

基于多模态感知的立体农业智能采摘与环境监测系统

刘渝博¹, 刘元林^{1,2*}, 唐健超¹, 张皓月¹, 黄巧薇¹, 曹燕玲¹, 金扬印³

1. 嘉应学院 物理与电子工程学院, 广东 梅州 514015

2. 广东省丘陵农业智能装备工程技术研究中心, 广东 梅州 514015

3. 梅州市鑫农机械制造有限公司, 广东 梅州 514015

DOI: 10.61369/TACS.2026070032

摘 要 : 针对立体农业采摘效率低、监测不精准、小型化部署成本高等问题, 设计一套轻量化嵌入式智能采摘与环境监测系统。以 K230 视觉模块与 STM32 为核心, 采用 YOLOv5 模型实现果实识别, 结合机械臂抓取与二维云台完成无损分拣, 并通过 LoRa 与 4G 实现环境数据远程传输。实验表明, 果实识别置信度 ≥ 0.85 , 抓取成功率 93%, 损伤率 $\leq 2.5\%$, 响应延迟 $< 200\text{ms}$, 系统具备低成本、轻量化、高可靠性特点, 可为丘陵立体农业智能化提供实用方案。

关 键 词 : 多模态感知; K230; 智能采摘; 立体农业; 环境监测

A Stereoscopic Agricultural Intelligent Picking and Environmental Monitoring System Based on Multimodal Perception

Liu Wengbo¹, Liu Yuanlin^{1,2*}, Tang Jianchao¹, Zhang Haoyue¹, Huang Qiaowei¹, Cao Yanling¹, Jin Yangyin³

1. School of Physics and Electronic Engineering, Jia Ying University, Meizhou, Guangdong 514015

2. Guangdong Province Engineering Technology Research Center for Intelligent Agricultural Equipment in Hilly Areas, Meizhou, Guangdong 514015

3. Meizhou Xin-Nong Machinery Manufacturing Co., Ltd, Meizhou, Guangdong 514015

Abstract : To address low efficiency, inaccurate monitoring, and high costs in three-dimensional agriculture, a lightweight embedded intelligent picking and environmental monitoring system is designed. The system uses a K230 vision module and STM32, adopts YOLOv5 for fruit recognition, a robotic arm and 2D pan-tilt for non-destructive sorting, and LoRa with 4G for remote data transmission. Experiments show fruit recognition confidence ≥ 0.85 , grasping success rate 93%, damage rate $\leq 2.5\%$, and response delay $< 200\text{ms}$. The system is low-cost, lightweight, and reliable, offering a practical solution for hilly three-dimensional agriculture.

Keywords : multimodal perception; K230; intelligent picking; three-dimensional agriculture; environmental monitoring

引言

立体农业在我国南方丘陵地区应用广泛, 但存在采摘效率低、人工成本高、作业空间受限等问题^[1]。传统采摘机器人多采用工控机 + 机械臂方案, 成本高、体积大, 难以在中小型果园普及。

近年来, 嵌入式视觉、运动控制和物联网这些技术发展挺快。YOLO 系列轻量模型在 K230、K210 这种边缘模块上部署得越来越顺^[2-4]; STM32 拿来搞舵机和电机的实时闭环控制也很常见; 柔性夹持、RGB-D 定位、LoRa 通信这些技术都给智能采摘打了底^[5-6]。

项目信息:

2025 年嘉应学院大学生创新训练计划省级项目 (S202510582010);

2021 年广东省线下一流本科课程 - 精密机械与仪器设计 (粤教高函 [2022] 10 号, 序号: 289);

2022 年度广东省课程思政改革示范项目 - 精密机械与仪器设计 (粤教高函 [2023] 14 号, 序号: 60);

2023 年度广东省教育科学规划课题 (高等教育专项: 2023GXJK490);

2023 年度广东省本科高校教学质量与教学改革工程项目 (粤教高函 [2024] 9 号, 序号: 108);

2024 年度广东省本科高校教学质量与教学改革工程项目 (粤教高函 [2024] 9 号, 序号: 60; 序号: 347);

广东省本科高校教学质量与教学改革工程项目 (粤教高函 [2026] 4 号, 序号: 288);

广东省丘陵农业智能装备工程技术研究中心 (2025GCZX010);

嘉应学院教学质量与教学改革工程项目 (ZLGC2023301, ZLGC2024404);

嘉应学院人才科研启动项目 (2024RC18);

2024 年度校企合作共建课程 (嘉院教 [2024] 39 号, 序号: 16)。

作者简介: 刘渝博 (2005—), 男, 本科在读, 学生, 研究方向为三维建模和机器视觉。

* 通讯作者: 刘元林 (1970—), 男, 硕士, 教授, 研究方向为机械结构损害无损检测技术、机械设计及理论。

不过现有系统还是有点毛病，比如遮挡环境下识别精度不够，嵌入式协同没做太完善，还有低成本采摘分拣机构挺缺的。

针对这些，本文搞了一套基于多模态感知的立体农业智能采摘与环境监测系统。核心是 K230 视觉加 STM32 控制，配上机械臂无损抓取和二维云台分拣，实现果实精准识别、自动采摘和分类投放；同时把环境监测和远程操控也集成进去。这套系统成本低、部署也方便，能给丘陵地区立体农业智能化提供一个实用的方案。

一、整体设计方案

系统按“感知-决策-执行-监测”这四层来搭（总体框架见图1）。

感知层：K230 跑 YOLOv5 做水果实时识别，RGB-D 相机拿深度信息来定位 [7]，SHT11、GY302、MG811 这些传感器负责采集环境数据。

决策层：K230、STM32 和工控机三级配合，K230 管视觉推理，STM32 管机械臂和舵机控制，工控机负责 SLAM 建图和汇总数据 [8]。

执行层：多自由度机械臂加上二维云台，完成无损抓取和分类投放，替代了传统的 6 自由度方案。

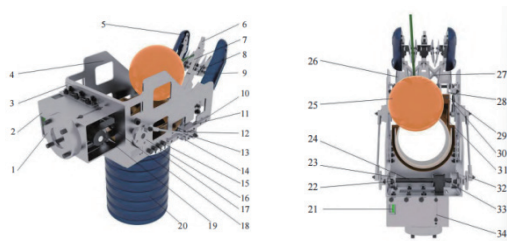
通信层：LoRa 搭局域网，4G 上传云平台，方便远程操控 [9]。硬件核心之间用串口直连，舵机由驱动板统一控制，7.4V 锂电池供电。



图1系统总体设计框架图

二、机械结构设计

系统机械结构采用模块化、轻量化设计，由移动底盘、视觉安装平台、二维云台、多自由度机械臂及末端柔性夹爪五部分组成，末端执行器结构如图2所示。移动底盘采用铝合金板材，作为 K230、云台与机械臂的固定底板。K230 固定于云台上方，随云台同步运动。二维云台为双轴舵机结构，X、Y 轴独立控制，配合 24 路舵机驱动板实现角度调节。机械臂采用 3-4 自由度舵机驱动，末端配备柔性夹爪，可自适应果实外形，配合云台摆动完成投放。云台初始居中，X 轴 PWM 范围 500~2400，Y 轴 300~2500，初始值均为 1600，软件限幅保障安全。采用 PID 算法与卡尔曼滤波对舵机运动平滑处理。工作流程：识别果实→机械臂伸出抓取→收回→云台左右摆动分类投放→复位。整体结构依靠柔性夹持与摆动分拣实现无损作业，维护便捷，适合中小型果园低成本部署。



a. 末端执行器正常状态下侧视 b. 末端执行器工作状态下俯视图

1. 法兰盘 2. 电机支架 3. 薄折弯板 4. 机架 5. 柔性指套 6. 远端指节 7. 中节 8. 方形折弯板 9. 刀片 10. 导轨 11. 滑块连接板 12. 鱼眼接头 13. 一级连杆 14. 滑块 15. 螺栓 16. 凸轮 17. 喉箍 18. 传动带 19. 一级同步轮 20. 运输管 21. 电机 22. 螺母 23. 轴承 24. 传动轴 25. 脐橙 26. 单指构型 27. 连接块 28. 连接底座 29. 斜面底座 30. 二级连杆 31. 隔板 32. 垫片 33. 二级同步轮 34. 电机安装板

图2末端执行器结构图

三、系统硬件设计

（一）轻量化嵌入式控制单元

这套系统用的是 K230 视觉模块加 STM32 运动控制器，双核心配合，替代了传统那种工控机配机械臂的高成本方案。

K230：自带 KPU 神经网络加速器，跑 YOLOv5 轻量模型，认苹果、香蕉、橘子，顺便定位并把坐标和类别通过串口发出去。

STM32：负责串口数据解码、规划机械臂轨迹、做云台 PID 闭环控制、卡尔曼滤波平滑处理，最后输出 PWM 信号。

协同逻辑也不复杂：K230 识别定位→下发类别和坐标→STM32 做运动规划→机械臂抓取→云台分拣→自动复位。

（二）多模态传感器组

视觉感知：K230 拍 RGB 图像，YOLOv5 分出水果类别；RGB-D 相机提供深度信息，完成果实的三维定位。

环境监测：温湿度用 SHT11，光照用 GY302，CO₂ 用 MG811，实时采集农田环境数据。

姿态反馈：机械臂关节和云台 X/Y 轴舵机都带位置反馈，闭环控制精度挺高。

（三）采摘与分拣执行机构

系统用了机械臂抓取加云台分拣的双执行机构，既保证采摘稳定，又兼顾分拣效率。

1. 机械臂抓取机构

3 到 4 自由度，舵机驱动，末端是柔性夹爪；STM32 根据 K230 给的坐标解算关节角度，自动对准、抓取、收回。

2. 二维云台分拣机构

X 轴舵机：S1 接口，PWM 500~2400；

Y 轴舵机: S2 接口, PWM 300~2500。

7.4V 供电, 运动平滑。

3. 动作逻辑

苹果→抓取→云台左摆→投放→回中;

橘子→抓取→云台右摆→投放→回中。

(四) 通信与能源模块

K230 与 STM32 硬件接线如图 3 所示。

通信接口: K230 TX ↔ STM32, USART3_RX,
RX ↔ USART3_TX, 5V/GND 共地;

数据帧: 0x55+0xaa+ 数据 +0xfa, 校验可靠, 防丢包防
错位;

供电: 7.4V 锂电池为主电源, 续航≥4 小时, 支持连续采摘
作业。

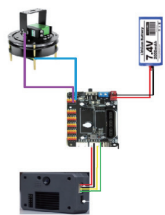


图3 K230 与 STM32 硬件接线图

四、软件程序设计

(一) 系统主程序流程

系统上电流程:

上电初始化→摄像头开启→水果识别与定位→数据帧发送
→STM32 解码→机械臂抓取→云台分拣→自动复位→环境数据采
集上传→循环执行。

下述为 K230 的系统主程序:

```
try:
```

```
while True:
```

os.exitpoint() # K230 系统退出点检测 (支持外部中断
退出)

with ScopedTiming("total", 1): # 耗时统计 (可选, 用
于性能调优)

```
# 1. 获取当前帧图像
```

```
img=pl.get_frame()
```

```
# 2. 模型推理
```

```
res=yolo.run(img)
```

```
# 3. 解析推理结果并通过串口输出
```

```
flag=res[0]
```

```
if flag == -1: # 无有效检测结果 (置信度 <0.65)
```

uart.write(bytearray([0x55, 0xaa, 0x00, 0xfa])) # 自定义串
口指令

if flag == 0: # 检测到目标 (具体标签需结合 res 后
续参数解析)

```
uart.write(bytearray([0x55, 0xaa, 0xff, 0xfa]))
```

```
# 4. LCD 绘制推理结果 (标签、置信度等)
```

```
yolo.draw_result(res, pl.osd_img)
```

```
#5.LCD 显示图像
```

```
pl.show_image()
```

```
#6. 手动回收内存
```

```
gc.collect()
```

(二) 视觉识别程序设计

基于 K230 平台部署 YOLOv5 轻量化分类模型, 实现水果实
时识别:

1. 图像流水线: PipeLine 获取 1280×720 原始帧, 缩放至
224×224 适配模型输入;

2. 模型推理: 加载 fruit_cls.kmodel, 置信度阈值 0.7;

3. 数据输出: 识别结果按帧格式发送至 STM32。

K230 水果实时识别效果如图 4 所示。

代码关键逻辑:

```
# K230 核心代码
```

```
labels = ["apple", "banana", "orange"]
```

```
conf_thresh = 0.7
```

```
if flag == 0: # 识别到苹果
```

```
uart.write(bytearray([0x55, 0xaa, 0x00, 0xfa]))
```

```
if flag == 1: # 识别到橘子
```

```
uart.write(bytearray([0x55, 0xaa, 0xff, 0xfa]))
```

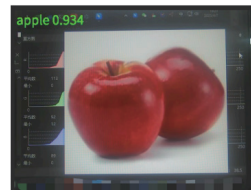


图4 K230 水果识别效果图

(三) STM32 运动控制程序设计

1. 串口数据解码

通过帧头 0x55、帧尾 0xfa 校验, 确保 K230 数据可靠接收,
串口数据帧格式如下所示。

```
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
uint8_t data[10]; // 帧头 0x55, 帧尾 0xfa, 数据 8 字节
```

2. 机械臂抓取控制

STM32 根据坐标完成机械臂关节角度解算、柔性夹爪闭合 /
张开控制、抓取 - 回位 - 等待分拣。

3. PID+ 卡尔曼滤波运动控制

为保证运动平稳性, 融合分段插补、位置式 PID 及卡尔曼滤
波。PID 控制积分项限幅 ±7000, 抑制积分饱和; 卡尔曼滤波平
滑运动步长; 云台单次摆动拆分为 200 步插补输出。核心控制逻
辑如下:

```
float pid_calculate(PID_TypeDef *PID, float CurrentValue)
```

```

{
PID->Err = PID->Target - CurrentValue;
PID->Integral += PID->Err;
if(PID->Integral > 7000){
PID->Integral = 7000;
}
if(PID->Integral < -7000){
PID->Integral = -7000;
} // 积分限幅
PID->PID_out = PID->Kp * PID->Err /* 比例 */
+ PID->Ki * PID->Integral /* 积分 */
+ PID->Kd * (PID->Err - PID->LastErr); /* 微分 */
PID->LastErr = PID->Err;
return PID->PID_out;
}

```

4. 舵机与机械臂驱动

STM32定时器输出20ms周期的PWM脉冲，脉宽 $500\mu\text{s} \sim 2500\mu\text{s}$ 对应 $0^\circ \sim 180^\circ$ 。分段插补把角度差拆成200步微小增量，再加个行程限幅保护。二维云台和机械臂各关节配合着来，完成伸臂、抓取、夹爪开合、回位这些动作。

五、系统关键技术与创新分析

（一）轻量化视觉—控制—执行一体化架构

用K230+STM32这套轻量嵌入式架构，替代传统工控机和大型机械臂，把视觉识别、运动控制、执行机构捏在一起。K230本地跑YOLOv5推理，STM32负责运动闭环控制和指令解析，整体体积小、功耗低、成本也控得住，适合南方丘陵立体农业那种空间窄、供电受限的作业场景。

（二）机械臂无损抓取+二维云台高效分拣

机械臂无损抓取加云台摆动分拣，两个执行机构配合。末端柔性夹爪能自适应果实外形，损伤率自然就降了；二维云台是双轴舵机驱动，结构简单响应快，左右摆动就能快速分类投放，维护成本低，跑着也稳。

（三）可靠串口通信与运动平滑算法

自定义串口数据帧配环形缓冲解读码，加上帧头帧尾校验，有效避免丢包和错位，通信更可靠。运动控制这边融合了PID、卡尔曼滤波和分段插补，抑制抖动和超调，定位精度和平稳性都上来了。

（四）多模态数据融合与远程操控

把视觉识别、环境监测、姿态反馈揉到一起做多模态融合，果园状态看得更全面。通信层是LoRa+4G混合组网，支持数据实时上传和远程查看，在复杂果园环境下的适应性和安全性都有提升^[10]。

六、实验测试结果

（一）果实识别性能测试

多场景测下来，系统对苹果、香蕉、橘子这三类的平均识别

置信度不低于0.85，60%遮挡下准确率还能超过82%，抗干扰能力挺强，满足果园复杂环境下的稳定识别需求。

（二）采摘与分拣性能测试

连续作业测试显示，抓取成功率93%，果实损伤率控制在2.5%以内，响应延迟不到200ms，而且能连续稳定工作4小时以上不出故障，作业效率和可靠性都达标。

（三）环境监测、通信测试与样机展示

LoRa和4G的传输延迟都低于200ms，空旷果园有效距离 $\geq 2\text{km}$ ，丢包率 $< 1\%$ ，环境参数采集稳，通信靠谱。系统整体样机如图5所示。



图5 系统整体样机实物图

七、结束语

针对立体农业采摘效率低、监测不准、部署成本高等问题，本文设计了以K230和STM32为核心，加上多自由度机械臂和二维云台的智能采摘与环境监测系统，实现了高精度识别、无损采摘、分类投放及环境数据远程传输。实验表明，识别置信度 ≥ 0.85 ，抓取成功率93%，损伤率 $\leq 2.5\%$ ，响应和传输延迟都小于200ms。这套系统轻量化、低成本、易部署，能给南方丘陵立体农业智能化提供一个可行方案。

参考文献

- [1] 陈青, 殷程凯, 郭自良, 等. 苹果采摘机器人关键技术研究现状与发展趋势 [J]. 农业工程学报, 2023, 39(04): 1-15.
- [2] 朱军, 陈钢, 蒋丽花, 等. 轻量化改进YOLOv8n的荔枝病虫害检测方法 [J]. 农业工程学报, 2025, 41(13): 207-215.
- [3] 戈刚, 杨璐, 刘毅, 等. 改进YOLOv8n模型对复杂农业场景下西红柿目标的检测 [J]. 农业工程学报, 2025, 41(09): 143-153.
- [4] 杨森, 张鹏超, 王磊, 等. 集成改进YOLOv8n与通道剪枝的轻量化番茄叶片病虫害识别方法 [J]. 农业工程学报, 2025, 41(02): 206-214.
- [5] 蒲应俊, 罗锦艳, 张贵贤, 等. 仿手指拨动切割式柑橘采摘末端执行器设计与试验 [J]. 农业工程学报, 2025, 41(24): 63-73.
- [6] 涂德浴, 赵虎, 朱庆, 等. 一种刚柔结合柔性夹持器的设计 [J]. 现代制造工程, 2023, (10): 120-125.
- [7] 周宏平, 金寿祥, 周磊, 等. 基于多模态图像的自然环境下油茶果识别 [J]. 农业工程学报, 2023, 39(10): 175-182.
- [8] 朱泓逸, 李秋洁. 基于起伏地形的果园仿真环境搭建与SLAM算法测试 [J]. 林业工程学报, 2024, 9(02): 133-142.
- [9] 冀汶莉, 王佳豪, 王新伟. 基于LoRa的农业大田土壤多参数监测系统设计 [J]. 无线工程, 2023, 53(02): 456-464.
- [10] 赵春江, 范贝贝, 李瑾, 等. 农业机器人技术进展、挑战与趋势 [J]. 智慧农业(中英文), 2023, 5(04): 1-15.