

自然环境下柑橘采摘机器人识别定位系统研究

杨长辉^{1,2} 刘艳平² 王毅² 熊龙焯² 许洪斌² 赵万华¹

(1. 西安交通大学机械工程学院, 西安 710049; 2. 重庆理工大学机械工程学院, 重庆 400054)

摘要: 为了准确理解柑橘采摘机器人在自然环境下的作业场景, 获取采摘目标及周围障碍物的位置信息, 构建了基于卷积神经网络和 Kinect V2 相机的识别定位系统。首先, 对采摘场景中的果树提出 5 类目标物分类准则, 包含 1 类可采摘果实和 4 类障碍物目标; 然后, 在 YOLO V3 (You only look once) 卷积层模块中添加 3 层最大池化层, 对预测候选框进行 K-means 聚类分析, 增强模型对枝叶类物体特征的提取能力, 实现采摘场景的准确理解; 最后, 采用 Kinect V2 相机的深度图映射得到采摘目标和障碍物的三维信息, 并在自然环境下进行了避障采摘作业。实验结果表明, 构建的识别定位系统对障碍物和可采摘果实的识别综合评价指数分别为 83.6% 和 91.9%, 定位误差为 5.9 mm, 单帧图像的处理时间为 0.4 s, 采摘成功率和避障成功率分别达到 80.51% 和 75.79%。

关键词: 采摘机器人; 目标识别; 三维定位; YOLO V3; K-means

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2019)12-0014-09

Research and Experiment on Recognition and Location System for Citrus Picking Robot in Natural Environment

YANG Changhui^{1,2} LIU Yanping² WANG Yi² XIONG Longye² XU Hongbin² ZHAO Wanhua¹

(1. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

2. College of Mechanical Engineering, Chongqing University of Technology, Chongqing 400054, China)

Abstract: For citrus picking robot in natural environment, the accurate recognition and location vision system is one of the key factors ensuring the efficiency and safety of picking operations. In order to make the robot not only acquire the location information of the picking target accurately but also the surrounding obstacles, a novel obstacle recognition and location system based on Kinect V2 and improved you only look once (YOLO V3) algorithm was proposed. Firstly, five classification principles of citrus tree in natural orchard were defined, including one class that the fruit can be picked directly and four obstacle classes. Secondly, three maximum pooling layers were added to the convolution module of the YOLO V3 structure and K-means clustering analysis was conducted on anchor box to enhance the feature extraction performance of branches and leaves of the convolution neural network. Finally, three-dimensional coordinates of the classification targets were obtained by using the Kinect V2 depth mapping to guide obstacle-avoiding picking operation. The experimental results showed that the F_1 - scores of obstacles and normal fruits were 83.6% and 91.9%, respectively, the positioning error was 5.9 mm and the processing time of each frame was 0.4 s, the success picking rate was 80.51% and success rate of obstacle avoidance was 75.79%. The research results provided a basis and guide for the picking path planning and obstacle avoidance of robotic harvesting task in natural scene.

Key words: picking robot; obstacle recognition; 3D positioning; YOLO V3; K-means

收稿日期: 2019-04-21 修回日期: 2019-05-23

基金项目: 重庆市重点产业共性关键技术创新专项 (cstc2015zdcy-ztxx70003) 和重庆市基础科学与前沿技术研究一般项目 (cstc2016jcyjA0444)

作者简介: 杨长辉 (1976—), 男, 博士生, 重庆理工大学副教授, 主要从事智能农业机械和机器视觉研究, E-mail: chhyang@stu.xjtu.edu.cn

通信作者: 赵万华 (1965—), 男, 教授, 博士生导师, 主要从事智能装备和精密数控机床集成理论研究, E-mail: whzhao@mail.xjtu.edu.cn

0 引言

近年来,由于人力成本的上升导致果蔬采摘成本大幅提高,果实采摘机器人逐渐成为研究热点,有关番茄、苹果、甜椒^[1-4]等采摘机器人的研究居多。识别定位系统是采摘机器人的重要组成部分,通常采用机器视觉方法实现对采摘环境的感知^[5]。文献[6]通过 K-means 聚类分割和 Hough 圆拟合实现柑橘果实的识别分割,文献[7]采用贝叶斯分类器实现樱桃树树干识别,文献[8]通过 K-means 聚类算法和 HSV 阈值分割方法实现柑橘果实的识别。上述传统机器视觉方法在自然环境中对光照变化较为敏感,不能准确提取遮挡果实及枝干等障碍物的特征信息,从而导致机器人不能准确感知周围环境,在采摘过程中经常会因为碰撞障碍物导致采摘失败,对执行器或机械臂造成破坏。

目前,深度学习在物体检测领域取得了巨大进展,可较好地解决机器视觉在自然环境下识别物体泛化性较弱的问题。如文献[9]针对田间环境下的多簇猕猴桃果实识别,提出基于卷积神经网络的识别方法,实现了多类果实的同时识别。文献[10]采用基于 SSD (Single shot multibox detector) 卷积神经网络的多类水果识别模型,通过加深网络模型和数据集增强处理,提升识别正确率。文献[11]采用基于 Darknet19 的网络识别柑橘果实,该方法对果实的识别率达到 86.9%,实现了柑橘果实的多场景识别。文献[12]针对苹果树枝干提出了基于 R-CNN (Regions-convolutional neural network) 的识别方法。文献[13]基于 SSD 检测模型和 Kinect V2 三维测量模型实现物体的快速准确定位。

以上研究成果在自然环境下均能实现对采摘果实和采摘场景的识别。本文针对传统机器视觉方法在自然环境下不能准确理解采摘场景的问题,构建柑橘采摘机器人识别定位系统,提出障碍物的分类准则。通过改进 YOLO V3 识别算法实现障碍物和采摘果实的识别分类,采用 Kinect V2 深度图映射得到柑橘果实和障碍物的位置信息,实现目标物柑橘和障碍物的快速识别和精准定位,并在自然环境下使用自行研制的柑橘采摘机器人进行实验验证。

1 目标识别定位流程与算法

1.1 识别定位系统组成

柑橘采摘机器人识别与定位系统由物体识别和三维定位两个模块组成。物体识别模块采用卷积神经网络作为识别算法,将原图的目标物进行分类并

得到目标物位置边框信息。定位模块将每个目标物的位置边框信息转换为目标物二维中心点坐标,通过二维中心点坐标在 Kinect V2 深度映射图上获取该坐标点深度值。最后采用 Kinect V2 三维定位技术将目标物二维中心点坐标转换为三维空间坐标。系统组成如图 1 所示。

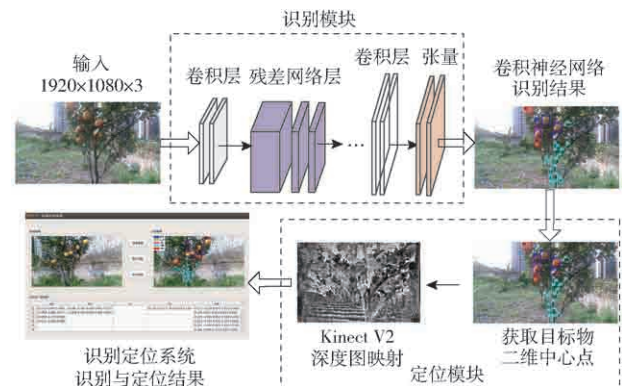


图 1 识别定位系统组成

Fig. 1 Schematic of recognition and position system

1.2 果实与障碍物位置关系分类

柑橘采摘机器人采摘场景中主要由柑橘果实、树叶和枝干 3 类物体组成,本文根据 3 类物体在柑橘树上不同的空间位置关系定义 4 类障碍物目标和 1 类可采摘果实,详细定义如下:

(1) 枝干遮挡果实 (Branch occluded), 简称 BO。定义枝干对果实造成的遮挡为枝干遮挡果实类。

(2) 树叶遮挡果实 (Leaf occluded), 简称 LO。设树叶遮挡面积为 S_L , 被遮挡果实拟合圆面积为 S_f , 二者之比为

$$P = S_L / S_f \quad (1)$$

定义 $P \geq 1/3$ 为树叶遮挡果实类。

(3) 重叠果实 (Overlapping), 简称 OL。设果实重叠数量为 N , 定义 $2 \leq N \leq 4$ 为重叠果实类。

(4) 枝干类 (Main branch), 简称 MB。果树的一、二级枝干承担着果树的营养运输任务,为了保护果树的一、二级枝干不受破坏,同时也为了确保采摘机器人的安全,定义果树的一、二级枝干为枝干类。

(5) 正常无遮挡果实 (Normal), 简称 NM。果实前面无遮挡物,树叶遮挡面积小于 $1/3$ 均为正常无遮挡果实类。

以上 5 类中,正常无遮挡果实能直接进行采摘。由于本课题组设计的末端执行器只能针对单个果实进行采摘作业^[14],因此树叶遮挡果实、枝干遮挡果实和重叠果实均不能直接采摘,将其和枝干类一起作为采摘过程中的障碍物。如图 2 所示。

1.3 障碍物识别检测模型

YOLO V3^[15-17] 具有识别速度快、识别准确率高



图 2 识别目标物类别

Fig. 2 Classification of recognition object

等优点。该算法将整幅图像用作输入,并将图像分割成 $S \times S$ 大小的网格,若某个格子中包含检测目标,则该格子负责检测该目标,并预测位置边框和物体置信度。位置边框信息为该格子位置的偏移宽度和高度,置信度反映是否包含物体以及包含物体情况下的准确性。位置边框采用 logistic 进行预测

$$\begin{cases} b_x = \sigma(t_x) + c_x \\ b_y = \sigma(t_y) + c_y \\ b_w = p_w e^{t_w} \\ b_h = p_h e^{t_h} \end{cases} \quad (2)$$

式中 c_x, c_y ——网格的坐标偏移量
 p_w, p_h ——预设 anchor 框的边长
 b_x, b_y, b_w, b_h ——最终得到的边框坐标值
 t_x, t_y, t_w, t_h ——网络学习目标

通过对每种尺度预测多个边框来提高多尺度目标预测准确率。

对于枝干和树叶遮挡的果实,特征信息相似度较高,相互干扰较为严重,因此采用最大池化层强化卷积神经网络在训练过程中对枝干和树叶遮挡纹理特征的提取,以提高识别准确率。在 YOLO V3 第 2 个卷积层模块中,添加 3 层最大池化层,第 1 层尺寸为 5×5 ,第 2 层尺寸为 9×9 ,第 3 层尺寸为 13×13 ,第 2 和第 3 层最大池化层融合卷积模块中第 1 层卷积层输出的特征信息,如图 3 所示。

1.4 障碍物三维定位

获取 Kinect V2 视场内障碍物的三维坐标 (X, Y, Z) ,其中三维坐标中的深度 Z 通过 Kinect V2 自带的 SDK 函数 (MapColorFrameTo DepthSpace) 获得,该函数将彩色相机 (1 920 像素 $\times$ 1 080 像素) 的像素点映射到深度相机空间中,得到彩色图中每个像素点的深度 Z 。通过彩色相机的标定得到内外参

数矩阵,从而建立像素坐标系到相机坐标系的映射关系,得到坐标 (X, Y, Z) 。

1.4.1 像素坐标点与基坐标空间点的转换

采用张正友标定法^[18]和 Matlab 标定工具箱对 Kinect V2 彩色镜头进行标定。得到标定结果内部参数矩阵 K_{rgb} 为

$$K_{rgb} = \begin{bmatrix} 1\,092.213\,7 & 0 & 963.451\,5 \\ 0 & 1\,093.432\,7 & 540.945\,0 \\ 0 & 0 & 1 \end{bmatrix}$$

则彩色镜头焦距为 $(f_x, f_y) = (1\,092.213\,7, 1\,093.432\,7)$,主光点为 $(u_0, v_0) = (963.451\,5, 540.945\,0)$ 。

物体的三维空间坐标与图像坐标存在线性关系^[19],假设相机坐标系中采摘点 P_c 的坐标为 (X_c, Y_c, Z_c) ,像素坐标系上与之对应的坐标为 $p(x, y)$,其关系式为

$$\begin{cases} X_c = \frac{Z_c}{f_x}(x - u_0) \\ Y_c = \frac{Z_c}{f_y}(y - v_0) \end{cases} \quad (3)$$

其中 Z_c 由深度镜头获取, u_0, v_0, f_x, f_y 为前面标定结果,代入式 (3) 可得

$$\begin{cases} X_c = \frac{Z_c(x - 963.451\,5)}{1\,092.213\,7} \\ Y_c = \frac{Z_c(y - 540.945\,0)}{1\,093.432\,7} \end{cases} \quad (4)$$

利用式 (4) 可计算物体标记点在相机坐标系中的三维坐标。实际采摘中使用机械臂坐标系为基准,因此需要将相机坐标系的三维坐标转换为对应基坐标系中的三维坐标。在机械臂末端放置带有标记点的标记板,通过控制移动机械臂到基坐标系中某点 $P_{rob}(X_{rob}, Y_{rob}, Z_{rob})$ 处,利用式 (4) 计算点 P_{rob} 位置处标记点在相机坐标系中的三维点坐标 (X_c, Y_c, Z_c) 。按照上述步骤重复测量 80 组基坐标系与相机坐标系对应的三维点坐标,使用 Opencv 函数 EstimateAffine3D 建立相机坐标系与基坐标系的转换关系,该函数为计算 2 个三维点集的理想仿射变换矩阵,得到仿射变换矩阵

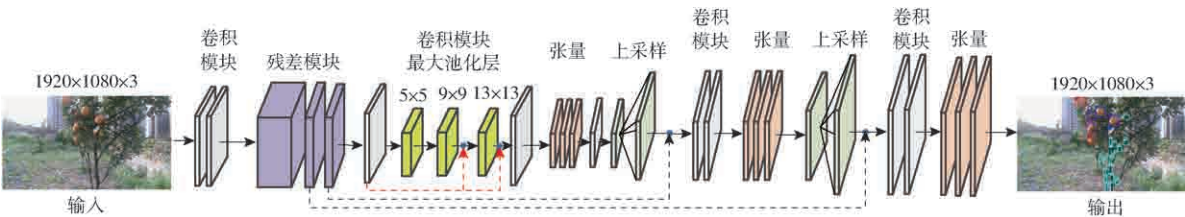


图 3 障碍物识别检测网络模型

Fig. 3 Network model for obstacle recognition and detection

$${}^r\mathbf{M}_c = \begin{bmatrix} -0.244\ 8 & 0.207\ 0 & 0.742\ 0 & 0.137\ 0 \\ -1.049\ 8 & -0.079\ 4 & -0.152\ 0 & 0.445\ 8 \\ -0.052\ 4 & -0.988\ 9 & 0.152\ 0 & 1.156\ 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

则

$$\begin{bmatrix} X_{\text{rob}} \\ Y_{\text{rob}} \\ Z_{\text{rob}} \\ 1 \end{bmatrix} = {}^r\mathbf{M}_c \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (5)$$

由式(5)可计算物体中心点在基坐标系中的三维坐标,其中 X_{rob} 、 Y_{rob} 、 Z_{rob} 为采摘机器人基坐标系中三维坐标值。

1.4.2 目标物空间特征获取

Kinect V2 彩色相机采集的图像经过本文识别模型识别后,输出结果包含目标物的类别信息和位置边框,位置边框信息为像素坐标系中左上角 $P_{\text{LT}}(X_L, Y_L)$ 点坐标和右下角 $P_{\text{RD}}(X_R, Y_R)$ 点坐标,如图 4 所示。

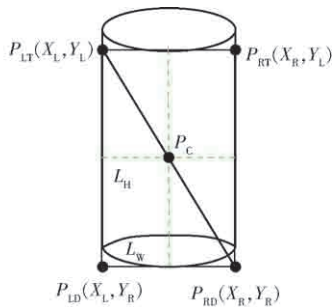


图 4 障碍物三维空间示意图

Fig. 4 Digram of 3D space area of obstacles

由点 P_{LT} 和点 P_{RD} 像素坐标即可求得点 P_C 像素坐标为 $\left(\frac{X_L + X_R}{2}, \frac{Y_L + Y_R}{2}\right)$ 。由式(4)、(5)可求得点 P_C 三维坐标为 (X_C, Y_C, Z_C) 。将点 P_C 的深度 Z_C 作为点 P_{LT} 、点 P_{LD} 、点 P_{RT} 和点 P_{RD} 的深度,则点 P_{LT} 和点 P_{RD} 在相机坐标系中三维坐标为 (X_L, Y_L, Z_C) 和 (X_R, Y_R, Z_C) 。采摘机器人采摘作业过程中需要获取 5 类目标物详细的位置特征信息。计算三维坐标点 $P_{\text{LD}}(X_L, Y_R, Z_C)$ 和点 $P_{\text{RD}}(X_R, Y_R, Z_C)$ 距离 L_W ,以及点 $P_{\text{RT}}(X_R, Y_L, Z_C)$ 和点 $P_{\text{RD}}(X_R, Y_R, Z_C)$ 距离 L_H ,则

$$\begin{cases} L_W = |X_L - X_R| \\ L_H = |Y_L - Y_R| \end{cases} \quad (6)$$

点 $P_C(X_C, Y_C, Z_C)$ 为 NM 类的采摘点。以点 P_C 为圆柱体表面中心, $\min(L_W, L_H)$ 为直径、 $\max(L_W, L_H)$ 为高度构造的圆柱体为障碍物 BO、LO、OL 和 MB 类的三维空间区域。

2 数据集制作与模型训练

2.1 图像采集

图像采集平台包括微软公司生产的 Kinect V2 相机、相机三脚架和联想 Think Pad E535 型便携式计算机,如图 5a 所示。Kinect V2 相机结构如图 5b 所示,彩色镜头分辨率为 $1\ 920$ 像素 $\times$ $1\ 080$ 像素,深度镜头分辨率为 512 像素 $\times$ 424 像素,垂直和水平方向视场角分别为 60° 和 70° ,深度测量范围为 $0.5 \sim 4.5$ m。光照强度测试仪为胜利仪器生产的 VC1010A 型。

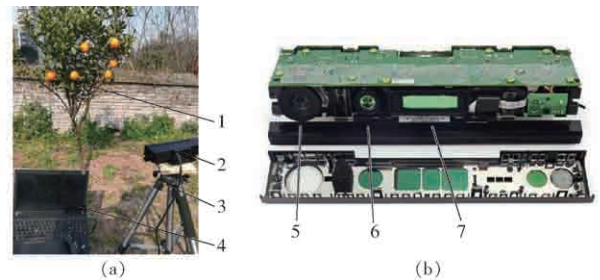


图 5 图像采集平台硬件

Fig. 5 Platform for collecting experimental dataset

1. 柑橘树 2. Kinect V2 相机 3. 相机三脚架 4. 计算机终端
5. 彩色镜头 6. 深度镜头 7. 红外镜头

本文数据集采自于重庆市北碚金果园(果园 A)与重庆理工大学柑橘实验基地(果园 B),柑橘树树龄为 12 a。相机感受视野与拍摄距离 L_C 相关,设相机视场角 $\alpha = 60^\circ$,树冠视为直径为 d 的类球体。则

$$L_C = \frac{d}{2} \left(\cot \frac{\alpha}{2} - 1 \right) \quad (7)$$

统计果园 A 和果园 B 中随机分布的 160 株 12 a 树龄柑橘树树冠直径,树冠直径直方图如图 6 所示。由图可知,树冠直径集中在 $1.0 \sim 1.1$ m,取该区间平均值 1.05 m 作为树冠直径。由式(7)可知 L_C 为 0.38 m。结合 Kinect V2 相机最小测距距

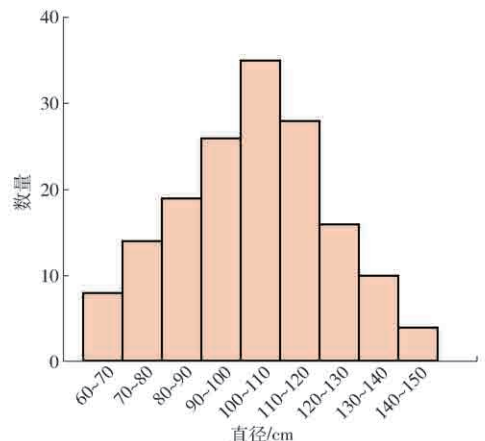


图 6 树冠直径直方图

Fig. 6 Histogram of canopy diameter

离为 0.50 m,本文使用的采摘机械臂最大采摘距离为 0.90 m,因此 L_c 应满足: $0.50\text{ m} < L_c < 0.90\text{ m}$ 。

为了模拟自然环境下不同光照强度的环境,设计 3 种拍摄角度:顺光、侧光和逆光,所有训练数据集均按此拍摄角度进行采集。

2.2 数据集建立

在果园 A 和果园 B 中采集 1 400 幅图像,从中随机选用 1 200 幅图像按照 3:1 配置训练集和测试集,200 幅作为验证集。采用 labelImg 标注工具进行数据集标注,用鼠标框选目标物,形成的四边形框外接目标物轮廓边缘。本文对形状不规则的枝干采用形状规则的四边形离散化标记,四边形的尺寸没有强制性要求,其包含的像素点面积尽量和柑橘果实包含的像素点面积一致,四边形边框两条对角线中的任一对角线应在枝干区域内。标记样例如图 7 所示。



图 7 训练数据集标注样例

Fig.7 Example of training dataset annotation

2.3 anchor boxes 参数设置

YOLO V3 使用 anchor boxes 对图像中的目标进行检测,其尺寸设置与数据集的标记框尺寸有关。本文对形状不规则的枝干采用离散化标记方法,枝干类和果实类的标记边框尺寸差异明显,因此需要对数据集的标记框尺寸做 K-means 聚类分析^[20],得到合适的 anchor boxes 参数。

本文 K 选取 $[1, 40]$,分别对训练集中标记边框尺寸进行 K-means 聚类分析,得到的结果如图 8 所示。由图可知,在 $K = 12$ 之后,曲线下降趋势减缓,逐渐趋于平缓,所以 anchor boxes 设置为 12。

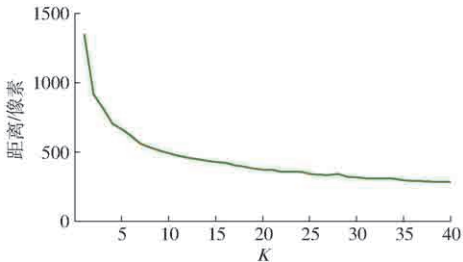


图 8 K-means 聚类分析结果

Fig.8 Result of K-means cluster analysis

2.4 模型训练与性能评估

识别模型使用迁移学习训练,训练过程中对数据集旋转、增加对比度、增加曝光进行数据增强,实现 3 倍数据集扩充。训练阶段动量项为 0.9,每一个 batch 包含 16 幅图像,衰减系数为 0.000 5。迭代训练 50 000 次,其中 1 000 ~ 5 000 次迭代训练中,权值学习率为 0.02,加速模型的收敛。5 000 ~ 15 000 次迭代训练中,权值学习率为 0.002,精调模型,15 000 ~ 24 000 次迭代训练中,权值学习率为 0.000 2,优化模型。模型训练过程中每 1 000 次迭代保存一次网络的权重文件。

模型迭代训练 50 000 次,其迭代训练过程中的损失率曲线如图 9 所示。由图 9 可知,训练模型在迭代训练 20 000 次左右时损失率不再下降。为了查看 m_{AP} (Mean average precision) 在迭代训练 20 000 次左右的变化趋势,在迭代训练 20 000 次基础上增加 8 000 次,即选用 28 000 以内的权重文件计算 m_{AP} ,其 m_{AP} 曲线如图 10 所示。

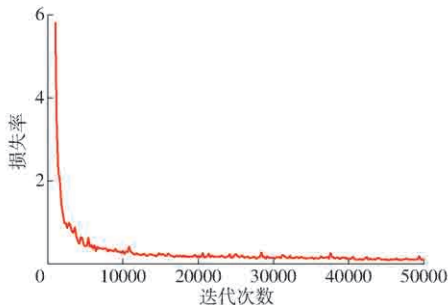


图 9 识别模型训练损失率曲线

Fig.9 Loss curve of proposed model

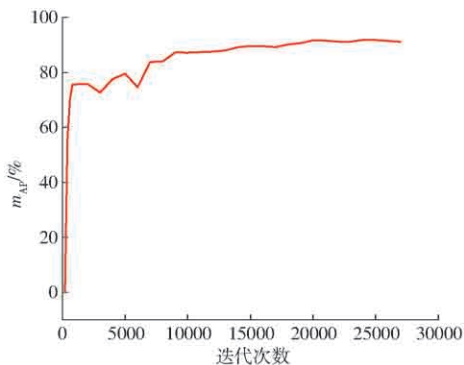


图 10 识别模型 m_{AP} 曲线

Fig.10 m_{AP} curve of proposed model

m_{AP} 计算公式为

$$m_{AP} = \frac{1}{5} \sum_{k=0.5}^N P(k) \Delta R(k) \quad (8)$$

其中

$$P = \frac{T_p}{T_p + F_p} \times 100\% \quad (9)$$

$$R = \frac{T_p}{T_p + F_N} \times 100\% \quad (10)$$

式中 $P(k)$ ——在阈值 k 时的准确率

$\Delta R(k)$ ——在阈值 k 时的召回率变化量
 T_p ——真正正样本数量
 F_p ——虚假正样本数量
 F_N ——虚假的负样本数量
 k ——阈值 P ——准确率
 R ——召回率 N ——引用阈值的数量

由图 10 可知,在迭代 20 000 次之后, m_{AP} 达到 85% 左右,并逐渐趋于平稳,因此本文选择迭代 20 000 次的权重文件作为识别模型。

3 实验结果与分析

3.1 障碍物与采摘果实识别实验

F_1 值是物体检测模型综合评价指标,是精确率和召回率的调和均值,因此本文使用 F_1 对模型的识别性能进行评估,其计算公式为

$$F_1 = \frac{2PR}{P + R} \tag{11}$$

3.1.1 验证集识别结果

使用验证集的 200 幅图像对迭代训练 20 000 次的识别模型进行性能测试,同时与改进前的 YOLO V3 算法对比,部分识别效果如图 11 所示。统计识别模型对 5 类目标物识别的 F_1 值。障碍物的 F_1 值为非 NM 类 F_1 值的平均值。图 11a 中 1 号黄色框为漏识别 MB 类,在统计 MB 类的 F_1 值时,漏识别的 MB 数量参照枝干标记规则,此处应为 3 个 MB 类目标。实验统计结果如表 1 所示。

由表 1 统计结果可知,本文方法对枝干类的 F_1 最低,为 79.4%;对 NM 类和 OL 类的 F_1 值较高,分别为 91.9% 和 89.3%;对 BO 类和 LO 类的 F_1 值为 83.5% 和 82.0%,对障碍物的 F_1 值为 83.6%。与

改进前的 YOLO V3 相比,本文方法对 NM 类、OL 类和 MB 类的提升效果不明显。对 BO 类和 LO 类的 F_1 值提升了 3.3 个百分点和 3.4 个百分点,对障碍物的 F_1 值提升了 2.1 个百分点。

其原因在于改进前模型对 BO 类和 LO 类的误识别率较高,相互干扰较为严重。改进后的模型强化了模型对枝干和树叶的纹理特征提取,减少了 BO 类和 LO 类的误识别,小幅提升了 F_1 值,图 11a 中 2 号黄色框与图 11e 中 4 号黄色框所示,YOLO V3 方法对 BO 类均出现了误识别。从图 11a 的识别效果可以发现,识别模型对 MB 类在枝干底部区域有较好的识别效果,在靠近树冠区域开始出现漏识别现象。其原因在于从柑橘树的生长姿态来看,靠近树冠区域的枝干其直径比底部区域枝干的直径要小。且由于靠近树冠区域,背景开始出现树叶和柑橘的干扰,降低 MB 类的 F_1 值。而 NM 类的特征较为明显,且训练数据集中 NM 类是最多的类别。因此 NM 类的 F_1 值是最高的。OL 类由于是多果重叠,其形状特征明显,尺度较大,且遮挡物干扰较少,有利于模型对 OL 类的特征提取。图 11a 中 OL 类的位置边框尺度为 198 像素 $\times$ 215 像素和 117 像素 $\times$ 220 像素,NM 类的边框尺度为 115 像素 $\times$ 127 像素。而图中较远的目标物对采摘机器人没有识别意义,不作为识别目标物。图 11c 中 3 号黄色框的柑橘尺寸为 38 像素 $\times$ 42 像素,而离相机最近的柑橘树上最小的柑橘尺寸为 72 像素 $\times$ 75 像素。BO 类和 LO 类由于枝叶的遮挡,造成模型对遮挡面积过大的目标难以识别。图 11e 中 5 号黄色框内的目标物,枝干遮挡面积超过了果实面积的 60%,且遮挡面积较多的目标物在训练集中较少出现,降低了模

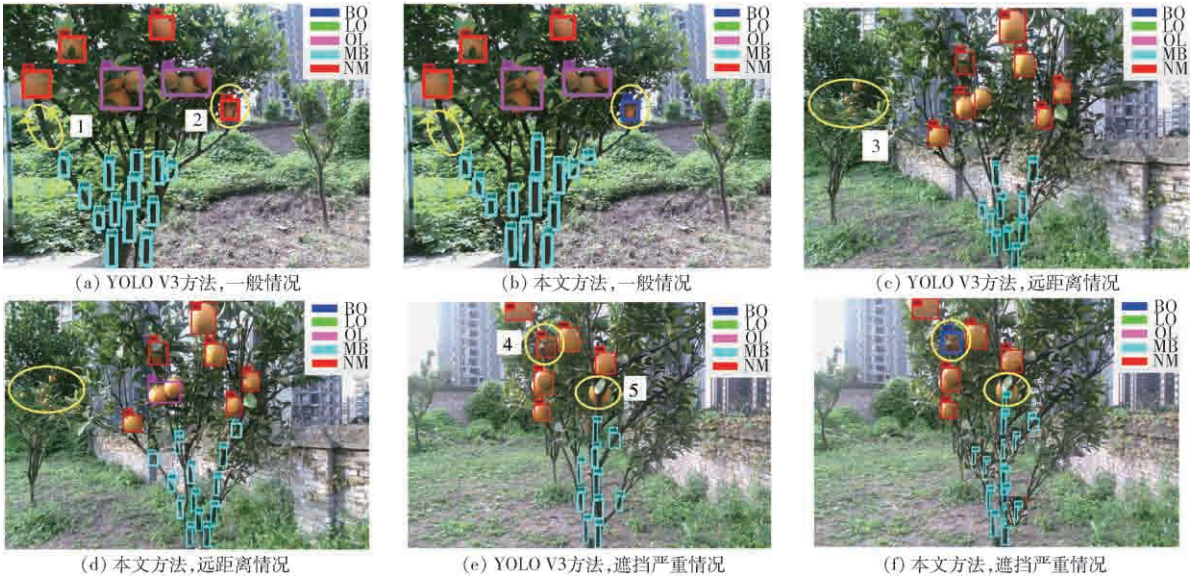


图 11 部分识别效果

Fig. 11 Partial recognition results

型对该类别的 F_1 值。综合来看,本文改进后的方法识别效果优于改进前的 YOLO V3 方法。

表 1 识别模型性能指标 F_1 统计结果
Tab.1 Recognition model performance indicators F_1

类别	YOLO V3 方法	本文方法
BO	80.2	83.5
LO	78.6	82.0
OL	88.6	89.3
MB	78.7	79.4
NM	92.3	91.9
障碍物	81.5	83.6

3.1.2 不同光照环境识别结果

由于本文的识别模型要满足采摘机器人在果园环境中的采摘要求,测试识别模型在不同光照环境中的识别性能。由 3.1.1 节可知,改进后的识别模型识别效果优于改进前,因此只对改进后的识别模型进行不同光照测试实验。按照采集数据集时的 3 种拍摄角度采集图像,每次拍摄采集时记录当前拍摄角度的光照强度。每种拍摄角度采集 20 幅图像,共计 60 幅图像作为模型在变光照环境中性能评估的验证集,统计 3 种拍摄角度的光照强度变化范围分别为 6 300 ~ 11 800 lx、48 000 ~ 76 000 lx 和 21 000 ~ 43 000 lx。计算模型在每一种光照强度区间中对 5 类目标物的平均 F_1 值。

在光照强度较弱的 6 300 ~ 11 800 lx 和 21 000 ~ 43 000 lx 环境中模型对 5 类物体的 F_1 最高为 83.8% 和 84.5%,识别效果最好。在该环境中,每类物体表面没有出现过曝区域,纹理特征明显,有利于卷积神经网络的识别。在强光照 48 000 ~ 76 000 lx 环境中, F_1 为 81.1%,比弱光环境低 3 个百分点左右。如图 12a 所示,图中拍摄时的光照强度为 69 600 lx。图中黄色框 1、2 均为目标物过曝区域,均失去枝干表面特征信息,造成卷积神经网络无法识别 MB 类目标物。图 12b 光照强度为 47 500 lx,图中 3、4 号黄色框中的枝干处于曝光不足,造成枝干表面出现黑色区域,不能真实反映枝干的特征,同样造成模型对曝光不足的 MB 类识别率较低。因此在曝光不足或过曝区域,本文识别模型对 MB 类的识别容易出现漏识别现象,降低模型的 F_1 值。



图 12 不同光照环境识别结果

Fig. 12 Recognition results in variable illuminations

3.2 三维定位与避障采摘实验

3.2.1 定位精度评估

为了评估本文识别与定位系统的定位精度,在自然环境下采用本课题组研发的采摘机器人进行定位测试实验^[21],采摘机器人平台如图 13 所示。由 Kinect V2 相机、履带式移动底盘、采摘机械臂、控制主机和自组研制的末端执行器组成。控制主机配置如下:CPU 为 Intel i7 7800X, GPU 为 2 块 11 GB NVIDIA1080Ti,主机运行内存 32 GB。



图 13 柑橘采摘机器人样机

Fig. 13 Overview of citrus picking robot

1. 采摘执行器 2. 采摘机械臂 3. Kinect V2 相机 4. 机器人控制器 5. 履带式移动底盘

Kinect V2 相机基于 TOF 实现物体深度值测量,其测量过程受光照变化影响较小^[22]。为了验证强光照环境中 Kinect V2 相机对采摘场景中目标物的深度值测量性能,设计如下实验:采用光照强度测试仪测量目标物表面光照强度,如图 14 所示。在 2019 年 5 月 16 日 13:00—17:00 时间段内,采用顺光拍摄角度采集图像和记录对应的光照强度,最后选取光照强度最强图像进行深度测量,如图 15 所示。Kinect V2 相机采集的深度图中每个像素点的灰度为 16 位二进制数,其数值在 0 ~ 65 535 之间,图像效果较暗。为了提高深度图的可视化效果,将深度图中每个像素点 16 位的灰度左移两位进行深度图增强,增强深度图只作为可视化效果图。



图 14 光照强度测试

Fig. 14 Measurement of light intensity on fruit surface

使用本文的三维坐标计算模型得到目标物标记点 P_{rob} 的三维坐标 $(X_{rob}, Y_{rob}, Z_{rob})$ 。在采摘机械臂末端关节安装长度为 45 mm 的探针,将坐标 $(X_{rob}, Y_{rob}, Z_{rob} - 45)$ 作为采摘机械臂的运动位置点,待采摘机械臂末端探针移动到 P_{rob} 位置时,测量探针端点 P_z 与标记点 P_{rob} 的距离 ΔE ,设定位系统在 X 、 Y 和



图 15 识别图与增强深度图

Fig. 15 Recognition result and enhanced depth images

Z 方向上的定位误差为 ΔX 、 ΔY 和 ΔZ , 则

$$\Delta E = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2} \tag{12}$$

ΔE 即为识别定位系统的综合定位误差。统计结果如表 2 所示。

表 2 柑橘果实中心点三维坐标点计算结果

Tab. 2 Results of 3D coordinate of citrus fruits central points

序号	光照强度/lx	X/m	Y/m	Z/m	ΔE /m
1	5.06×10^4	0.345	-0.056	0.848	0.006
2	5.16×10^4	0.284	-0.348	0.868	0.005
3	3.22×10^4	0.165	-0.412	0.921	0.007
4	5.45×10^4	0.124	0.281	0.836	0.005
5	3.65×10^4	0.118	-0.162	0.832	0.008
6	4.74×10^4	-0.037	0.324	0.751	0.004
7	4.36×10^4	-0.098	0.273	0.792	0.006
8	3.88×10^4	-0.236	-0.173	0.823	0.007
9	4.16×10^4	-0.184	-0.327	0.934	0.005
10	4.97×10^4	-0.168	-0.403	0.913	0.006
平均误差					0.005 9

由表 2 可知,测试期间,自然环境光照强度最强在 54 500 lx 左右,部分柑橘果实表面由于树叶的遮挡,其表面光照强度有所减弱。实验结果表明 Kinect V2 相机在光照强度 30 000 ~ 55 000 lx 的自然环境中,可实现三维坐标测量。平均定位误差达到 5.9 mm。定位误差主要来源于相机的标定误差和手动测量误差。同时由于测量过程中的实时性,自然环境中的风向造成目标物的随机微弱摆动对三维坐标的实时测量也有一定影响。

3.2.2 避障采摘实验

由本文设计的识别定位系统得到目标物边框中心点坐标,结合 Kinect V2 相机三维坐标计算模型和式(5)获得柑橘果实和枝干在基坐标系中的三维坐标点,将该坐标点作为果实的采摘点和避障路径特征坐标点。应用快速扩展随机树算法 (Rapid-exploration Random Trees, RRT) 进行避障^[23],该算法是一种基于随机采样的规划算法,能够快速有效地搜索高维空间。通过随机采样点,把搜索导向空白区域,从而寻找到一条从起始点到目标点的规划路径。

采摘机械臂完成一次采摘动作需要 14 s 左右,

本文设定每隔 5 s 进行一次图像采集和识别定位。设计每组采摘实验为采摘一棵柑橘树机械臂采摘范围内的可采摘柑橘果实,共进行 20 组实验。统计指标为采摘过程中成功采摘的柑橘果实数量和采摘机械臂与障碍物发生碰撞的次数。识别定位系统界面如图 16 所示,图中左上方为采集原图,右上方为识别效果图,下方包含 5 类目标物坐标信息,其中 NM 类为采摘中心点坐标,BO、LO、OL 和 MB 类为中心点坐标和三维圆柱体的直径和高度。采摘过程如图 17 所示。实验统计结果如图 18 所示。设可采摘果实数量为 Q ,成功采摘果实数量为 M ,与障碍物发生碰撞次数为 P ,则采摘成功率 ΔM 为

$$\Delta M = \frac{M}{Q} \times 100\% \tag{13}$$

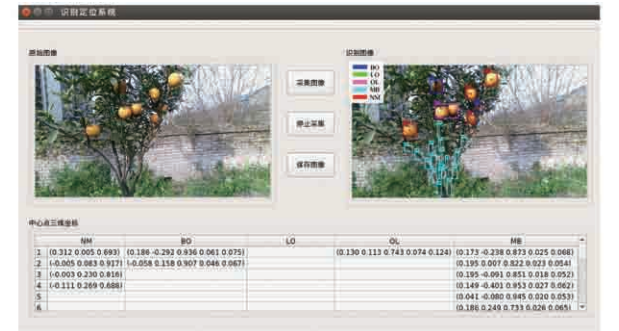


图 16 识别定位系统界面

Fig. 16 Recognition and location system interface

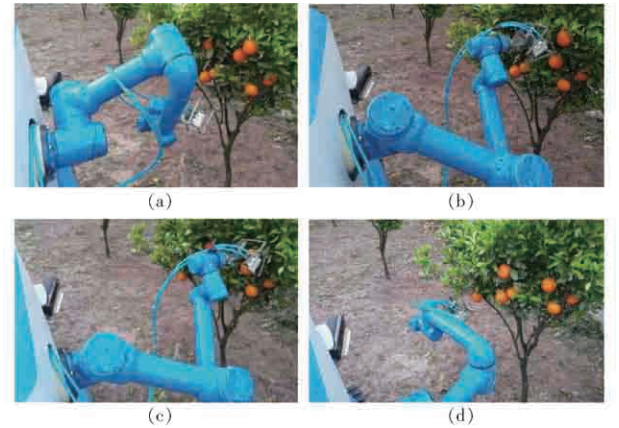


图 17 采摘实验过程示意图

Fig. 17 Process of picking experiment

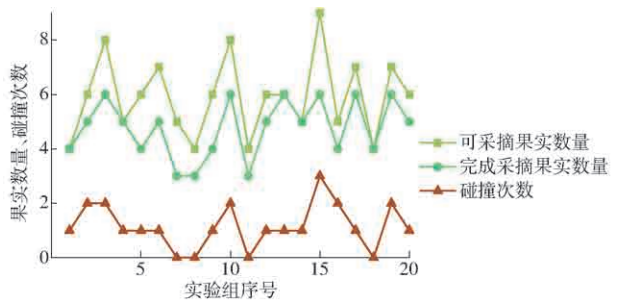


图 18 每组采摘实验统计结果

Fig. 18 Statistical results of harvesting experiments

避障成功率 ΔP 为

$$\Delta P = \frac{M - P}{M} \times 100\% \quad (14)$$

从统计结果可知,本文设计的识别定位系统在自然环境下采摘成功率达到 80.51%,避障成功率达到 75.79%。采摘实验中,存在采摘失败的主要原因在于采摘路径和采摘姿态求解失败^[24],导致采摘机械臂不能完成采摘动作。采摘过程中,障碍物的误识别和避障算法本身不足是导致避障失败的主要原因。采摘过程中,采摘机器人识别 1 幅图像的时间为 0.4 s,小于完成一次采摘时间 14 s,满足采摘机器人的实时性采摘要求。

4 结论

(1) 针对自然环境下柑橘采摘机器人障碍物的识别提出了障碍物分类准则,对不规则枝干的识别

提出离散化标记方法,解决了不规则枝干的识别问题。在自然环境下的识别实验中,本文的识别模型对障碍物和可采摘果实的识别综合评价指数为 83.6% 和 91.9%,在 6 300 ~ 76 000 lx 光照强度区间范围内对 5 类目标物的识别综合评价指数均值为 83.13%,在自然环境下有良好的泛化性。

(2) 通过自然环境下定位精度实验验证了 Kinect V2 相机在自然光照条件下的深度测量性能,并得到定位模块的综合定位误差为 5.9 mm。整个识别定位系统实时处理一幅图像的时间为 0.4 s。

(3) 自然环境下识别定位系统实验表明,采摘机器人采摘成功率达到 80.51%,避障成功率达到 75.79%。在满足自然环境下采摘作业的同时,提高了机器人采摘作业过程中的安全性。该系统可应用于其他各类采摘机器人识别定位系统中,具有良好的泛化性。

参 考 文 献

- [1] TIAN Y, YANG G, WANG H, et al. Apple detection during different growth stages in orchards using the improved YOLO-V3 model[J]. Computers and Electronics in Agriculture, 2019, 157: 417-426.
- [2] QINGCHUN F, WEI C, JIANJUN Z, et al. Design of structured-light vision system for tomato harvesting robot[J]. International Journal of Agricultural and Biological Engineering, 2014, 7(2): 19-26.
- [3] ZHAO Y, GONG L, HUANG Y, et al. Robust tomato recognition for robotic harvesting using feature images fusion[J]. Sensors (Basel, Switzerland), 2016, 16(2): 173.
- [4] BAC C W, HEMMING J, VAN HENTEN E J. Robust pixel-based classification of obstacles for robotic harvesting of sweet-pepper[J]. Computers and Electronics in Agriculture, 2013, 96: 148-162.
- [5] 项荣, 应义斌, 蒋焕煜. 田间环境下果蔬采摘快速识别与定位方法研究进展[J/OL]. 农业机械学报, 2013, 44(11): 208-223. XIANG Rong, YING Yibin, JIANG Huanyu. Development of real-time recognition and localization methods for fruits and vegetables in field[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2013, 44(11): 208-223. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20131137&flag=1. DOI: 10.6041/j.issn.1000-1298.2013.11.037. (in Chinese)
- [6] 熊俊涛, 邹湘军, 彭红星, 等. 扰动柑橘采摘的实时识别与采摘点确定技术[J/OL]. 农业机械学报, 2014, 45(8): 38-43. XIONG Juntao, ZOU Xiangjun, PENG Hongxing, et al. Real-time identification and picking point localization of disturbance citrus picking[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(8): 38-43. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20140807&flag=1. DOI: 10.6041/j.issn.1000-1298.2014.08.007. (in Chinese)
- [7] AMATYA S, KARKEE M, GONGAL A, et al. Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting[J]. Biosystems Engineering, 2016, 146: 3-15.
- [8] 李扬. 基于双目视觉的柑橘采摘机器人目标识别及定位技术研究[D]. 重庆: 重庆理工大学, 2017. LI Yang. Research on target recognition and localization of citrus harvesting robot based on binocular vision[D]. Chongqing: Chongqing University of Technology, 2017. (in Chinese)
- [9] 傅隆生, 冯亚利, TOLA Elkanil, 等. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报, 2018, 34(2): 205-211. FU Longsheng, FENG Yali, TOLA Elkanil, et al. Image recognition method of multi-cluster kiwifruit in field based on convolutional neural networks[J]. Transactions of the CSAE, 2018, 34(2): 205-211. (in Chinese)
- [10] 彭红星, 黄博, 邵园园, 等. 自然环境下多类水果采摘目标识别的通用改进 SSD 模型[J]. 农业工程学报, 2018, 34(16): 155-162. PENG Hongxing, HUANG Bo, SHAO Yuanyuan, et al. General improved SSD model for picking object recognition of multiple fruits in natural environment[J]. Transactions of the CSAE, 2018, 34(16): 155-162. (in Chinese)
- [11] 毕松, 高峰, 陈俊文, 等. 基于深度卷积神经网络的柑橘目标识别方法[J/OL]. 农业机械学报, 2019, 50(5): 181-186. BI Song, GAO Feng, CHEN Junwen, et al. Detection method of citrus based on deep convolution neural network[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 181-186. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20190521&journal_id=jcsam. DOI: 10.6041/j.issn.1000-1298.2019.05.021. (in Chinese)

- [14] 陈霓,刘正怀,夏劲松,等.基于 Petri 网模型的收获机轴流式脱分选装置参数化设计[J].农业机械学报,2017,48(11):123-129.
CHEN Ni,LIU Zhenghui,XIA Jingsong,et al. Parametric design for axial flow threshing-separating-cleaning unit based on Petri model[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2017,48(11):123-129. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20171115&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2017.11.015. (in Chinese)
- [15] 庄力骏,孙颖,陈发,等.关于采棉机速比系数 K 值的探讨[J].新疆农机化,2001(2):28-29.
- [16] 中国农业机械化科学研究院.农业机械设计手册[M].北京:中国农业科技出版社,2007.
- [17] 李俊江.弹齿滚筒式棉桃采摘装置的设计研究[D].石河子:石河子大学,2015.
LI Junjiang. Design and research on the cotton boll picking device of spring tooth roller-type[D]. Shihezi: Shihezi University, 2015. (in Chinese)
- [18] LI Teng, FANG Xianfa, WANG Decheng, et al. Tensile strength experiment of seed cotton pulled out of cotton[C]//2018 ASABE Annual International Meeting, 2018.
- [19] BAKER K D, HUGHS E, FOULK J. Cotton quality as affected by changes in spindle speed[J]. Applied Engineering in Agriculture, 2010, 26(3):363-369.
- [20] 李腾,郝付平,韩增德,等.水平摘锭采棉理论分析与试验[J/OL].农业机械学报,2018,49(增刊):233-238.
LI Teng,HAO Fuping,HAN Zengde,et al. Theoretical analysis and experiment of picking cotton with horizontal spindle[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(Supp.):233-238. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2018S031&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2018.S0.31. (in Chinese)
- [21] 镇江农业机械学院.农业机械理论及设计[M].北京:中国工业出版社,1961.

(上接第 22 页)

- [12] ZHANG Q,ZHANG X,ZHANG J, et al. Branch detection for apple trees trained in fruiting wall architecture using depth features and regions-convolutional neural network (R-CNN)[J]. Computers and Electronics in Agriculture, 2018,155:386-393.
- [13] 黄玲涛,王彬,倪涛,等.基于 Kinect 的机器人抓取系统研究[J/OL].农业机械学报,2019,50(1):390-399.
HUANG Lingtao,WANG Bin,NI Tao,et al. Research on robotic grasping system based on Kinect camera[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2019,50(1):390-399. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20190145&flag=1. DOI:10.6041/j.issn.1000-1298.2019.01.045. (in Chinese)
- [14] 王毅,许洪斌,张茂,等.仿蛇嘴咬合式柑橘采摘末端执行器设计与实验[J/OL].农业机械学报,2018,49(10):54-64.
WANG Yi,XU Hongbin,ZHANG Mao,et al. Design and experiment of bite-model end-effector for citrus harvesting by simulating with mouth of snake[J/OL]. Transactions of the Chinese Society for Agricultural Machinery,2018,49(10):54-64. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20181007&flag=1. DOI:10.6041/j.issn.1000-1298.2018.10.007. (in Chinese)
- [15] REDMON J,DIVVALA S,GIRSHICK R, et al. You only look once: unified, real-time object detection[J]. arXiv:1506.02640 [cs.CV], 2015.
- [16] REDMON J,FARHADI A. YOLO9000: better, faster, stronger[C]//Computer Vision and Pattern Recognition,2017. IEEE Conference on. IEEE, 2017:6517-6525.
- [17] JOSEPH R,ALI F. YOLOv3: an incremental improvement[J]. arXiv:1804.02767 [cs.CV]. 2018.
- [18] ZHANG Z Y. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11):1330-1334.
- [19] 徐德,李原,谭民.机器人视觉测量与控制[M].北京:国防工业出版社,2016.
- [20] HARTIGAN J A,WONG M A. Algorithm AS 136: a K-means clustering algorithm[J]. Journal of the Royal Statistical Society, 1979, 28(1):100-108.
- [21] WANG Y,YANG Y,YANG C, et al. End-effector with a bite mode for harvesting citrus fruit in random stalk orientation environment[J]. Computers and Electronics in Agriculture, 2019,157:454-470.
- [22] PETER F, BLOESCH M,RODRIGUEZ D, et al. Kinect v2 for mobile robot navigation: evaluation and modeling[C]//2015 International Conference on Advanced Robotics (ICAR),Istanbul, 2015, 388-394.
- [23] 王维,李焱.基于 RRT 的虚拟人双臂操控规划方法[J].系统仿真学报,2009,21(20):6515-6518.
WANG Wei, LI Yan. RRT-based manipulation planning method for both arms of virtual human[J]. Journal of System Simulation, 2009,21(20):6515-6518. (in Chinese)
- [24] 王毅,张哲,马冀桐,等.采摘机器人奇异位型分析[J].机械传动,2019,43(1):124-130.
WANG Yi,ZHANG Zhe, MA Jitong,et al. Singularity analysis of harvesting robot[J]. Journal of Mechanical Transmission, 2019,43(1):124-130. (in Chinese)