

# 基于辅助光投影的单目果实视觉定位方法

郭家祥, 马兆敏\*

(广西科技大学自动化学院, 广西柳州 545006)

**摘要:**【目的】针对果实采摘机器人中单目视觉定位操作简单但精度不足问题, 根据光线投影原理引入辅助光获取果实的地面投影, 提出了一种基于果实投影的单目视觉定位方法。【方法】该视觉定位系统包括两盏投影灯和一台相机。首先通过不同位置的两盏投影灯在地面形成交叉投影区域, 然后利用神经网络对地面交叉投影区域进行标定, 建立交叉投影区域内图像坐标系与世界坐标系间的坐标映射模型; 再将视觉定位系统放置到果实前, 在地面交叉投影区域产生果实投影, 并求出果实投影质心在图像中的坐标; 接着通过坐标映射模型获取果实投影的地面坐标; 最后结合投影灯和摄像头的位置信息在三维空间建立直线方程, 通过广义逆矩阵求方程的最小二乘解, 得到果实位置。为验证定位算法精度, 分别进行了果实位置测量和果实采摘的仿真实验。【结果】在果实位置测量中, 各个坐标轴平均误差在1cm以内; 在果实采摘仿真中, 成功率为93%。【结论】实验表明, 该定位方法的精度可以满足果实采摘的需求, 为果实采摘设备的开发提供了理论基础。

**关键词:** 投影; 视觉; 标定; 最小二乘解; 定位

## 引言

果实采摘作为农业生产的基础性环节, 占用了40%–50%的劳动力, 且目前依然以人工作业为主<sup>[1]</sup>, 费时费力<sup>[2]</sup>。由于果实重叠、晃动<sup>[3]</sup>、绿熟<sup>[4]</sup>、易损坏<sup>[5]</sup>、实时性<sup>[6]</sup>、外界光照环境的影响<sup>[7]</sup>等, 果实形态十分复杂, 不但形状各异、大小不同, 甚至有的需要对一整串进行定位, 导致对其进行自动化采摘十分困难。

通过机器人实现果实的自动化采摘是一个可行的技术路线<sup>[8]</sup>, 其中果实的定位问题是农业自动化采摘的关键步骤, 对自动化采摘的效率、成本都有着重要的影响。目前, 果实的定位环境已经从早期的温室大棚<sup>[9]</sup>逐渐扩展至野外果园环境, 定位目标从对静止、无遮挡的果实到晃动<sup>[10]</sup>、有遮挡<sup>[11]</sup>的果实, 从单个果实到黏结的果实, 从而增加了对果实位姿的定位要求。

目前, 果实采摘定位的主要感知方法是深度相机、双目视觉技术、单目视觉技术等<sup>[12]</sup>。以结构光和ToF相机为代表的深度相机技术在果实采摘定位上也

有较多应用, 但是由于在野外工作时, 强自然光会导致结构光相机投射的编码光被淹没<sup>[13–14]</sup>, 而ToF相机在野外工作时需要较大功率<sup>[15–16]</sup>, 因此深度相机难以在野外实际场景中应用<sup>[17–18]</sup>。双目视觉测量精度与标定精度强相关<sup>[19]</sup>, 但是装在车辆等运动平台上时, 由于野外工作环境的颠簸, 对机械结构的稳定性要求很高, 面临每隔一定时间就需要重新标定的问题<sup>[20]</sup>。同时双目视觉标定复杂, 深度信息的量程和精度受基线距离和相机分辨率限制<sup>[21–23]</sup>。且视差图计算量大, 导致实时性较低<sup>[24]</sup>。

单目相机结构简单、成本低、便于标定和识别<sup>[25]</sup>, 但是只能获取二维空间信息, 无法直接实现果实的三维定位<sup>[26]</sup>, 需要通过其他手段获取另外一个维度的空间信息。双目定位精度高于单目定位本质是由于双目使用两个摄像头获取果实的位置信息, 而单目摄像头需要其他辅助手段获取果实的深度信息, 有观点认为该方法优于单纯的单目视觉定位方法<sup>[27]</sup>。目前的辅助设备主要有运动机构<sup>[28]</sup>、反射镜镜组<sup>[29]</sup>、惯性传感器<sup>[30]</sup>、激光雷达<sup>[31]</sup>、激光测距云台<sup>[32]</sup>等。单目测量获取位置信息多是利用辅助设备的运动产生图像的差异, 所以对辅助设备的运动精度要求较高<sup>[33]</sup>。不但成本高昂, 且在野外环境使用时难以保持设备的精度。实际上, 根据光线投影原理, 可以通过果实的地面投影测量深度信息, 完成果实的定位。因此, 本文提出了一种基于果实投影的单目视觉定位方法。

## 1 投影定位原理

视觉定位就是由目标对象的已知图像的坐标点

**基金项目:** 中央引导地方科技发展资金资助(桂科ZY22096023)——丘陵山地经济作物农业机械化与智能化创新中心。

**作者简介:** 郭家祥, 研究生, 主要从事视觉测量与模式识别研究。Email: 2259898232@qq.com

**通讯作者:** 马兆敏, 副教授, 主要从事视觉测量与模式识别研究。Email: 314044282@qq.com

计算未知三维世界坐标点的方法。构建视觉定位系统可获得投影灯位置、果实投影等已知信息，每个果实投影中包含了二维平面的位置信息，两个投影灯之间又包含了一个维度的信息。因此通过两个投影灯的位置及果实投影的位置可以计算出现实空间中一点的坐标。

定位系统原理如图1所示，当光线照射在三维空间的物体中时，会在投影平面上形成对应的阴影。 $P_s$ 为投影灯光心位置， $P$ 为灯光照射的果实， $P_e$ 为果实在地面产生的投影。其中， $P_s$ 为固定点可提取测量得到， $P_e$ 可通过标定后的相机测得，为定位系统中的已知点。果实位置 $P$ 为定位系统需要测量的未知点。

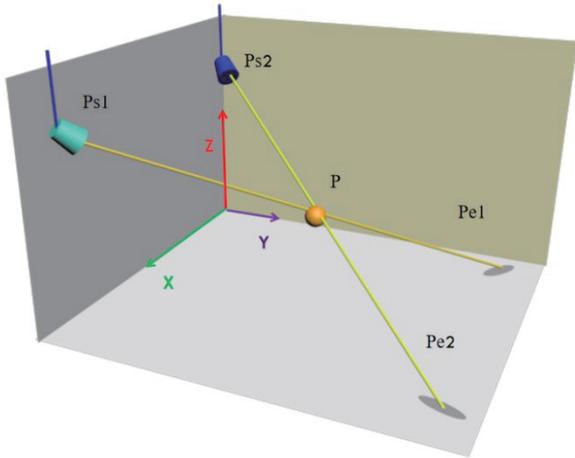


图1 世界坐标系投影定位原理图

在得到投影灯位置、果实投影位置等已知量时，可以凭借这些已知量构建定位方程。可知投影灯光源的中心 $P_s$ 与果实 $P$ 、果实投影 $P_e$ 在同一射线上。采用三维射线方程来描述这一投影过程，其中 $s$ 为变量， $\vec{u}$ 为射线的方向向量。在灯光对果实的投影过程中，已知两个光源的光心位置  $p_{s1} \leftrightarrow p_{s2}$  和各自对应的果实投影  $p_{e1} \leftrightarrow p_{e2}$ ，因此我们可以建立射线方程组。 $s$ 需大于等于0，确保射线只有一个延伸方向，由此可得 $l$ 的射线方程。

$$\begin{cases} p = p_{s1} + s \cdot \vec{u}_1 \\ p = p_{s2} + t \cdot \vec{u}_2 \quad (s > 0, t > 0) \\ \vec{u}_1 = p_{e1} - p_{s1} \\ \vec{u}_2 = p_{e2} - p_{s2} \end{cases} \quad (2)$$

求解两条射线的交点，由于光心和投影的位置存在一定的误差，两条直线一般不会相交，由此可知两条射线间的最近点为果实位置。射线方程组的最

小二乘解即为最近点，因此可用广义逆矩阵来求解  $[s \ t]^T$ 。

$$\begin{bmatrix} s \\ t \end{bmatrix} = (M^T M)^{-1} M^T B \quad (3)$$

其中：

$$M = [u_1^T \ -u_2^T], \quad B = [p_{s2}^T - p_{s1}^T] \quad (4)$$

得到 $[s \ t]^T$ 后，代入直线方程，分别得 $P_s$ 和 $P_t$ 。如果两点相同，则两直线有交点，为 $P_s$ 或 $P_t$ 。如果两点不同，则两条直线没有交点，求其最近点  $(P_s + P_t) / 2$ 。由此可得果实 $P$ 的位置。

## 2 投影定位方法

该视觉定位系统的搭建包括投影灯的布置、相机的布置、世界坐标系的确定等步骤。在搭建好系统后，通过相机标定、采集果实投影图像，计算果实位置。



图2 定位流程图

### 2.1 搭建视觉定位系统

在侧面放置两台投影灯，使得灯光交叉在地面形成两个投影，形成交叉投影区域。在棋盘格前方放置一台摄像头，使得相机可以拍摄到交叉投影区域。布置完成后，测量灯光位置可知，白色投影灯镜头位置为 $[-1221, 543, 848]$ ，蓝色投影灯镜头位置为 $[-980, -14, 800]$ ，单位mm。

### 2.2 神经网络标定

利用神经网络对地面交叉投影区域进行标定，建立交叉投影区域内图像坐标系与世界坐标系间的坐标映射模型。在交叉投影区域放置一块单位长度为10cm的10\*10的棋盘格标定板，如图3所示，以标定

板左下角角点为世界坐标系原点，以横轴为X轴，纵轴为Y轴。提取棋盘格角点的图像坐标，以图像坐标为输入，世界坐标为输出，训练神经网络。计算图像中角点神经网络预测值与实际值间的误差，可知X轴和Y轴的误差都在0.5mm以下，可以满足定位要求。

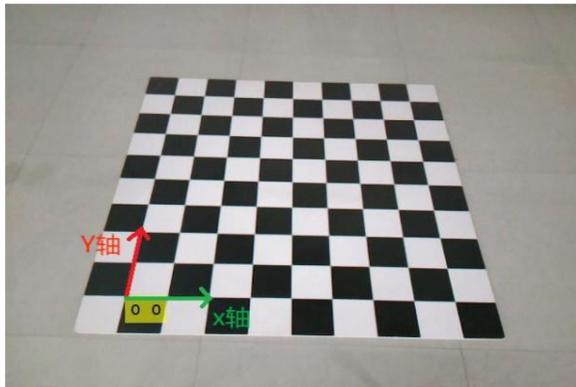


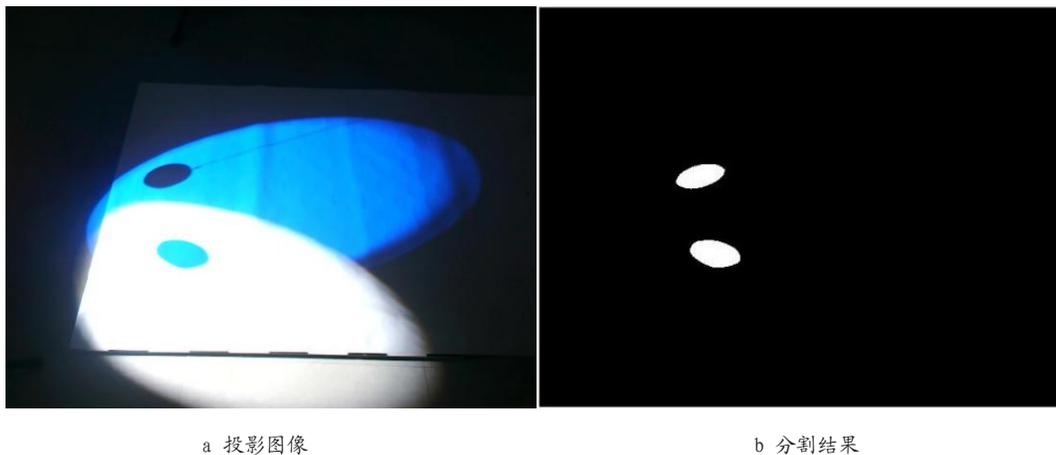
图3 标定场景图

### 2.3 果实投影图像处理

首先用白布遮盖标定板，拍摄一张无果实投影的图像作为背景图像。随后在侧面放置一个垂吊着一个果实的三脚架，两台投影灯使得灯光打在果实上，使得圆球正好投影在交叉投影区域上形成两个果实投影，拍摄作为前景图像。使用投影图像减去背景图像分割出果实投影，去除小面积区域。放置果实前后两幅图像相减，获得如图4所示图像。以图4为例，用两质心代表投影位置，对两个投影区域提取质心[284, 172]和[345, 377]。

### 2.4 果实位置测量

输入神经网络，得到果实投影的世界坐标[301, 580, 0]和[460, -16, 0]。结合前文测得的光心位置，可得投影灯与果实投影的方向向量 $u_1(1281, 595, -800)$ 、 $u_2(1681, -559, -848)$ 。根据投影



a 投影图像

b 分割结果

图4 投影实验图

的方向向量和投影灯位置构建射线定位方程，可得向量矩阵M和位置向量B，通过广义逆矩阵计算定位方程的参数。将参数代入定位方程，可得两射线最近点，即果实三维坐标为[-373, 261, 420]。

移动果实位置，视觉定位系统其他部分保持不变，使其重新在地面上产生投影，再次使用相机测量得到果实投影的世界坐标，结合投影灯位置构建定位方程，对定位方程求最小二乘解得到果实位置，使用激光测距仪和直尺测量果实的真实位置，完成一次新的定位实验。上述步骤重复多次，得到足够的实验数据用于后续分析。

## 3 实验结果与分析

由于机械臂在进行果实采摘时会受到遮挡物碰撞等因素而干扰采摘的成功率，因此本文实验分为两次，分别为位置测量实验和采摘实验。

### 3.1 位置测量实验

测量果实实际位置，使用激光测距仪和直尺对果实实际位置进行测量。将激光测距仪放置在地面上移动，使其激光点恰好打在果实底部中心位置。使用激光测距仪多次测量此时果实的高度，计算平均值，得到果实的Z轴坐标位置。标记激光测距仪所处位置，使用直尺等测量工具测量该标记在地面的位置，

得到果实的XY轴坐标。由此得到果实的实际坐标。

根据本文提出的基于辅助光投影的单目果实视觉定位方法测量果实的实际位置。比较本文方法定位结果与实际位置，可绘制如图5所示的定位结果图。实际坐标和本文算法计算的坐标各个坐标轴平均绝对误差分别为：X轴6.87mm，Y轴8.07mm，Z轴4.97mm，可以满足果实定位需求。

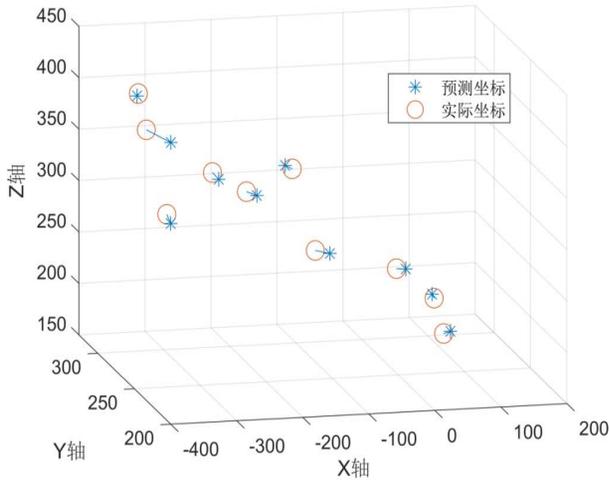


图5 果实定位结果图

### 3.2 机械臂抓取实验

定位精度对机械臂自动化采摘的结果有决定性影响，不够准确的定位结果将导致机械臂无法准确地抵达目标位置。但是，采摘的成功与否也受果实环境的遮挡、机械臂精度等因素的影响。因此，为了验证本文算法的实际定位效果，搭建了一套采摘系统对果实进行采摘实验。采摘系统主要由相机、投影灯、上位机和机械臂组成。采摘机器人系统由上位机处理相机拍摄的图像，根据图像投影点计算出果实的世界坐标。将机械臂放置在此前确定的世界坐标系原点位置，使机械臂基坐标系与世界坐标重合，这样通过投影点确定的世界坐标便是机械臂要抓取的基坐标。根据逆运动学方程计算机械臂的转动角度，上位机控制机械臂抓取果实。

实验时将果实放在交叉投影区域，在地面形成两个投影，使用相机测量出果实投影的世界坐标，结合投影灯位置，可按本文方法计算出果实位置，控制机械臂抓取。本次实验对果实进行了30次抓取，成功抓取果实28次，成功率93%。验证了本文算法的有效性。

## 4 讨论

### 4.1 误差分析

分析果实各个坐标轴上的定位误差可知，误差呈现逐渐增大的趋势。可对其进行线性回归估计定位范围。

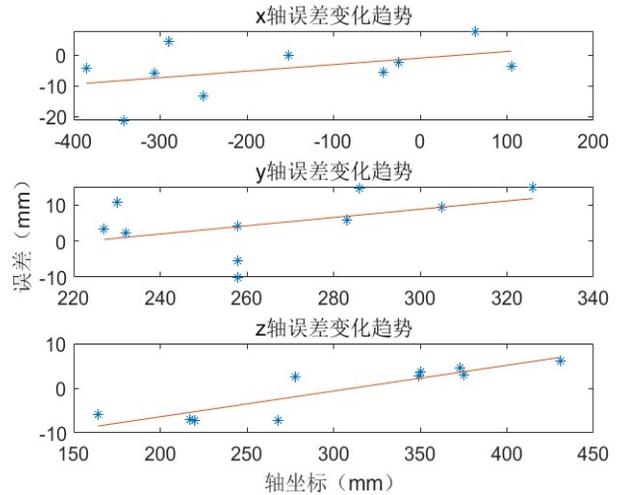


图6 各轴定位误差图

$$\Delta p = \begin{bmatrix} 0.0211 & & \\ & 0.115 & \\ & & 0.0577 \end{bmatrix} p + \begin{bmatrix} -1.0272 \\ -25.67 \\ -17.938 \end{bmatrix} \quad (5)$$

定位要求误差在2cm以内才算定位成功。根据公式计算误差变化情况，可知定位区域在 $\{p(x, y, z) | x \in (-899, 997) \& y \in (49, 397) \& z \in (-36, 658)\}$ 所示范围内，误差期望在2cm以下，该范围内可以满足定位要求。

### 4.2 误差来源

在计算过程中误差的产生不可避免，但精确度是相对的，误差产生的原因有以下四种，

(1) 模型误差，世界坐标与图像坐标的映射模型的计算结果与实际结果之间出现的误差。分析该误差发现，其误差范围在5mm以内，对果实的大小而言，已经具备足够的精度。

(2) 测量误差，测量投影灯光心的位置时产生的误差。保证精度主要依靠事先确定。事先确定了激光测距仪的精度为10米内误差为2mm。直尺的最小单位为1mm。

(3) 截断误差，由于本方法的方程组不能求出问题的解析解，只能使用数值计算方法求解，数值解与解析解之间存在一定的误差。

(4) 椭圆偏心误差, 圆形目标的投影多为椭圆, 投影过程中将产生偏心误差, 椭圆形变导致椭圆质心位置发生偏移。偏心误差与相机拍摄方位有关, 当投影灯倾角大于 $60^\circ$ 时, 偏心误差会快速增大, 因此在布置投影灯时需要控制投影灯投影角度或使用算法消除偏心误差。

为了避免上述误差对定位结果造成较大的影响, 两台投影灯到投影的射线方向距离不能相近, 且两台投影灯彼此之间的距离不能相近。这是由于当两条直线几乎平行时, 由误差导致的直线斜率的微小变化, 会导致交点位置的巨大变化的特点, 从而造成定位错误。所以为了确保定位成功, 投影灯的间距要根据距离果实的位置而调整。

### 4.3 定位范围

投影灯的投影定位范围由投影灯的投射比和投影角度决定。普通投影机的投射比通常在1.5-1.9之间。当投射比小于1时, 我们通常称之为短焦镜头; 当投射比小于0.6时, 我们通常称之为超短焦镜头。只有位于两台投影灯光路相交的区域的果实才能在地面产生两个投影, 从而构建三维射线方程进行定位。由于投影灯的投射比不同, 因此必然存在部分位置不会产生投影的情况, 导致投影定位出现盲区。在相交区域内的任意一点都能在地面产生两个对应的投影点, 而定位盲区内的任意一点都最多只有一个投影点, 无法以本文方法定位。因此在进行定位时, 可以选择不同投射比的投影灯进行投影。

### 5 结论

根据两次实验可知, 本文提出的定位方法具有较高的准确性, 可以满足果实采摘的精度要求, 且在指示机械臂采摘果实时也能拥有较高的采摘成功率。该方法相较于传统的双目相机等方法不需要复杂的标定工作且机械结构稳定性要求不高, 相较于深度相机等果实定位方法不需要较大功率且成本更低, 相较于使用运动机构定位的单目视觉定位系统不需要精密的运动机构。本视觉定位系统具有系统定位精度较高、结构简单、成本较低的优点。为未来果实自动化采摘设备的开发提供了理论基础。

### 参考文献

[1] 郑刚, 刘佳, 李旭. 现代温室采摘机器人发展概况[J]. 农业工程技术, 2019, 39(31):35-40.

- [2] Feng Q C, Zou W, Fan P F, 等. 樱桃番茄采摘机器人设计与试验[J]. 中国农业文摘-农业工程, 2018, 30(5):81.
- [3] 李国利, 姬长英, 顾宝兴. 基于单目视觉与超声检测的振荡果实采摘识别与定位[J]. 农业机械学报, 2015, 46(11):1-8. [知网]
- [4] 柳长源, 赖楠旭, 毕晓君. 基于深度图像的球形果实识别定位算法[J]. 农业机械学报, 2022, 53(10):228-235.
- [5] 张勤, 陈建敏, 李彬, 等. 基于RGB-D信息融合和目标检测的番茄串采摘点识别定位方法[J]. 农业工程学报, 2021, 37(18):143-152.
- [6] 刘德儿, 朱磊, 冀炜臻, 等. 基于RGB-D相机的脐橙实时识别定位与分级方法[J]. 农业工程学报, 2022, 38(14):154-165.
- [7] 曹春卿, 张吴平, 李富忠, 等. 自然场景下多目标苹果识别定位融合算法研究[J]. 湖北农业科学, 2022, 61(7):145-151.
- [8] VALVERDE J. Harvesting and processing of mushrooms[J]. Edible and Medicinal Mushrooms: Technology and Applications, 2017, 9(1):261-270.
- [9] 王海楠, 弋景刚, 张秀花. 番茄采摘机器人识别与定位技术研究进展[J]. 中国农机化学报, 2020, 41(5):188-196.
- [10] 李敏, 冯亚丽, 吴东林. 采摘机器人动态果实目标检测与跟踪技术研究——基于云存储[J]. 农机化研究, 2020, 42(9):207-211.
- [11] 李寒, 陶涵虢, 崔立昊, 等. 基于SOM-K-means算法的番茄果实识别与定位方法[J]. 农业机械学报, 2021, 52(1):23-29.
- [12] 郑太雄, 江明哲, 冯明驰. 基于视觉的采摘机器人目标识别与定位方法研究综述[J]. 仪器仪表学报, 2021, 42(9):28-51.
- [13] 刘志海, 代振锐, 田绍鲁, 等. 非接触式三维重建技术综述[J]. 科学技术与工程, 2022, 22(23):9897-9908.
- [14] 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, 209:107834.
- [15] 段争光. TOF成像系统光功率分析及仿真[J]. 西南民族大学学报(自然科学版), 2020, 46(6):643-647.
- [16] 郭宁博, 陈向宁, 薛俊诗. 基于飞行时间法的红外相机研究综述[J]. 兵器装备工程学报, 2017, 38(3):152-159.
- [17] 刘玉轩, 张力, 艾海滨, 等. 光场相机三维重建研究进展与展望[J]. 电子学报, 2022, 50(7):1774-1792.
- [18] Chu P Y, Li Z J, Zhang K X, et al. High-precision fruit localization using active laser-camera scanning:robust laser line extraction for 2D-3D transformation[J]. Smart Agricultural Technology, 2024, 7:100391.
- [19] Real-Moreno O, Rodriguez-Quiñonez J C, Flores-Fuentes W, et al. Camera calibration method through multivariate quadratic regression for depth estimation on a stereo vision system[J]. Optics and Lasers in Engineering, 2024, 174:107932.
- [20] 郑太雄, 黄帅, 李永福, 等. 基于视觉的三维重建关键技术研究综述[J]. 自动化学报, 2020, 46(4):631-652.
- [21] Yang S B, Gao Y, Liu Z, et al. A calibration method for binocular stereo vision sensor with short-baseline based on 3D flexible control field[J]. Optics and Lasers in Engineering, 2020, 124:105817.
- [22] 吕耀文, 徐熙平. 双目视觉测量系统结构参数设计[J]. 光电子技术, 2016, 36(2):98-102.
- [23] 吴彰良, 孙长库, 杨中东. 视觉测量传感器系统结构参数设计分析[J]. 光电工程, 2009, 36(12):56-61+90.
- [24] ZHANG Z Q, LI S C, LI C Y, et al. Navigation Path Detection Method for a Banana Orchard Inspection Robot Based on Binocular Vision[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(21):9-15.
- [25] 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.
- [26] Zhou H Y, Wang X, Au W, et al. Intelligent robots for fruit harvesting: recent developments and future challenges[J]. Precision Agriculture, 2022, 23(5):1856-1907.
- [27] Weiss U, Biber P. Plant detection and mapping for agricultural robots using a 3D LIDAR sensor[J]. Robotics and Autonomous Systems, 2011, 59(5):265-273.
- [28] 王粮局, 张立博, 段运红, 等. 基于视觉伺服的草莓采摘机器人果实定位方法[J]. 农业工程学报, 2015, 31(22):25-31.
- [29] 李桂华, 李涛, 孙卫庆, 等. 基于双平面镜双相机组合测量全场变形的的方法[J]. 光学学报, 2023, 43(2):132-140.
- [30] 王丹丹, 石峰, 翟亚芳, 等. 基于UKF的苹果果实定位估计算法[J]. 昆明理工大学学报(自然科学版), 2020, 45(4):50-56.
- [31] Kang H W, Wang X, Chen C. Accurate fruit localisation using high resolution LiDAR-camera fusion and instance segmentation[J]. Computers and Electronics in Agriculture, 2022, 203:107450.
- [32] 张凯良, 杨丽, 张铁中. 草莓采摘位置机器视觉与激光辅助定位方法[J]. 农业机械学报, 2010, 41(4):151-156.
- [33] 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.