sup>。为了应对大规模果园不同区域中的自然力学扰动、光照变化等动态因素,视觉建图系统需要具备较强的实时性和动态更新能力。基于图优化和回环检测技术可有效减少累计建图误差。此外,采用增量式建图方法可实现在采集新数据的同时更新地图,适应动态环境变化<sup>[69]</sup>。

基于视觉方法的果园建图效果如图6所示,研究情况见表3。

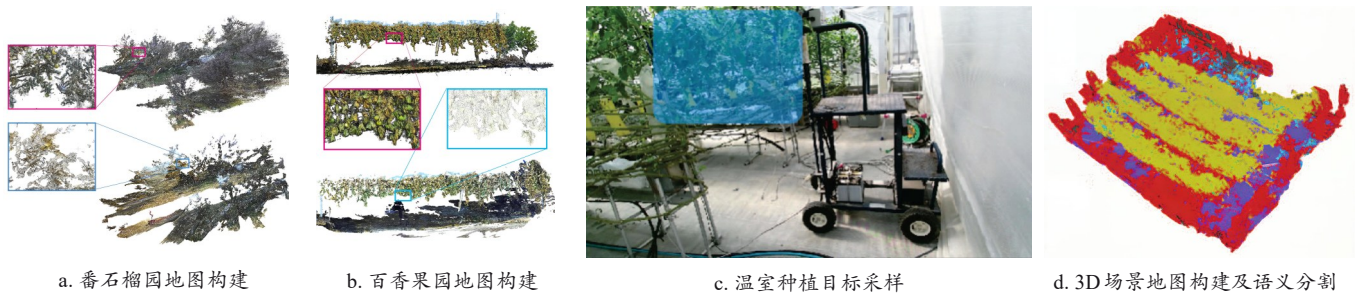


图6 基于视觉方法的果园建图效果示意图

Fig. 6 Schematic diagram of the effects of orchard mapping based on vision method

表3 近年基于视觉的果园建图方法研究情况

Table 3 Research on vision-based orchard mapping methods in recent years

编号	作者	年份	果园类型	立体视觉类型	地图类型
1	陈明猷等 <sup>[3]</sup>	2021	百香果、番石榴、柑橘、青枣	双目视觉	稠密点云地图
2	Nellithimaru 和 Krajnik <sup>[65]</sup>	2019	葡萄	双目视觉	语义点云地图
3	Matsuzaki 等 <sup>[66]</sup>	2018	番茄(温室)	RGB-D 视觉	语义点云地图
4	Dong 等 <sup>[67]</sup>	2020	适用多类果园	RGB-D 视觉	语义点云地图
5	Liu 等 <sup>[68]</sup>	2023	适用多类果园	RGB-D 视觉	稠密点云地图
6	Zhang 等 <sup>[69]</sup>	2022	适用多类果园	双目视觉	稠密点云地图

目前,果园环境下视觉信息的高效获取与高质量后处理仍然面临较多难题,导致视觉建图效果受限,主要体现在三个方面。首先,视觉建图系统的正常运行依赖于高质量帧间关联,要求基础视觉算法能持续稳定地提取场景中的稀疏特征或光流信息,并具备较强的异常状态响应与校正能力。然

而,受户外果园中的复杂因素影响,实际采样的全局序列图像广泛存在动态失焦、照度不均、视场突变等现象,难以满足系统长时间稳定运行的需求<sup>[70,71]</sup>;其次,以视觉里程计、回环检测、图优化、局部点云重建等为代表的视觉建图关键技术对非结构化农业环境的针对性优化不足,存在计算效



率低、全局累计误差大、点云分辨率低等应用受限问题；此外，为了实现高质量全局产量估计和实时移动避障路径规划，往往需要构建数据量较大的高精度稠密语义地图。由于野外环境下硬件性能以及能源供给的限制，已有农业视觉建图系统仍难以实现对实时性、效率与精度的兼顾<sup>[72, 73]</sup>。Capua等<sup>[74]</sup>曾指出，为了使主流建图系统适应复杂的农业场景，需要对视觉算法、系统硬件架构乃至其基础工作逻辑进行大幅修改与优化。

## 2.2 基于多传感融合的果园建图

除了视觉传感器外，常见于地图构建任务的传感器还有激光传感器、惯性测量单元（Inertial Measurement Unit, IMU）和全球导航卫星系统（Global Navigation Satellite System, GNSS）等。激光传感器可提供精确的场景距离与目标表型信息，惯性导航单元可提供运动学与动力学信息，而卫星导航模块则可为机器人提供稳定的全局位置信息。将这些传感器与视觉传感器结合，可降低基于单一视觉传感信息建图的模型误差，提高复杂干扰下果园建图的精度和稳定性。

视觉传感器与激光传感器融合感知的方案可基本满足3D场景建模对颜色、纹理与几何信息的数据量需求，为建图系统提供多重冗余数据，提高系统在全果园产量评估、植株表型重构与几何参数提取、林冠模型数据库构建等任务下的综合性能表现<sup>[75-78]</sup>。

视觉传感器与IMU融合的方案可大大强化机器人对自身运动状态的感知与实时调节能力，在视觉传感器数据不充分或发生遮挡时补充加速度、角速度与位置信息，优化关键帧的生成速率和帧间匹配精度，提升建图过程的稳定性<sup>[79]</sup>。

视觉传感器与GNSS融合的方案中，卫星数据可为视觉建图系统提供准确可测的全局坐标系，确保帧间关联精度。对于大规模、长时间跨度下的水果采摘任务，该坐标系是建图系统校正和补偿累积位姿估计误差的重要基准<sup>[80]</sup>。

进一步地，同时结合视觉传感器、激光传感器、IMU和GNSS，可构建具备强鲁棒性的建图系统，进一步提升建图过程的精度和稳定性<sup>[81-83]</sup>。基于多传感融合方法的果园建图效果如图7所示，研究情况见表4。

表4 近年基于多传感融合的果园建图方法研究情况

Table 4 Recent researches on multi-sensor-based orchard mapping methods

编号	作者	年份	建图对象	传感来源	地图类型
1	Shalal等 <sup>[75]</sup>	2015	适用多类果园	视觉+激光	稀疏点云地图
2	Underwood等 <sup>[76]</sup>	2016	杏仁园	视觉+激光	全局产量地图
3	Habibic等 <sup>[77]</sup>	2017	苹果园	视觉+激光	占用栅格地图
4	Chen等 <sup>[78]</sup>	2021	适用多类果园	视觉+激光	稠密点云地图
5	Sun等 <sup>[79]</sup>	2023	适用多类果园	视觉+IMU	稠密点云地图
6	Li等 <sup>[80]</sup>	2023	适用多类果园	视觉+GNSS	稠密点云地图
7	Gan等 <sup>[81]</sup>	2017	柑橘园	视觉+GNSS+IMU	全局产量地图
8	Hua等 <sup>[82]</sup>	2023	适用多类果园	视觉+GNSS+IMU	全局轨迹地图
9	Zhu等 <sup>[83]</sup>	2023	适用多类果园	视觉+激光+GNSS+IMU	稠密点云地图

基于多传感融合的方法在提高建图精度和稳定性方面具有显著优势，但也存在明显局限性。一方面，多传感器系统的集成和全局标定相对复杂，需要在不同传感器之间建立精确的时空同步机制，大大增加了系统与算法设计的复杂度，也对硬件的算力、成本及其运行实时性提出更高要求。另一方面，果园中的天气与光照条件变化、传感器遮挡等因素也会影响多传感器系统的整体性能。例如，果园枝叶遮挡和动态光照容易导致视觉传感器采样质量下降，而雨雪等恶劣天气容易影响激光传感器和

GNSS数据的可靠性。如何在复杂、多变的果园环境中保持多传感器系统的高效稳定运行仍然是一个亟待解决的问题。

## 2.3 分析与评价

视觉建图方法可有效获取果园内水果、枝叶和障碍物空间语义信息，并通过图像的颜色、形状和灰度特征来评估产量和生长状态，为移动底盘和机械臂的作业行为规划提供重要参考。然而，视觉建图在非结构化环境中易受动态失焦、光照变化和遮挡等因素干扰，需进一步优化光照一致性、算法鲁



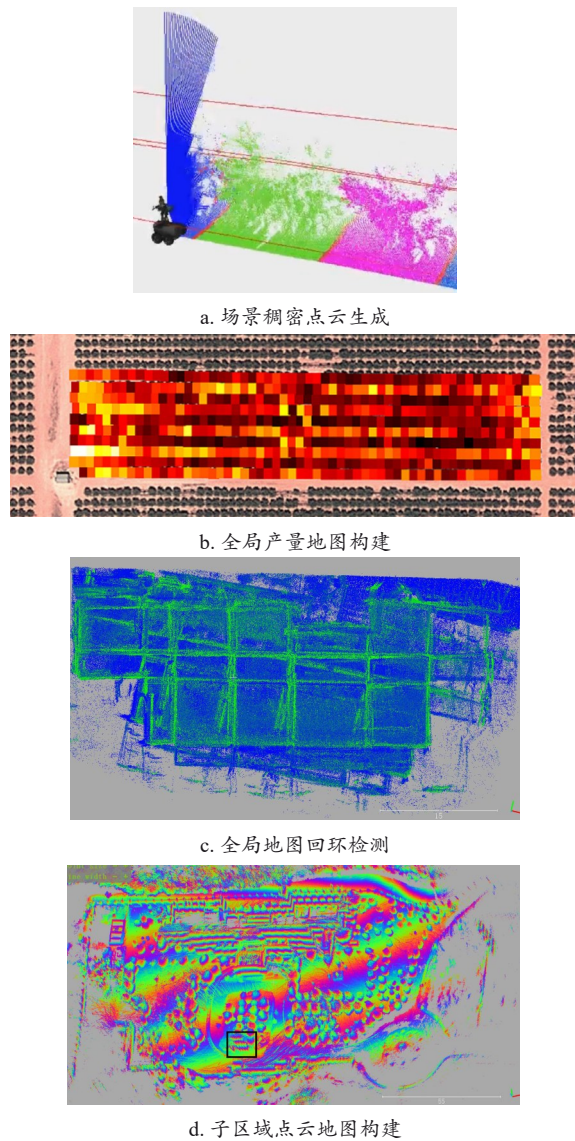


图7 基于多传感融合方法的果园建图效果示意图

Fig. 7 Schematic diagram of the effects of multi-sensor-based orchard mapping method

棒性和点云拼接的实时性。

多传感融合建图方法结合了视觉传感器、激光传感器、IMU和GNSS等数据,提高了果园建图的精度和稳定性。然而,数据融合的过程显著提升了系统的复杂性及其集成难度,对硬件性能和算法设计提出了更高要求。同时,恶劣天气和动态光照变化等环境因素对多传感系统的性能影响问题仍需解决。

### 3 果园自主作业行为规划

在视觉感知的基础上,机器人应能自主连贯地移动到果园中的多个不同位点并实施采摘。因此,其自主作业行为主要包括在全局果园中的自主移动

和在局部场景位点下的自主采摘。

#### 3.1 全局果园中的自主移动行为规划

采摘机器人的自主移动行为规划主要包括全局尺度下的移动路径规划,及其在移动过程中的局部路径规划。

在非农业场合,自主移动机器人已成功应用于环境探测<sup>[84,85]</sup>、室内作业<sup>[86-90]</sup>、楼道救援<sup>[91]</sup>、安全管理<sup>[92-94]</sup>和虚拟仿真<sup>[95]</sup>等众多领域,目前不乏成熟案例。相比之下,复杂果园环境下的自主移动机器人技术仍处于发展阶段。为了实现全果园范围内的连贯移动与采收作业,机器人不但需要具备基本的自定位和运动能力,还需要具备基于果园语义进行移动决策与规划,准确导航至最优采摘位点并与采摘模块高效协同的能力,这涉及地图后处理、基于产量的实时调度、高安全性路径规划、高精度自定位和异常状态校正等问题,对已有移动机器人技术提出了更高的要求。

针对移动过程中的采样质量问题,可考虑构建面向全景采样的视觉感知模块,设计跨场景光照自适应与补偿算法<sup>[8]</sup>,部署结构光硬件系统<sup>[96]</sup>,提高机器人跨片区移动时对场景光照与视野变化的自适应性。同时,可借助大角度回旋机构带动传感器采样,进一步扩大自主导航系统的感知范围<sup>[97]</sup>。

针对非结构化果园道路识别问题,由于不同区域中道路的形状、材料属性与表型特征存在差异,传统基于单一阈值的视觉识别方法往往难以适应该类工况。部署特征自学习的神经网络、设计针对道路区域的注意力模块、采用无监督学习等,均是提高视觉算法对复杂路面特征识别的鲁棒性的有效策略<sup>[98,99]</sup>。在道路特征不明显的情况下,以道路两侧种植行为基准进行路径规划也是一种有效的方案<sup>[100]</sup>。

针对自主移动中地图与定位数据的稳定性问题,需合理设计机器人内部各节点模块的数据交换逻辑<sup>[101]</sup>,基于果园先验知识构建SLAM状态估计器,降低场景目标状态不确定性对底盘全局运动位姿估计的影响<sup>[102,103]</sup>。全果园自主移动行为规划效果如图8所示。

需要指出的是,目前采摘机器人移动行为规划相关研究多聚焦于机器人本体的实时自定位与移动避障问题,而自定位过程中所构建的全局地图利用率较低。实际上,对于以水果收获为目标的采摘机器人而言,通过对高分辨率全局三维地图进行后处

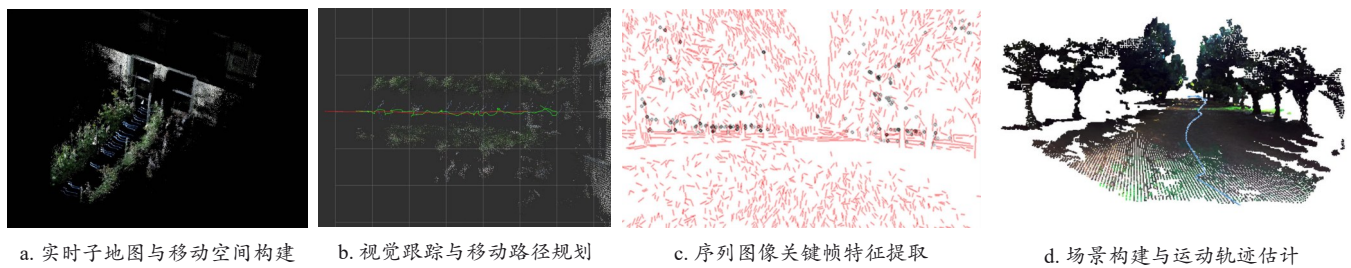


图8 全果园自主移动行为规划效果示意图

Fig. 8 Schematic diagram of the effects of autonomous movement behavior planning in an entire orchard

理与细节提取，可以为机器人提供全局水果目标的生长状况、空间分布和成熟度等重要收获信息，提高机器人移动决策与路径规划的效率与目标指向性。因此，地图的结构细节信息应被进一步融合至移动行为规划的技术流程中。总的来说，机器人的移动模块需要同时具备较强的自定位、地图后处理，以及基于地图语义的实时信息提取与决策能力，这对机器人的硬件架构、算法鲁棒性，以及工作逻辑设计提出了更高的要求，亟待进一步深入的研究。

### 3.2 局部场景下的自主采摘行为规划

机器人随底盘移动到特定采摘位点后，需基于当前视野下水果的几何表型、障碍物分布及实时环境状况，自主规划最优采样视点以获取高质量采摘位姿，并规划其最优采摘避障路径。因此，机器人自主采摘行为规划主要包括机器人在视觉采样过程中的主动视点规划，以及在采摘进给过程中的避障路径规划。

在采摘的过程中，系统需实时评估潜在的障碍物碰撞、果实损伤与定位误差等情况，建立基于实时视觉信息的视点与路径校正机制。“眼在手上”的Eye-in-Hand视觉系统通过机械臂带动相机在不同视野下进行采样，具有较强的部署灵活性与环境交互性，是高质量视觉伺服控制与自主采摘行为规划的关键载体<sup>[104-107]</sup>。

针对采摘视点规划问题，通过离线预设的运动

规则建立末端运动学模型，控制“手-眼”视觉系统在物体前方多个视点进行精细采样，可有效筛选最优采摘视点，获取完整的目标表面信息并生成实时障碍物空间，提高机器人对采摘区域的感知精度和全面性<sup>[108]</sup>。然而有研究指出，由于场景中存在未知障碍物和随机扰动，基于固定运动学模型的遍历式视点规划策略无法完全解决遮挡问题，需进一步辅以二次视点规划与校正处理<sup>[57]</sup>。在离线运动学模型的基础上增加远-近协同的相机视点，远视点用于快速筛选可采摘区域，近视点用于对场景进行局部重采样与精定位，从而引导机器人完成高质量采摘动作<sup>[109]</sup>。

针对采摘避障路径规划问题，基于“手-眼”协同与高频视觉伺服的末端行进方案在本领域较常见。在末端靠近采摘点的过程中，“手-眼”视觉系统以较高的频率持续计算并更新目标位姿信息，可实时校正末端的采摘路径，实现对果实真实位置的精确逼近<sup>[100,110]</sup>。近年来，基于深度强化学习的局部路径规划方法研究取得较大进展。通过在仿真空间内收集机械臂-环境行为交互的经验信息，可引导机械臂学习最优采摘运动学参数，实现端对端采摘避障路径规划。由于深度学习方法具有较强的特征提取与复杂环境适应能力，其在构建强鲁棒采摘行为规划系统方面具有较大发展潜力<sup>[111]</sup>。局部场景下自主采摘行为规划效果如图9所示。

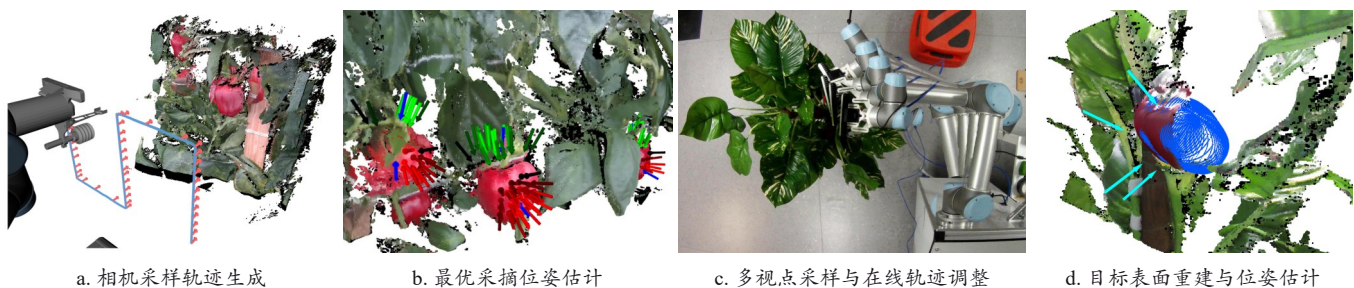


图9 局部场景下自主采摘行为规划效果示意图

Fig. 9 Schematic diagram of the effects of autonomous picking behavior planning in local scenes



从已有案例不难发现,“手-眼”协同是实现高质量自主采摘行为规划的关键。然而,已有的“手-眼”视觉系统仍然存在局限性。一方面,环绕目标的多视点规划与采样需要耗费大量计算资源和时间,从而导致采摘效率下降,这是“手-眼”多视点采样模式的共性问题<sup>[112]</sup>。另一方面,已有的“手-眼”协同机制过于单一,过度依赖人工设计的运动学模型与规则,缺乏对果园中动态目标和场景的自适应机制,难以满足多元化水果收获需求。为了进一步提升“手-眼”视觉系统对果园采摘任务的适应性,有必要深入研究具备环境交互与自学习能力的行为决策模型,形成动态、智能与柔性化的自主采摘行为。

### 3.3 分析与评价

机器人的自主作业行为规划包括其在全局果园中的自主移动行为规划和在果园局部位点下的自主采摘行为规划。

在自主移动行为规划方面,目前相关研究侧重于机器人的自定位问题,对全局地图中细节信息的利用率较低。为了提高机器人自主移动效率,有必要进一步关注对全局地图中的语义、结构与产量等与采摘过程密切相关的细节信息的挖掘与后处理,形成针对果园采摘任务的新型自主移动行为规划方法与框架。

在自主采摘行为规划方面,“手-眼”协同的采摘模式在户外环境下具有较强的灵活性与机动性,是实现高效自主采摘的关键载体。然而,已有基于手眼协同的视点规划与避障路径规划方法大多存在采样耗时长、采摘规划机制单一等问题。研究具备动态环境交互与自学习能力的行为决策模型,形成动态与柔性化的自主采摘作业行为,是未来进一步实现智能水果收获的关键。

## 4 采摘机器人自主作业系统构建

采摘机器人是一个复杂的系统<sup>[113]</sup>。为了构建实用型作业机器人,实现自主高效采摘,系统不但需要具备前述局部感知、全局建图与作业行为规划等基础功能,还需要基于一定的策略与准则集成感知、移动和采摘模块,解决“眼-身-手”行为高效协同和全果园连贯作业的关键技术瓶颈问题。

### 4.1 采摘机器人系统

21世纪初,诸多集成度高、结构紧凑的采摘机器人系统被相继推出,在柑橘<sup>[114]</sup>、番茄<sup>[115]</sup>、茄

子<sup>[116]</sup>、草莓<sup>[117,118]</sup>、樱桃<sup>[119]</sup>、黄瓜<sup>[120]</sup>、苹果<sup>[121,122]</sup>等果蔬采摘任务中崭露头角。2014年前后,随着并行计算硬件的快速迭代升级,以深度学习为代表的人工智能技术取得较大进步,并在随后数年迁移应用至农业工程领域,有效解决了以往采摘机器人在感知精度、控制鲁棒性和规划自适应性等方面的诸多基础技术问题。目前,采摘机器人相关研究已进入新的阶段,朝着系统化、集成化与实用化和商业化方向发展。感知、移动和采摘模块的高效协同是构建强自主与实用型采摘机器人的关键瓶颈问题。

感知模块与采摘模块的“手-眼”协同是机器人多功能模块协同的基础。通过相机标定、手眼标定与机械臂正运动学参数标定可精确求解视觉系统与机械臂基坐标系的转换关系,进而建立“手-眼”视觉测量模型<sup>[104]</sup>。基于端对端的深度目标检测网络、语义分割网络和路径规划网络,可有效提高“手-眼”视觉系统的识别、定位与测量精度,实现对采摘过程的高质量视觉伺服控制<sup>[123]</sup>。对于不同外形的水果,应针对设计与其适配的末端爪形,以及抓取-分离-放置策略<sup>[124]</sup>,确保视觉伺服控制效率。针对水果脆弱易损的生物学特性,可全面考虑牵拉、旋转、按压、折断、按压-折断组合、牵拉-旋转组合等多种摘取模式,并基于所选择的最优摘取模式设计“手-眼”视觉系统、运动控制系统和末端控制系统的运行逻辑与采摘动作序列<sup>[125]</sup>,实现高效、无损水果采收<sup>[126]</sup>。此外,采用采摘-放置一体化的集成式末端是也降低视觉算法复杂度,提高“手-眼”协同效率的有效方法之一<sup>[127]</sup>。

上述“手-眼”协同模式的局限性在于其仅能作用于果园局部位点。为了进一步提高机器人的有效作业范围,可在“手-眼”协同的基础上添加移动平台节点,实现全果园作业。目前,果园移动系统以人工操控或半自动操控居多,其依赖特定的运动规则或场地硬件设施,如借助承载平台垂直上升和下降<sup>[128]</sup>、以轨道车的形式搭载机械臂在特制化道路移动等<sup>[27]</sup>。这种相对固定的策略可确保机器人移动过程稳定性,但大大限制了其作业过程的自主性和智能性。为了构建自主连贯作业系统,移动平台节点需进一步参与至“手-眼”协同过程中,通过分析来自“手-眼”视觉系统的场景几何结构、语义信息,并结合自身的运动学信息来规划全果园移动序列<sup>[129-131]</sup>,形成“眼-身-手”闭环协同模式,这将涉及更复杂的环境感知、路径规划、视点规划

算法与数据流设计<sup>[132, 133]</sup>问题。

机器人在全果园不同片区之间来回移动与采摘的过程中，其视场范围和光照条件处于动态变化的状态，这是影响机器人“眼-身-手”行为协同稳定性的重要因素。针对光照自适应性问题，可设计专用照明模块进行物理补光，提高采样图像的光照一致性<sup>[134, 135]</sup>。通过部署基于深度学习的光照补偿网络，以数据驱动的方式动态增强图像亮度，也是提高复杂光照条件下视觉伺服控制系统稳定性的有效手段之一<sup>[136]</sup>。针对动态视场下的尺度自适应性问题，一方面可通过控制移动底盘主动靠近目标的策略来缩小视场，并采用主动视点规划和物理清除的方式排除障碍物，提高目标像素占比<sup>[25, 137]</sup>；另一方面，可在强鲁棒性深度学习模型的基础上，结合骨架细化算法提高视觉系统对细小图像目标的敏感

性<sup>[19]</sup>，引导机器人在动态、大视场条件下实现连贯的感知-移动-收获行为。

除了尽可能提升基础算法与硬件的性能外，还应面向农户的实际生产需求，适当考虑系统通用性与采收效率之间的平衡问题。例如，对于大批量采收任务，可以通过分析当前水果的生物学特点<sup>[138, 139]</sup>，定制化设计机器人本体架构、感知算法和末端采摘逻辑，力求降低系统设计和模块间协同的复杂度。该策略的局限性在于可能会降低系统对其他类型水果采摘任务的通用性，但构建系统所需的经济和时间成本将大大下降。此外，部署多机械臂系统，研究多臂运动的顺序优化与高效调度机制，是提高采收效率的另一可行方案<sup>[140-144]</sup>。

近年报道的代表性采摘机器人系统如图 10 所示，研究情况见表 5。

表 5 近年报道的采摘机器人自主作业系统模块协同及其应用情况

Table 5 Recent reports on the coordination of modules in autonomous picking robot systems and their applications

编号	作者	年份	参与协同模块	应用对象	系统功能
1	Li 等 <sup>[19]</sup>	2020	感知模块+采摘模块	荔枝	细小果梗识别、路径规划、采摘
2	Ye 等 <sup>[24]</sup>	2019	感知模块+采摘模块	荔枝	视点规划、果实识别、采摘
3	Arad 等 <sup>[27]</sup>	2020	感知模块+采摘模块	甜椒	果实识别、采摘路径规划
4	Silwal 等 <sup>[122]</sup>	2016	感知模块+采摘模块	苹果	果实识别、采摘
5	Fu 等 <sup>[123]</sup>	2024	感知模块+采摘模块	猕猴桃	多果实时采摘
7	Wang 等 <sup>[124]</sup>	2023	感知模块+采摘模块	黄桃	果实识别、无碰撞路径规划
8	Gao 等 <sup>[125]</sup>	2024	感知模块+采摘模块	樱桃番茄	果实识别、采摘
9	Zhuang 等 <sup>[132]</sup>	2023	感知模块+采摘模块	通用对象	水果无损采摘
10	Lee 等 <sup>[134]</sup>	2019	感知模块+采摘模块	甜椒	光照补偿、果实识别、采摘
11	Sun 等 <sup>[135]</sup>	2023	感知模块+采摘模块	通用对象	光照补偿、果实识别、采摘
12	Zeeshan 等 <sup>[136]</sup>	2024	感知模块+采摘模块	通用对象	光照补偿、果实识别、采摘
13	Wang 等 <sup>[137]</sup>	2023	感知模块+采摘模块	荔枝	果实识别、主动障碍物清除、采摘
14	Onishi 等 <sup>[138]</sup>	2019	感知模块+采摘模块	通用对象	高度定制化模式下的果实识别、采摘
15	Williams 等 <sup>[140]</sup>	2019	感知模块+采摘模块	猕猴桃	多臂协同采摘
16	Ling 等 <sup>[141]</sup>	2019	感知模块+采摘模块	番茄	双臂协作采摘
17	Xiong 等 <sup>[142]</sup>	2020	感知模块+采摘模块	草莓	主动障碍物分离、双臂协作采摘
18	Wang 等 <sup>[18]</sup>	2017	感知模块+移动模块+采摘模块	温室番茄	果实识别、自主导航、采摘
19	Roshanianfard 等 <sup>[130]</sup>	2018	感知模块+移动模块+采摘模块	南瓜	果实识别、移动路径规划、采摘
20	De Preter 等 <sup>[131]</sup>	2018	感知模块+移动模块+采摘模块	草莓	果实识别、移动路径规划、采摘
21	陈明猷等 <sup>[133]</sup>	2024	感知模块+移动模块+采摘模块	通用对象	果实识别、移动路径规划、采摘
22	Almendral 等 <sup>[139]</sup>	2018	感知模块+移动模块+采摘模块	橙子	多果实同时识别与采摘

注：虽然部分采摘机器人系统具备移动模块，但由于该模块未能与其他模块进行实质性数据交换与行为协同，因此未在“参与协同模块”中列出。



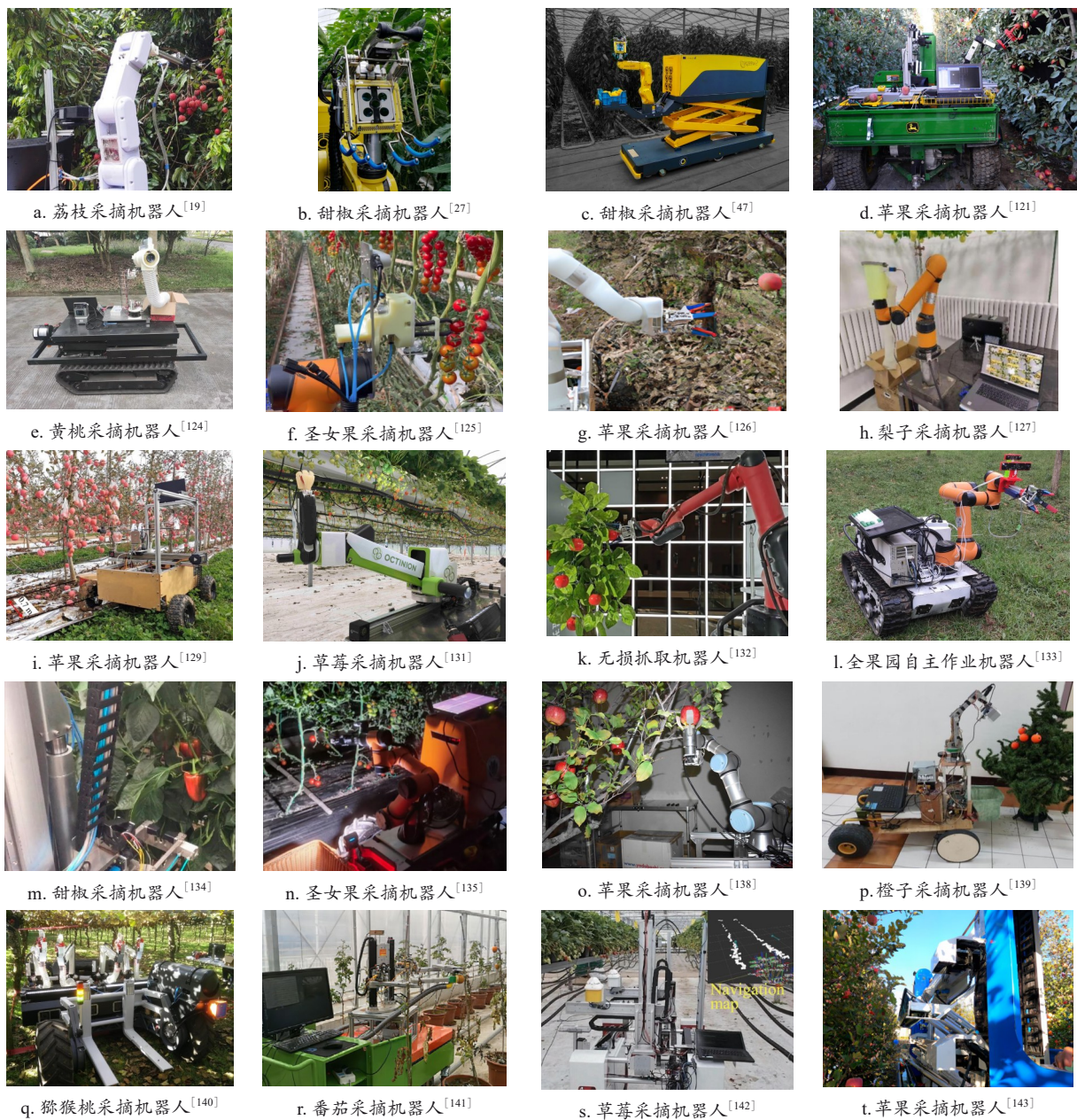


图 10 代表性采摘机器人系统示意图

Fig. 10 Schematic diagram of representative picking robot systems

## 4.2 分析与评价

国内外报道的采摘机器人系统均具备识别水果的基础功能,同时一定程度上具备移动-采摘综合作业能力,可为实现大规模无人化水果采摘提供重要理论与技术支持。然而,已有采摘系统的感知、移动、采摘模块功能相对割裂,协同程度不高,未能切实满足全果园自主高效作业需求,极大限制了采摘机器人在真实大规模场景中的部署应用。

已有采摘机器人系统构建案例中,可见针对局部目标感知、大规模果园建图、移动与采摘行为规

划等单独任务的案例或功能演示,但是具备“眼-身-手”行为协同与全果园一体化作业能力的采摘系统不多见。可查及的大部分移动平台仅充当简单的机械承载部件,机器人不具备对感知、移动、采摘模块进行高效协同并实现全果园连贯作业的能力。部分具备自主移动能力的采摘机器人系统需依赖固定轨道、人工轨迹或标记点,对场地布置要求较高,可移植性和实用性有待提高。目前为止,采摘机器人的模块协同多限于“手-眼”协同。综上所述,尽管部分关键技术取得突破,目前采摘机器



人系统相关研究仍未成熟，离自主、智能和实用化的要求还有较大距离。

## 5 展 望

围绕采摘机器人全果园自主连贯作业问题，本文对涉及的果园局部目标感知、全局果园地图构建和自主作业行为规划三项关键技术进行深入调研和分析。在此基础上，进一步报道了已有采摘机器人自主作业系统构建的相关工作，结合果园采摘需求对其性能进行分析和评估。

目前为止，采摘机器人在大规模场景、长时间跨度工况下的稳定性、可靠性和作业效率方面仍面临诸多挑战：视觉算法普遍依赖人工特征或大规模训练样本，缺乏针对非结构化果园的专用视觉算法架构，导致对复杂水果目标的感知质量偏低，信息缺失；建图模块易受动态失焦、光照变化和遮挡等非结构化因素干扰，其采样范围、光照一致性和场景拼接的实时性不足，且对恶劣天气和自然条件变化的适应性亟待提高；移动行为规划模块对全局地图中的语义、结构与产量等细节信息的挖掘与后处理不足，导致全局地图使用效率不高；采摘行为规划模块大多局限于“手-眼”协同，存在采样耗时长、采摘规划机制单一等问题，缺乏动态环境交互与自学能力，一定程度限制了采摘成功率。总的来说，构建可替代人类采摘员的自主作业系统仍是本领域的长期目标，其未来展望如下。

(1) 立体视觉与深度神经网络需有更深层的结合。目前，应用于采摘机器人的深度学习网络大多以2D图像为处理对象。实际上，深度神经网络可以经过进一步的结构优化而内嵌于立体匹配、3D点云后处理和作业行为规划技术流程中，辅助构建具备更强鲁棒性和稳定性的3D视觉感知模块和自主作业模块。

针对视觉算法的稳定性与抗扰性问题，未来可考虑构建面向复杂农业场景的轻量型深度立体匹配网络，通过自动提取高级特征实现端对端视差图生成，避免传统立体匹配对人工阈值的依赖。针对全局建图中序列图像的光照一致性问题，可考虑在深度学习范式下构建新型智能光照补偿算法，通过赋予照明模块参数化描述，使其成为独立于感知、移动和采摘模块以外的另一具备可学习参数的模块，并参与到多模块协同作业的过程中；同时，可考虑构建大型深度3D语义分割网络，辅以生成对抗网络或扩散模型进行场景样本扩充，实现基于小样本

自学的大型果园地图构建。针对机器人自主作业行为规划策略单一、作业效率较低的问题，可在深度强化学习框架下构建虚拟-实景结合的训练环境，将视觉系统的动作参数序列化为可训练的动作基元，以环境交互式学习的形式获取机器人最佳移动-采摘策略，建立具备强动态与环境自适应性的自主作业行为规划模型。

需要指出的是，大规模部署深度神经网络将显著增加系统硬件的计算负荷，这意味着需进一步提高感知、移动、采摘等模块中并行计算单元的数量规模，提供更富余和更稳定的外部能源输入，并快速排除因大规模计算产生的大量余热。如何在能源供给相对受限、气温环境条件相对不可控、所允许的硬件部署体积相对较小的户外采摘作业任务中实现模块性能、硬件成本、户外能源供给与系统热稳定性间的平衡，这是一个复杂但有意义的工程问题。

(2) 进一步提升视觉系统的感知范围。针对视觉数据采样完整性与障碍物遮挡问题，目前已有基于多视点联合采样与多传感融合的成功尝试。在未来工作中，需进一步提高视觉系统对目标对象采样的覆盖率，扩大感知范围，如通过在移动底盘上部部署全景低畸变相机阵列，建立多个相机在移动过程中的坐标协同与数据融合模式。在多视角信息融合感知方面，除了基于机器人运动学模型进行点云拼接外，可进一步借助自然环境中丰富的三维特征信息构建具备场景自适应性的多视图几何约束，实现大视场、跨场景自适应采样。此外，可考虑借鉴工业界的多模态信息融合与柔顺控制策略，采集类型更全面的传感数据，研究视觉传感器与力/触/滑觉传感器在农业环境下的最优部署方案，构建对复杂地形扰动具备柔顺补偿能力的自适应运动控制模型，提高视觉感知全面性与自适应性。

(3) 进一步提高“感知-移动-采摘”模块的协同程度。一方面，需系统化探明果园场景目标分布的共性特点，并将其作为视觉感知、路径规划、运动控制等一般化算法的额外约束，进而建立针对果园场景的专用算法体系，完善采摘机器人“眼-身-手”高效协同机制。另一方面，需建立面向机器人全果园连贯作业的综合性能评估体系。除了采用针对视觉识别、路径规划、移动导航、水果采摘等独立过程的性能描述指标外，还需进一步设计面向“感知-移动-采摘”全过程的综合性指标，以量化多模块的协同程度以及移动采摘机器人的整体作业能



力, 满足未来更为广泛和深入的系统化测试需求, 辅助构建高性能、低成本和面向农户的实用型自主作业机器人系统。

**利益冲突声明:** 本研究不存在研究者以及与公开研究成果有关的利益冲突。

### 参考文献:

- [1] 李会宾, 史云. 果园采摘机器人研究综述[J]. 中国农业信息, 2019, 31(6): 1-9.  
LI H B, SHI Y. Review on orchard harvesting robots[J]. China agricultural informatics, 2019, 31(6): 1-9.
- [2] 张鹏, 张丽娜, 刘铎, 等. 农业机器人技术研究现状[J]. 农业工程, 2019, 9(10): 1-12.  
ZHANG P, ZHANG L N, LIU D, et al. Research status of agricultural robot technology[J]. Agricultural engineering, 2019, 9(10): 1-12.
- [3] CHEN M Y, TANG Y C, ZOU X J, et al. 3D global mapping of large-scale unstructured orchard integrating eye-in-hand stereo vision and SLAM[J]. Computers and electronics in agriculture, 2021, 187: ID 106237.
- [4] JIN Y C, YU C C, YIN J J, et al. Detection method for table grape ears and stems based on a far-close-range combined vision system and hand-eye-coordinated picking test[J]. Computers and electronics in Agriculture, 2022, 202: ID 107364.
- [5] CHEN M Y, TANG Y C, ZOU X J, et al. Three-dimensional perception of orchard banana central stock enhanced by adaptive multi-vision technology[J]. Computers and electronics in agriculture, 2020, 174: ID 105508.
- [6] JADHAV T, SINGH K, ABHYANKAR A. Volumetric estimation using 3D reconstruction method for grading of fruits[J]. Multimedia tools and applications, 2019, 78(2): 1613-1634.
- [7] WU J W, XUE X Y, ZHANG S C, et al. Plant 3D reconstruction based on LiDAR and multi-view sequence images[J]. International journal of precision agricultural aviation, 2018, 1(1): 37-43.
- [8] 李盛辉, 夏春华, 姬长英, 等. 自主导航农业车辆的全景视觉同时定位与地图创建[J]. 江苏农业学报, 2017, 33(3): 598-609.  
LI S H, XIA C H, JI C Y, et al. Simultaneous localization and mapping for autonomously-navigating agricultural vehicle based on panoramic vision[J]. Jiangsu journal of agricultural sciences, 2017, 33(3): 598-609.
- [9] URVINA R P, GUEVARA C L, VÁSCONEZ J P, et al. An integrated route and path planning strategy for skid - steer mobile robots in assisted harvesting tasks with terrain traversability constraints[J]. Agriculture, 2024, 14(8): ID 1206.
- [10] XIONG J T, HE Z L, LIN R, et al. Visual positioning technology of picking robots for dynamic litchi clusters with disturbance[J]. Computers and electronics in agriculture, 2018, 151: 226-237.
- [11] ZOU X G, YE M, LUO C Y, et al. Fault-tolerant design of a limited universal fruit-picking end-effector based on vision-positioning error[J]. Applied engineering in agriculture, 2016, 32(1): 5-18.
- [12] LIN G, TANG Y, ZOU X, et al. Guava detection and pose estimation using a low-cost RGB-D sensor in the field[J]. Sensors (basel), 2019, 19(2): ID E428.
- [13] WANG Z, WALSH K B, VERMA B. On-tree mango fruit size estimation using RGB-D images[J]. Sensors (basel, Switzerland), 2017, 17(12): ID E2738.
- [14] CHEN M Y, TANG Y C, ZOU X J, et al. High-accuracy multi-camera reconstruction enhanced by adaptive point cloud correction algorithm[J]. Optics and lasers in engineering, 2019, 122: 170-183.
- [15] ZHUANG J J, HOU C J, TANG Y, et al. Computer vision-based localisation of picking points for automatic litchi harvesting applications towards natural scenarios[J]. Biosystems engineering, 2019, 187: 1-20.
- [16] WANG C L, TANG Y C, ZOU X J, et al. A robust fruit image segmentation algorithm against varying illumination for vision system of fruit harvesting robot[J]. Optik, 2017, 131: 626-631.
- [17] SUN S S, JIANG M, HE D J, et al. Recognition of green apples in an orchard environment by combining the Grab-Cut model and Ncut algorithm[J]. Biosystems engineering, 2019, 187: 201-213.
- [18] WANG L L, ZHAO B, FAN J W, et al. Development of a tomato harvesting robot used in greenhouse[J]. International journal of agricultural and biological engineering, 2017, 10(4): 140-149.
- [19] LI J H, TANG Y C, ZOU X J, et al. Detection of fruit-bearing branches and localization of litchi clusters for vision-based harvesting robots[J]. IEEE access, 2020, 8: 117746-117758.
- [20] OLATUNJI J R, REDDING G P, ROWE C L, et al. Reconstruction of kiwifruit fruit geometry using a CGAN trained on a synthetic dataset[J]. Computers and electronics in agriculture, 2020, 177: ID 105699.
- [21] WANG C L, LUO T H, ZHAO L J, et al. Window zooming-based localization algorithm of fruit and vegetable for harvesting robot[J]. IEEE access, 2019, 7: 103639-103649.
- [22] TAO Y T, ZHOU J. Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking[J]. Computers and electronics in agriculture, 2017, 142: 388-396.
- [23] LIN G C, ZHU L X, LI J H, et al. Collision-free path planning for a guava-harvesting robot based on recurrent deep reinforcement learning[J]. Computers and electronics in agriculture, 2021, 188: ID 106350.
- [24] YE L, DUAN J L, YANG Z, et al. Collision-free motion planning for the litchi-picking robot[J]. Computers and electronics in agriculture, 2021, 185: ID 106151.
- [25] CAO X M, ZOU X J, JIA C Y, et al. RRT-based path planning for an intelligent litchi-picking manipulator[J]. Computers and electronics in agriculture, 2019, 156: 105-118.
- [26] BAC C W, VAN HENTEN E J, HEMMING J, et al. Harvesting robots for high-value crops: State-of-the-art review and challenges ahead[J]. Journal of field robotics, 2014, 31(6): 888-911.
- [27] ARAD B, BALENDONCK J, BARTH R, et al. Development of a sweet pepper harvesting robot[J]. Journal of field robotics, 2020, 37(6): 1027-1039.
- [28] FU L S, GAO F F, WU J Z, et al. Application of consumer

- RGB-D cameras for fruit detection and localization in field: A critical review[J]. *Computers and electronics in agriculture*, 2020, 177: ID 105687.
- [29] GONGAL A, AMATYA S, KARKEE M, et al. Sensors and systems for fruit detection and localization: A review[J]. *Computers and electronics in agriculture*, 2015, 116: 8-19.
- [30] WU G, LI B, ZHU Q B, et al. Using color and 3D geometry features to segment fruit point cloud and improve fruit recognition accuracy[J]. *Computers and electronics in agriculture*, 2020, 174: ID 105475.
- [31] GE L, YANG Z, SUN Z, et al. A method for broccoli seedling recognition in natural environment based on binocular stereo vision and Gaussian mixture model[J]. *Sensors (basel, Switzerland)*, 2019, 19(5): ID E1132.
- [32] DAI N, XIE H, YANG X J, et al. Recognition of cutting region for pomelo picking robot based on machine vision[C]// 2019 ASABE Annual International Meeting. St. Joseph, Missouri, USA: American Society of Agricultural and Biological Engineers, 2019.
- [33] FU L S, TOLA E, AL-MALLAHI A, et al. A novel image processing algorithm to separate linearly clustered kiwi-fruits[J]. *Biosystems engineering*, 2019, 183: 184-195.
- [34] WANG D D, HE D J, SONG H B, et al. Combining SUN-based visual attention model and saliency contour detection algorithm for apple image segmentation[J]. *Multimedia tools and applications*, 2019, 78(13): 17391-17411.
- [35] GE Y Y, XIONG Y, FROM P J. Symmetry-based 3D shape completion for fruit localisation for harvesting robots[J]. *Biosystems engineering*, 2020, 197: 188-202.
- [36] LUO L F, TANG Y C, ZOU X J, et al. Vision-based extraction of spatial information in grape clusters for harvesting robots[J]. *Biosystems engineering*, 2016, 151: 90-104.
- [37] XIE H, DAI N, YANG X J, et al. Research on recognition methods of pomelo fruit hanging on trees base on machine vision[C]// 2019 ASABE Annual International Meeting. St. Joseph, Missouri, USA: American Society of Agricultural and Biological Engineers, 2019.
- [38] QIU C, TIAN G Z, ZHAO J W, et al. Grape maturity detection and visual pre-positioning based on improved YOLOv4[J]. *Electronics*, 2022, 11(17): ID 2677.
- [39] 肖张娜, 罗陆锋, 陈明猷, 等. 基于改进YOLO-v4的果园环境下葡萄检测[J]. *智能化农业装备学报(中英文)*, 2023, 4(2): 35-43.  
XIAO Z N, LUO L F, CHEN M Y, et al. Detection of grapes in orchard environment based on improved YOLOv4[J]. *Journal of intelligent agricultural mechanization*, 2023, 4(2): 35-43.
- [40] MAJEED Y, ZHANG J, ZHANG X, et al. Apple tree trunk and branch segmentation for automatic trellis training using convolutional neural network based semantic segmentation[J]. *IFAC-PapersOnLine*, 2018, 51(17): 75-80.
- [41] PENG H X, CHEN H, ZHANG X, et al. Retinnet\_G2S: A multi-scale feature fusion-based network for fruit detection of punna navel oranges in complex field environments[J]. *Precision agriculture*, 2024, 25(2): 889-913.
- [42] TAO Z Q, LI K, RAO Y, et al. Strawberry maturity recognition based on improved YOLOv5[J]. *Agronomy*, 2024, 14(3): ID 460.
- [43] GUO C M, ZHU C H, LIU Y C, et al. End-to-End lightweight Transformer-Based neural network for grasp detection towards fruit robotic handling[J]. *Computers and electronics in agriculture*, 2024, 221: ID 109014.
- [44] GE Y Y, XIONG Y, TENORIO G L, et al. Fruit localization and environment perception for strawberry harvesting robots[J]. *IEEE access*, 2019, 7: 147642-147652.
- [45] ZHANG T, HUANG Z, YOU W, et al. An autonomous fruit and vegetable harvester with a low-cost gripper using a 3D sensor[J]. *Sensors (basel, Switzerland)*, 2019, 20(1): ID E93.
- [46] FANG C Y, CHEN H B, LI L, et al. A novel Adaptive Zone-fusion network for precise waxberry semantic segmentation to improve automated-harvesting in a complex orchard environment[J]. *Computers and electronics in agriculture*, 2024, 221: ID 108937.
- [47] BARTH R, HEMMING J, VAN HENTEN E J. Angle estimation between plant parts for grasp optimisation in harvest robots[J]. *Biosystems engineering*, 2019, 183: 26-46.
- [48] GENÉ-MOLA J, SANZ-CORTIELLA R, ROSELL-POLO J R, et al. Fruit detection and 3D location using instance segmentation neural networks and structure-from-motion photogrammetry[J]. *Computers and electronics in agriculture*, 2020, 169: ID 105165.
- [49] ALTAHERI H, ALSULAIMAN M, MUHAMMAD G. Date fruit classification for robotic harvesting in a natural environment using deep learning[J]. *IEEE access*, 2019, 7: 117115-117133.
- [50] SALIM F, SAEED F, BASURRA S, et al. DenseNet-201 and Xception pre-trained deep learning models for fruit recognition[J]. *Electronics*, 2023, 12(14): ID 3132.
- [51] TANG Y, CHEN M, WANG C, et al. Recognition and localization methods for vision-based fruit picking robots: A review[J]. *Frontiers in plant science*, 2020, 11: ID 510.
- [52] LIN G C, TANG Y C, ZOU X J, et al. In-field citrus detection and localisation based on RGB-D image analysis[J]. *Biosystems engineering*, 2019, 186: 34-44.
- [53] BARNEA E, MAIRON R, BEN-SHAHAR O. Colour-agnostic shape-based 3D fruit detection for crop harvesting robots[J]. *Biosystems engineering*, 2016, 146: 57-70.
- [54] LIN G C, TANG Y C, ZOU X J, et al. Color-, depth-, and shape-based 3D fruit detection[J]. *Precision agriculture*, 2020, 21(1): 1-17.
- [55] WANG Y W, CHEN Y F. Fruit morphological measurement based on three-dimensional reconstruction[J]. *Agronomy*, 2020, 10(4): ID 455.
- [56] NGUYEN T T, VANDEVOORDE K, WOUTERS N, et al. Detection of red and bicoloured apples on tree with an RGB-D camera[J]. *Biosystems engineering*, 2016, 146: 33-44.
- [57] MEHTA S S, BURKS T F. Vision-based control of robotic manipulator for citrus harvesting[J]. *Computers and electronics in agriculture*, 2014, 102: 146-158.
- [58] KACZMAREK A L. Stereo vision with Equal Baseline Multiple Camera Set (EBMCS) for obtaining depth maps of plants[J]. *Computers and electronics in agriculture*, 2017, 135: 23-37.
- [59] LIU H J, LEE S H, CHAHL J S. Registration of multi-spectral 3D points for plant inspection[J]. *Precision agriculture*



- culture, 2018, 19(3): 513-536.
- [60] NGUYEN T T, SLAUGHTER D C., TOWNSLEY B T., et al. In-field plant phenotyping using multi-view reconstruction: An investigation in eggplant[C]// International Conference on Precision Agriculture (ICPA). St. Louis, Missouri, USA: International Society of Precision Agriculture, 2016.
- [61] LATIF R, SADDIK A. SLAM algorithms implementation in a UAV, based on a heterogeneous system: A survey[C]// 2019 4th World Conference on Complex Systems (WCCS). Piscataway, New Jersey, USA: IEEE, 2019.
- [62] CHEN X, LU H M, XIAO J H, et al. Distributed monocular multi-robot SLAM[C]// 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER). Piscataway, New Jersey, USA: IEEE, 2018: 73-78.
- [63] 吴皓, 迟金鑫, 田国会. 基于视觉SLAM的物体实例识别与语义地图构建[J]. 华中科技大学学报(自然科学版), 2019, 47(9): 48-54.  
WU H, CHI J X, TIAN G H. Instance recognition and semantic mapping based on visual SLAM[J]. Journal of Huazhong university of science and technology (natural science edition), 2019, 47(9): 48-54.
- [64] KUSUMAM K, KRAJNÍK T, PEARSON S, et al. 3D-vision based detection, localization, and sizing of broccoli heads in the field[J]. Journal of field robotics, 2017, 34(8): 1505-1518.
- [65] NELLITHIMARU A K, KANTOR G A. ROLS: Robust object-level SLAM for grape counting[C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway, New Jersey, USA: IEEE, 2019.
- [66] MATSUZAKI S, MASUZAWA H, MIURA J, et al. 3D semantic mapping in greenhouses for agricultural mobile robots with robust object recognition using robots' trajectory[C]// 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Piscataway, New Jersey, USA: IEEE, 2018: 357-362.
- [67] DONG W B, ROY P, ISLER V. Semantic mapping for orchard environments by merging two-sides reconstructions of tree rows[J]. Journal of field robotics, 2020, 37(1): 97-121.
- [68] LIU T H, KANG H W, CHEN C. ORB-Livox: A real-time dynamic system for fruit detection and localization[J]. Computers and electronics in agriculture, 2023, 209: ID 107834.
- [69] ZHANG X G, ZHANG R D, WANG X K. Visual SLAM mapping based on YOLOv5 in dynamic scenes[J]. Applied sciences, 2022, 12(22): ID 11548.
- [70] GAO X Y, LI J H, FAN L F, et al. Review of wheeled mobile robots' navigation problems and application prospects in agriculture[J]. IEEE access, 2018, 6: 49248-49268.
- [71] CHEBROLU N, LOTTES P, SCHAEFER A, et al. Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields[J]. The international journal of robotics research, 2017, 36(10): 1045-1052.
- [72] AGUIAR A S, DOS SANTOS F N, CUNHA J B, et al. Localization and mapping for robots in agriculture and forestry: A survey[J]. Robotics, 2020, 9(4): ID 97.
- [73] ZHAO W, WANG X, QI B Z, et al. Ground-level mapping and navigating for agriculture based on IoT and computer vision[J]. IEEE access, 2020, 8: 221975-221985.
- [74] CAPUA F R, SANSONI S, MOREYRA M L. Comparative analysis of visual-SLAM algorithms applied to fruit environments[C]// 2018 Argentine Conference on Automatic Control (AADECA). Piscataway, New Jersey: IEEE, 2018.
- [75] SHALAL N, LOW T, MCCARTHY C, et al. Orchard mapping and mobile robot localisation using on-board camera and laser scanner data fusion - Part B: Mapping and localisation[J]. Computers and electronics in agriculture, 2015, 119: 267-278.
- [76] UNDERWOOD J P, HUNG C, WHELAN B, et al. Mapping almond orchard canopy volume, flowers, fruit and yield using lidar and vision sensors[J]. Computers and electronics in agriculture, 2016, 130: 83-96.
- [77] HABIBIE N, NUGRAHA A M, ANSHORI A Z, et al. Fruit mapping mobile robot on simulated agricultural area in Gazebo simulator using simultaneous localization and mapping (SLAM)[C]// 2017 International Symposium on Micro-NanoMechatronics and Human Science (MHS). Piscataway, New Jersey, USA: IEEE, 2017.
- [78] CHEN S B, ZHOU B D, JIANG C H, et al. A LiDAR/visual SLAM backend with loop closure detection and graph optimization[J]. Remote sensing, 2021, 13(14): ID 2720.
- [79] SUN Y, WANG Q, YAN C, et al. D-VINS: Dynamic adaptive visual-inertial SLAM with IMU prior and semantic constraints in dynamic scenes[J]. Remote sensing, 2023, 15(15): ID 3881.
- [80] LI D D, ZHANG F B, FENG J X, et al. LD-SLAM: A robust and accurate GNSS-aided multi-map method for long-distance visual SLAM[J]. Remote sensing, 2023, 15(18): ID 4442.
- [81] GAN H, LEE W S, ALCHANATIS V. A prototype of an immature citrus fruit yield mapping system[C]// 2017 ASABE International Meeting. St. Joseph, Missouri, USA: American Society of Agricultural and Biological Engineers, 2017.
- [82] HUA T, PEI L, LI T, et al. M2C-GVIO: Motion manifold constraint aided GNSS-visual-inertial odometry for ground vehicles[J]. Satellite navigation, 2023, 4(1): ID 13.
- [83] ZHU J, ZHOU H, WANG Z Y, et al. Improved multi-sensor fusion positioning system based on GNSS/LiDAR/Vision/IMU with semi-tight coupling and graph optimization in GNSS challenging environments[J]. IEEE Access, 2023, 11: 95711-95723.
- [84] SCHUSTER M J, SCHMID K, BRAND C, et al. Distributed stereo vision-based 6D localization and mapping for multi-robot teams[J]. Journal of field robotics, 2019, 36(2): 305-332.
- [85] 傅博, 焦艳梅, 丁夏清, 等. 一种鲁棒的多目视觉惯性即时定位与建图方法[J]. 载人航天, 2019, 25(5): 581-585.  
FU B, JIAO Y M, DING X Q, et al. A robust multi-camera visual-inertial simultaneous localization and mapping method[J]. Manned spaceflight, 2019, 25(5): 581-585.
- [86] LEE T J, KIM C H, CHO D I D. A monocular vision sensor-based efficient SLAM method for indoor service robots[J]. IEEE transactions on industrial electronics, 2019, 66(1): 318-328.

- [87] SHEIKH T S, AFANASYEV I M. Stereo vision-based optimal path planning with stochastic maps for mobile robot navigation[M]// *Intelligent Autonomous Systems 15*. Cham: Springer International Publishing, 2018: 40-55.
- [88] CHENG Y, BAI J Q, XIU C B. Improved RGB-D vision SLAM algorithm for mobile robot[C]// *2017 29th Chinese Control and Decision Conference (CCDC)*. Piscataway, New Jersey, USA: IEEE, 2017: 5419-5423.
- [89] RAVANKAR A A, RAVANKAR A, EMARU T, et al. A hybrid topological mapping and navigation method for large area robot mapping[C]// *2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. Piscataway, New Jersey, USA: IEEE, 2017: 1104-1107.
- [90] SHWE L L T, WIN W. Vision-based mobile robot self-localization and mapping system for indoor environment[J]. *American academic scientific research journal for engineering, technology, and sciences*, 2017, 38(1): 306-324.
- [91] CHEN X, ZHANG H, LU H M, et al. Robust SLAM system based on monocular vision and LiDAR for robotic urban search and rescue[C]// *2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*. Piscataway, New Jersey, USA: IEEE, 2017: 41-47.
- [92] KALOGEITON V S, IOANNIDIS K, SIRAKOULIS G C, et al. Real-time active SLAM and obstacle avoidance for an autonomous robot based on stereo vision[J]. *Cybernetics and systems*, 2019, 50(3): 239-260.
- [93] SRINIVASAN RAMANAGOPAL M, NGUYEN A P V, LE NY J. A motion planning strategy for the active vision-based mapping of ground-level structures[J]. *IEEE transactions on automation science and engineering*, 2018, 15(1): 356-368.
- [94] SUJIWO A, UNIVERSITY N, ANDO T, et al. Monocular vision-based localization using ORB-SLAM with LIDAR-aided mapping in real-world robot challenge[J]. *Journal of robotics and mechatronics*, 2016, 28(4): 479-490.
- [95] USLU E, ÇAKMAK F, ALTUNTAŞ N, et al. An architecture for multi-robot localization and mapping in the Gazebo/Robot Operating System simulation environment[J]. *Simulation*, 2017, 93(9): 771-780.
- [96] YU Y J, SUN Z P, ZHAO X G, et al. Design and implementation of an automatic peach-harvesting robot system[C]// *2018 Tenth International Conference on Advanced Computational Intelligence (ICACI)*. Piscataway, New Jersey, USA: IEEE, 2018: 700-705.
- [97] CHEN X Y, WANG S A, ZHANG B Q, et al. Multi-feature fusion tree trunk detection and orchard mobile robot localization using camera/ultrasonic sensors[J]. *Computers and electronics in agriculture*, 2018, 147: 91-108.
- [98] MASUZAWA H, MIURA J, OISHI S. Development of a mobile robot for harvest support in greenhouse horticulture: Person following and mapping[C]// *2017 IEEE/SICE International Symposium on System Integration (SII)*. Piscataway, New Jersey, USA: IEEE, 2017: 541-546.
- [99] 顾宝兴, 刘钦, 田光兆, 等. 基于改进YOLOv3的果树树干识别和定位[J]. *农业工程学报*, 2022, 38(6): 122-129.
- GU B X, LIU Q, TIAN G Z, et al. Recognizing and locating the trunk of a fruit tree using improved YOLOv3[J]. *Transactions of the Chinese society of agricultural engineering*, 2022, 38(6): 122-129.
- [100] 胡广锐, 孔微雨, 齐闯, 等. 果园环境下移动采摘机器人导航路径优化[J]. *农业工程学报*, 2021, 37(9): 175-184.
- HU G R, KONG W Y, QI C, et al. Optimization of the navigation path for a mobile harvesting robot in orchard environment[J]. *Transactions of the Chinese society of agricultural engineering*, 2021, 37(9): 175-184.
- [101] IVANOV M, SERGYIENKO O, TYRSA V, et al. Influence of data clouds fusion from 3D real-time vision system on robotic group dead reckoning in unknown terrain[J]. *CAA journal of automatica sinica*, 2020, 7(2): 368-385.
- [102] XU S L, RAI R. Vision-based autonomous navigation stack for tractors operating in peach orchards[J]. *Computers and electronics in agriculture*, 2024, 217: ID 108558.
- [103] ISLAM R, HABIBULLAH H, HOSSAIN T. AGRISLAM: A real-time stereo visual SLAM for agricultural environment[J]. *Autonomous robots*, 2023, 47(6): 649-668.
- [104] ZHANG X S, YAO M, CHENG Q, et al. A novel hand-eye calibration method of picking robot based on TOF camera[J]. *Frontiers in plant science*, 2022, 13: ID 1099033.
- [105] LEHNERT C, ENGLISH A, MCCOOL C, et al. Autonomous sweet pepper harvesting for protected cropping systems[J]. *IEEE robotics and automation letters*, 2017, 2(2): 872-879.
- [106] LEHNERT C F, MCCOOL C, PEREZ T. Lessons learnt from field trials of a robotic sweet pepper harvester[EB/OL]. arXiv: 1706.06203, 2017
- [107] LEHNERT C, SA I, MCCOOL C, et al. Sweet pepper pose detection and grasping for automated crop harvesting[C]// *2016 IEEE International Conference on Robotics and Automation (ICRA)*. Piscataway, New Jersey, USA: IEEE, 2016: 2428-2434.
- [108] RAMON SORIA P, SUKKAR F, MARTENS W, et al. Multi-view probabilistic segmentation of pome fruit with a low-cost RGB-D camera[M]// *ROBOT 2017: Third Iberian Robotics Conference*. Cham: Springer International Publishing, 2017: 320-331.
- [109] CHEN X Y, CHAUDHARY K, TANAKA Y, et al. Reasoning-based vision recognition for agricultural humanoid robot toward tomato harvesting[C]// *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Piscataway, New Jersey, USA: IEEE, 2015: 6487-6494.
- [110] BARTH R, HEMMING J, VAN HENTEN E J. Design of an eye-in-hand sensing and servo control framework for harvesting robotics in dense vegetation[J]. *Biosystems engineering*, 2016, 146: 71-84.
- [111] 熊春源, 熊俊涛, 杨振刚, 等. 基于深度强化学习的柑橘采摘机械臂路径规划方法[J]. *华南农业大学学报*, 2023, 44(3): 473-483.
- XIONG C Y, XIONG J T, YANG Z G, et al. Path planning method for citrus picking manipulator based on deep reinforcement learning[J]. *Journal of South China agricultural university*, 2023, 44(3): 473-483.
- [112] HEMMING J, BAC C, TUIJL B V, et al. A robot for harvesting sweet-pepper in greenhouses[C]// *Proceedings International Conference of Agricultural Engineering (ICFAE)*. Zurich, Switzerland: EurAgEng, 2014.



- [113] CHRISTOPH S, JULIAN P, JÖRG B, et al. A modular robot system for agricultural applications[C]// Proceedings International Conference of Agricultural Engineering (ICFAE). Zurich, Switzerland: EurAgEng, 2014.
- [114] LEE B S, ROSA U A. Development of a canopy volume reduction technique for easy assessment and harvesting of Valencia citrus fruits[J]. Transactions of the asabe, 2006, 49(6): 1695-1703.
- [115] LING P P, EHSANI R, TING K C, et al. Sensing and end-effector for a robotic tomato harvester[C]// 2004 ASAE annual meeting. St. Joseph, Missouri, USA: American Society of Agricultural and Biological Engineers, 2004.
- [116] HAYASHI S, GANNO K, ISHII Y, et al. Robotic harvesting system for eggplants[J]. Japan agricultural research quarterly, 2002, 36(3): 163-168.
- [117] HAYASHI S, SHIGEMATSU K, YAMAMOTO S, et al. Evaluation of a strawberry-harvesting robot in a field test[J]. Biosystems engineering, 2010, 105(2): 160-171.
- [118] KONDO N, MONTA M, ARIMA S. Strawberry harvesting robot on hydroponic system[J]. IFAC proceedings volumes, 1998, 31(5): 181-185.
- [119] TANIGAKI K, FUJIURA T, AKASE A, et al. Cherry-harvesting robot[J]. Computers and electronics in agriculture, 2008, 63(1): 65-72.
- [120] VAN HENTEN E J, SCHENK E J, VAN WILLIGENBURG L G, et al. Collision-free inverse kinematics of the redundant seven-link manipulator used in a cucumber picking robot[J]. Biosystems engineering, 2010, 106(2): 112-124.
- [121] SILWAL A, DAVIDSON J R, KARKEE M, et al. Design, integration, and field evaluation of a robotic apple harvester[J]. Journal of field robotics, 2017, 34(6): 1140-1159.
- [122] SILWAL A, DAVIDSON J, KARKEE M, et al. Effort towards robotic apple harvesting in Washington State[C]// 2016 ASABE International Meeting. St. Joseph, Missouri, USA: American Society of Agricultural and Biological Engineers, 2016.
- [123] FU M, GUO S, CHEN A, et al. Design and experimentation of multi-fruit envelope-cutting kiwifruit picking robot[J]. Frontiers in plant science, 2024, 15: ID 1338050.
- [124] WANG Y S, WU H Y, ZHU Z K, et al. Continuous picking of yellow peaches with recognition and collision-free path[J]. Computers and electronics in agriculture, 2023, 214: ID 108273.
- [125] GAO J, ZHANG F, ZHANG J X, et al. Picking patterns evaluation for cherry tomato robotic harvesting end-effector design[J]. Biosystems engineering, 2024, 239: 1-12.
- [126] BU L X, CHEN C K, HU G R, et al. Design and evaluation of a robotic apple harvester using optimized picking patterns[J]. Computers and electronics in agriculture, 2022, 198: ID 107092.
- [127] ZHANG H W, LI X G, WANG L, et al. Construction and optimization of a collaborative harvesting system for multiple robotic arms and an end-picker in a trellised pear orchard environment[J]. Agronomy, 2023, 14(1): ID 80.
- [128] FEI Z H, VOUGIOUKAS S G. A robotic orchard platform increases harvest throughput by controlling worker vertical positioning and platform speed[J]. Computers and electronics in agriculture, 2024, 218: ID 108735.
- [129] HU G R, CHEN C, CHEN J, et al. Simplified 4-DOF manipulator for rapid robotic apple harvesting[J]. Computers and electronics in agriculture, 2022, 199: ID 107177.
- [130] ROSHANIANFARD A, KAMATA T, NOGUCHI N. Performance evaluation of harvesting robot for heavy-weight crops[J]. IFAC-PapersOnLine, 2018, 51(17): 332-338.
- [131] DE PRETER A, ANTHONIS J, DE BAERDEMAEKER J. Development of a robot for harvesting strawberries[J]. IFAC-PapersOnLine, 2018, 51(17): 14-19.
- [132] ZHUANG M, LI G, DING K X, et al. Research on the application of impedance control in flexible grasp of picking robot[J]. Advances in mechanical engineering, 2023, 15(4): ID 168781322311610.
- [133] CHEN M Y, CHEN Z X, LUO L F, et al. Dynamic visual servo control methods for continuous operation of a fruit harvesting robot working throughout an orchard[J]. Computers and electronics in agriculture, 2024, 219: ID 108774.
- [134] LEE B, KAM D, MIN B, et al. A vision servo system for automated harvest of sweet pepper in Korean greenhouse environment[J]. Applied sciences, 2019, 9(12): ID 2395.
- [135] SUN T, ZHANG W, MIAO Z H, et al. Object localization methodology in occluded agricultural environments through deep learning and active sensing[J]. Computers and electronics in agriculture, 2023, 212: ID 108141.
- [136] ZEESHAN S, AIZED T, RIAZ F. In-depth evaluation of automated fruit harvesting in unstructured environment for improved robot design[J]. Machines, 2024, 12(3): ID 151.
- [137] WANG C L, LI C J, HAN Q Y, et al. A performance analysis of a litchi picking robot system for actively removing obstructions, using an artificial intelligence algorithm[J]. Agronomy, 2023, 13(11): ID 2795.
- [138] ONISHI Y, YOSHIDA T, KURITA H, et al. An automated fruit harvesting robot by using deep learning[J]. ROBO-MECH journal, 2019, 6(1): ID 13.
- [139] ALMENDRAL K A M, BABARAN R M G, CARZON B J C, et al. Autonomous fruit harvester with machine vision[J]. Journal of telecommunication, electronic and computer engineering, 2018, 10(1-6): 79-86.
- [140] WILLIAMS H A M, JONES M H, NEJATI M, et al. Robotic kiwifruit harvesting using machine vision, convolutional neural networks, and robotic arms[J]. Biosystems engineering, 2019, 181: 140-156.
- [141] LING X, ZHAO Y S, GONG L, et al. Dual-arm cooperation and implementing for robotic harvesting tomato using binocular vision[J]. Robotics and autonomous systems, 2019, 114: 134-143.
- [142] XIONG Y, GE Y Y, GRIMSTAD L, et al. An autonomous strawberry-harvesting robot: Design, development, integration, and field evaluation[J]. Journal of field robotics, 2020, 37(2): 202-224.
- [143] LI T, XIE F, ZHAO Z Q, et al. A multi-arm robot system for efficient apple harvesting: Perception, task plan and control[J]. Computers and electronics in agriculture, 2023, 211: ID 107979.
- [144] XIE F, SUN N, LI J H, et al. Fruit distribution acquisition with multi-vision for multi-arm harvesting robots[C]// 2023 8th International Conference on Control, Robotics and Cybernetics (CRC). Piscataway, New Jersey, USA: IEEE, 2024: 7-13.

# Orchard-Wide Visual Perception and Autonomous Operation of Fruit Picking Robots: A Review

CHEN Mingyou, LUO Lufeng\*, LIU Wei, WEI Huiling, WANG Jinhai,  
LU Qinghua, LUO Shaoming

(*College of Mechanical and Electrical Engineering, Foshan University, Foshan 528231, China*)

## Abstract:

**[Significance]** Fruit-picking robot stands as a crucial solution for achieving intelligent fruit harvesting. Significant progress has been made in developing foundational methods for picking robots, such as fruit recognition, orchard navigation, path planning for picking, and robotic arm control, the practical implementation of a seamless picking system that integrates sensing, movement, and picking capabilities still encounters substantial technical hurdles. In contrast to current picking systems, the next generation of fruit-picking robots aims to replicate the autonomous skills exhibited by human fruit pickers. This involves effectively performing ongoing tasks of perception, movement, and picking without human intervention. To tackle this challenge, this review delves into the latest research methodologies and real-world applications in this field, critically assesses the strengths and limitations of existing methods and categorizes the essential components of continuous operation into three sub-modules: local target recognition, global mapping, and operation planning.

**[Progress]** Initially, the review explores methods for recognizing nearby fruit and obstacle targets. These methods encompass four main approaches: low-level feature fusion, high-level feature learning, RGB-D information fusion, and multi-view information fusion, respectively. Each of these approaches incorporates advanced algorithms and sensor technologies for cluttered orchard environments. For example, low-level feature fusion utilizes basic attributes such as color, shapes and texture to distinguish fruits from backgrounds, while high-level feature learning employs more complex models like convolutional neural networks to interpret the contextual relationships within the data. RGB-D information fusion brings depth perception into the mix, allowing robots to gauge the distance to each fruit accurately. Multi-view information fusion tackles the issue of occlusions by combining data from multiple cameras and sensors around the robot, providing a more comprehensive view of the environment and enabling more reliable sensing. Subsequently, the review shifts focus to orchard mapping and scene comprehension on a broader scale. It points out that current mapping methods, while effective, still struggle with dynamic changes in the orchard, such as variations of fruits and light conditions. Improved adaptation techniques, possibly through machine learning models that can learn and adjust to different environmental conditions, are suggested as a way forward. Building upon the foundation of local and global perception, the review investigates strategies for planning and controlling autonomous behaviors. This includes not only the latest advancements in devising movement paths for robot mobility but also adaptive strategies that allow robots to react to unexpected obstacles or changes within the whole environment. Enhanced strategies for effective fruit picking using the Eye-in-Hand system involve the development of more dexterous robotic hands and improved algorithms for precisely predicting the optimal picking point of each fruit. The review also identifies a crucial need for further advancements in the dynamic behavior and autonomy of these technologies, emphasizing the importance of continuous learning and adaptive control systems to improve operational efficiency in diverse orchard environments.

**[Conclusions and Prospects]** The review underscores the critical importance of coordinating perception, movement, and picking modules to facilitate the transition from a basic functional prototype to a practical machine. Moreover, it emphasizes the necessity of enhancing the robustness and stability of core algorithms governing perception, planning, and control, while ensuring their seamless coordination which is a central challenge that emerges. Additionally, the review raises unresolved questions regarding the application of picking robots and outlines future trends, include deeper integration of stereo vision and deep learning, enhanced global vision sampling, and the establishment of standardized evaluation criteria for overall operational performance. The paper can provide references for the eventual development of robust, autonomous, and commercially viable picking robots in the future.

**Key words:** fruit picking robot; autonomous operation; local perception; global mapping; behavior planning

**Foundation items:** National Natural Science Foundation of China (32301704, 32171909); Natural Science Foundation of Guangdong Province (2024A1515010199, 2023A1515011255)

**Biography:** CHEN Mingyou, E-mail: chenmingyou1@126.com

**\*Corresponding author:** LUO Lufeng, E-mail: luolufeng617@163.com

(登录 [www.smartag.net.cn](http://www.smartag.net.cn) 免费获取电子版全文)