

果园机器人视觉导航行间位姿估计与果树目标定位方法

毕松 王宇豪
(北方工业大学电气与控制工程学院, 北京 100041)

摘要: 针对单目视觉导航中位姿信息不完整和果树定位精度低的问题,提出基于实例分割神经网络的偏航角、横向偏移、果树位置计算方法。首先,基于 Mask R-CNN 模型识别并分割道路与树干;其次,寻找道路掩码凸包并进行霍夫变换,由凸包中的边界方程计算消失点坐标;最后,根据建立的位姿-道路成像几何模型,计算偏航角、横向偏移与果树相对位置。实验结果表明:改进 Mask R-CNN 模型的边框回归平均精确度为 0.564,分割平均精确度为 0.559,平均推理时间为 110 ms。基于本文方法的偏航角估计误差为 2.91%、横向偏移误差为 4.82%,果树横向定位误差为 3.80%,纵向误差为 2.65%。该方法能在不同位姿稳定地提取道路与果树掩码、计算消失点坐标与边界方程,较准确地估计偏航角、横向位移和果树相对位置,为果园环境下的视觉自主导航提供有效参考。

关键词: 果园机器人;视觉自主导航;实例分割;单目相机模型;位姿估计;果树定位

中图分类号: TP242 **文献标识码:** A **文章编号:** 1000-1298(2021)08-0016-11 **OSID:** 

Inter-line Pose Estimation and Fruit Tree Location Method for Orchard Robot

BI Song WANG Yuhao
(College of Electrical and Control Engineering, North China University of Technology, Beijing 100041, China)

Abstract: Because the monocular camera lacks of depth information, it was difficult to use this kind of camera to estimate robot pose in the line of fruit tree and measure distance accurately. Under this research background, a method was proposed to calculate yaw angle, lateral offset, and fruit tree position based on instance segmentation neural network. Firstly, based on the Mask R-CNN model, the road and tree trunks were detected, and their masks were extracted. Secondly, to calculate the vanishing point, the boundary equations were identified based on convex hull and Hough transform, and the vanishing point coordinates were calculated by solving the equation. Finally, according to the established pose-road imaging geometric model, the yaw angle, lateral offset and relative position of the fruit tree were calculated. The experimental results showed that the boundary regression accuracy of the improved Mask R-CNN model was 0.564, the segmentation accuracy was 0.559, and the average inference time was 110 ms. Based on the method, the yaw angle estimation error was 2.91%, and the lateral offset error was 4.82%. For fruit tree positioning, the lateral error was 3.80% and the longitudinal error was 2.65%. At various data collection sites, the method could stably extract road and fruit tree masks, calculate vanishing point coordinates and boundary equations, in addition, the yaw angle, lateral displacement and relative position of fruit trees could be estimated more accurately. Under orchard conditions, it could further improve the visual navigation effect and the intelligent level of agricultural equipment.

Key words: orchard robot; visual autonomous navigation; instance segmentation; monocular camera model; pose estimation; fruit tree location

0 引言

自 1990 年至 2019 年,我国水果产量均位于世界首位^[1],果品及其相关产业已成为农民收入的主要来源,逐渐成为农村经济的重要支柱。目前,高度依赖人工的生产模式和快速增长的人工成本已成为制约水果生产的主要因素^[2],为进一步提高生产效率和水果品质,农业机械装备的智能化是未来发展

收稿日期: 2021-05-09 修回日期: 2021-05-31
基金项目: 北京市自然科学基金项目(4202022)和北方工业大学毓优青年人才培养计划项目
作者简介: 毕松(1983—),男,副教授,博士,主要从事智能农机装备研究,E-mail: bisongo@163.com

的必然趋势。

自主导航是智能农业装备的关键技术之一，可有效降低劳动强度，提高作业效率，广泛应用于采摘、喷药、施肥等任务。视觉导航具有语义信息丰富、探测范围广、成本可控等优点，逐渐成为主流的导航方式，实现果园自主导航的关键在于位姿估计和果树定位，进而引导农业装备自主行进和自主作业。

针对自主行进问题，研究思路集中在两方面：基于道路或天空生成导航线和基于作物检测拟合导航路径。基于道路或天空的导航方法对植株种类、形状、高度具有较高鲁棒性，是国内外学者研究的热点。基于作物检测的导航方法需要准确识别作物主干，对复杂道路环境具有较高的适应性。文献[3-10]利用边界点拟合导航线，根据导航线定性提取航向信息，在特定环境中实现自主导航，但难以准确地定量分析位姿，导航准确性、位姿可控性可进一步提升。

为实现导航信息的定量分析，对基于视觉的位姿估计问题展开研究。文献[11-13]基于单目相机模型实现距离测量，但果园自主导航不仅需要测量道路宽度，也需要估计横向偏移和偏航角，这些方法应用于果园自主导航仍具有一定的局限性。

针对基于单目视觉的果树定位和自主作业方法，国内外学者展开深度研究。文献[14-17]对果树识别方法展开探索，取得了较理想的结果。然而，上述研究停留在果树检测层面，并未深入研究果树定位方法。智能农业装备不仅需要自主行进、识别作物，更需要根据植株位置完成采摘、喷药、施肥等作业任务。

根据以上对果园自主导航研究成果的分析，目前视觉导航的局限性体现在以下两点：①导航信息不完整，以定性视觉自主导航为主，基于视觉的位姿估计方法难以有效提取偏航角和横向位移。②果树定位方法不完善，以果树检测研究为主，难以确定果树的相对位置。

针对现有方法的局限性，本文提出基于 Mask R-CNN 实例分割网络的位姿估计与果树定位方法。在改进并训练实例分割网络的基础上，完成道路掩码和果树掩码的提取，并进一步计算消失点^[18-21]和边界方程。利用树行平行性，建立位姿-道路几何成像模型，确定不同位姿条件下，世界坐标与像素坐标的映射关系，定量分析位姿对图像关键点的影响，基于消失点和边界方程估计偏航信息和横向位移。

1 方法与模型

受光照条件和季节变化影响，种植园场景的色

温、照度、路面环境、植株长势存在较大差异。与人工特征相比，深度神经网络具有更强大的特征提取与识别能力，可准确、稳定地提取复杂场景的语义信息。本文基于实例分割网络模型，提取道路与果树掩码，并进一步提出消失点与边界方程的计算方法。

为使智能农业装备成本可控，本文在果园条件下，研究单目相机成像过程。现代果园的可通行区域近似走廊，两侧树行平行且单侧树行共线。利用树行的平行性和共线性，建立位姿-道路几何成像模型，定量分析位姿、偏航角、边界方程三者的关系，推导果树相对位置的计算方法。

基于以上分析，本文主要工作由 4 部分组成：道路与果树掩码提取、消失点与边界方程计算、偏航角与横向偏移估计、果树相对位置估计。通过建立道路几何成像模型，仅用单目相机实现位姿估计与果树定位，对应的算法流程如图 1 所示。

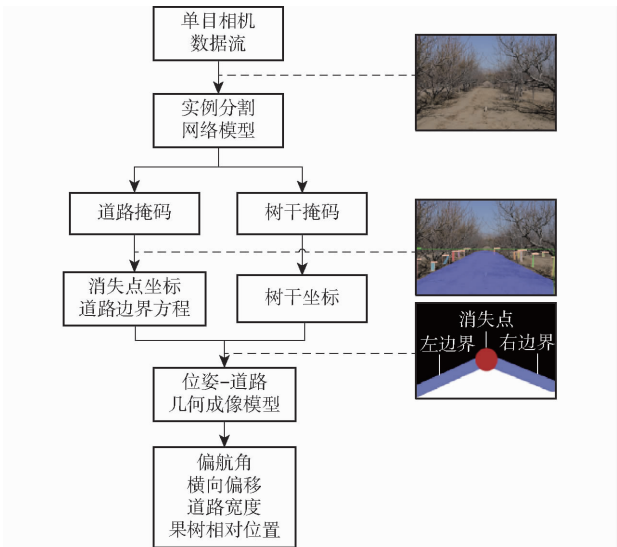


图 1 位姿估计与果树定位方法示意图

Fig. 1 Schematic of pose estimation and fruit tree positioning method

1.1 基于 Mask R-CNN 网络的道路与果树掩码提取方法

深度卷积神经网络不仅可对纹理、形状、轮廓、颜色等低层次特征进行提取，而且可进一步提取高层次抽象特征，具有较强的分类、检测和分割能力。Mask R-CNN 为基于深度卷积神经网络的实例分割模型，具有较高的计算效率和精确率，本文采用该网络完成道路和树干掩码提取，该模型主要由特征提取主干网络、区域建议网络、RoI Align、检测和分割分支组成，方法流程如图 2 所示。

基于该模型提取道路和果树掩码，但由于三维信息在单目相机中的成像过程为小孔成像，目标尺度在图像中表现为近大远小，其尺度存在较大差异性。为了获得多尺寸表示能力，要求特征提取能够

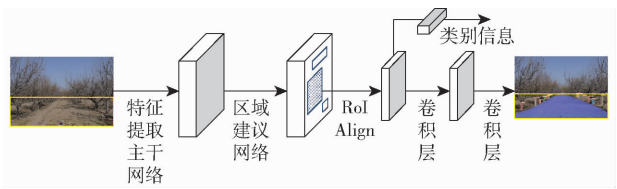


图 2 Mask R - CNN 网络结构示意图

Fig. 2 Schematic of Mask R - CNN network structure

以较大范围的感受野来描述不同尺寸的目标。基于以上原因,本文将 Mask R - CNN 模型的特征提取主干网络改进为 Res2Net^[23] 和特征金字塔网络的组合,与 ResNet^[24] 网络相比,该模型将特征图分成 4 组,除第 1 组特征 X_1 外,每组特征图经 3×3 卷积后,以残差的方式与下一组进行连接,Res2Net 模块结构如图 3 所示。

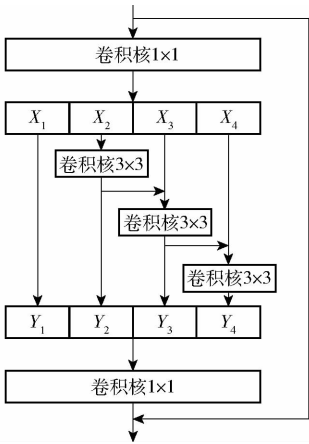


图 3 Res2Net 模块结构示意图

Fig. 3 Schematic of Res2Net module structure

由图 3 可知,除 X_1 特征组外,每个 3×3 的卷积核可以接收来自前层所有特征,其输出的特征图具有更大的感受野。基于该组网方式,Res2Net 的输

出包含不同尺寸及数量的感受野,同时,特征图输出前经过 1×1 卷积运算,可有效融合不同尺度信息。此外,为增强对不同尺寸物体的识别能力,本文在 Res2Net 网络的下一级增加特征金字塔网络,该模型通过提取多尺度的特征信息,并进行深度融合,从而提高检测的精确度。Res2Net 网络和特征金字塔网络结构如图 4 所示。

由图 4 可知,主干网络提取特征的流程如下:首先,原始图像经卷积层和 Res2Net 模块提取不同层次的图像特征;其次,将 C1 ~ C4 阶段的特征图进行 1×1 卷积运算得到特征图 T1 ~ T4,通过卷积上采样并与下一阶段的特征图叠加得到特征金字塔 M1 ~ M4;最后,将 M1 ~ M4 的特征图进行 3×3 卷积运算后,输出特征图 P1 ~ P4 至区域建议网络。

1.2 消失点与边界方程提取方法

针对 Mask R - CNN 分割的道路掩码,研究道路边界和消失点的提取方法。受到栽培精度、植株生长不确定性和行间平整度等因素的影响,可通行区域边界较不平滑,掩码区域的边界通常呈不规则锯齿线,如图 5b 所示。

由图 5b 可知,道路掩码属于非规则多边形,其左右边界呈波浪或锯齿曲线,增加了提取边界线方程和消失点的难度和复杂度。因此,本文首先寻找道路掩码的凸包,如图 5c 所示,掩码 Q 的凸包是一个最小的凸多边形 P ,满足 Q 中的每个点都在 P 的边界或者其内部。其次,基于霍夫变换和角度阈值计算左右边界方程,如图 5d 中绿色线所示,联合边界方程所得的交点坐标为道路消失点,如图 5d 中红色圆点所示。基于以上分析,所提方法由消失点提取、边界方程提取和果树坐标提取组成,具体算法流

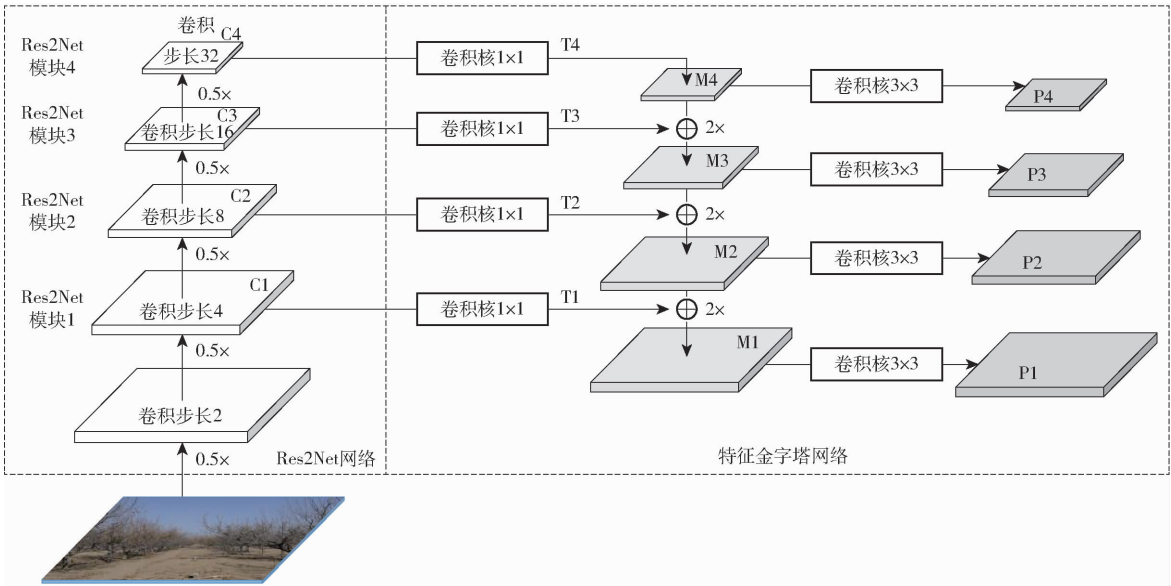


图 4 特征提取主干网络结构示意图

Fig. 4 Schematic of backbone network structure for feature extraction

程如图 6 和图 7 所示。

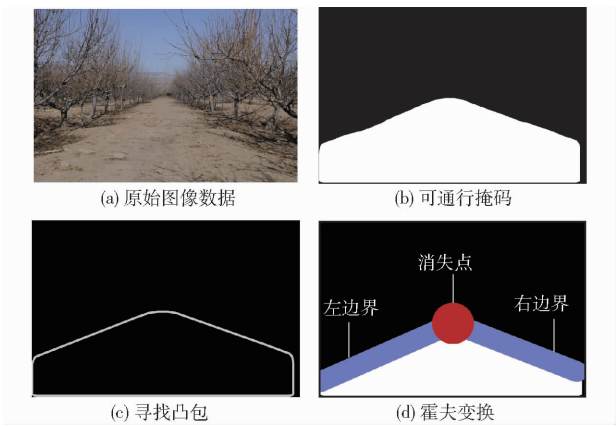


图 5 消失点与边界线提取方法

Fig. 5 Schematics of vanishing point and boundary line extraction method

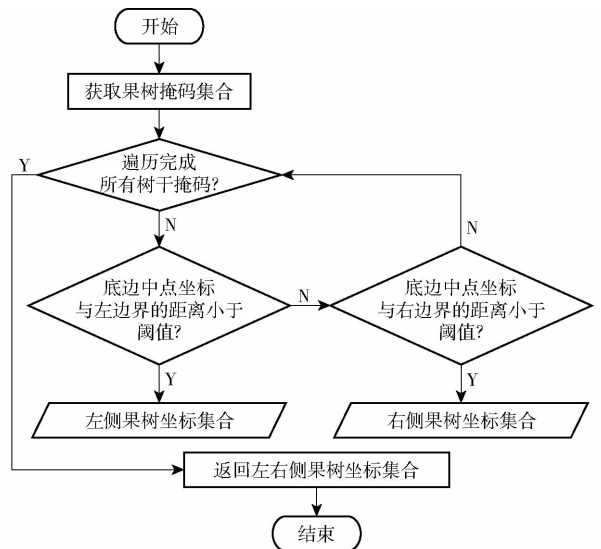


图 7 树干与地面交点的像素坐标计算方法流程图

Fig. 7 Pixel coordinate calculation method of intersection point of tree trunk and ground

1.3 姿态-道路几何成像模型

偏航角、横向偏移和植株相对位置是实现自主导航的关键信息。本文基于单目相机成像原理,利用果园树行的共线性和平行性,间接测量以上信息。现就存在偏航角和横向位移的情况,对单目相机成像过程进行几何建模,其示意图如图 8 所示。

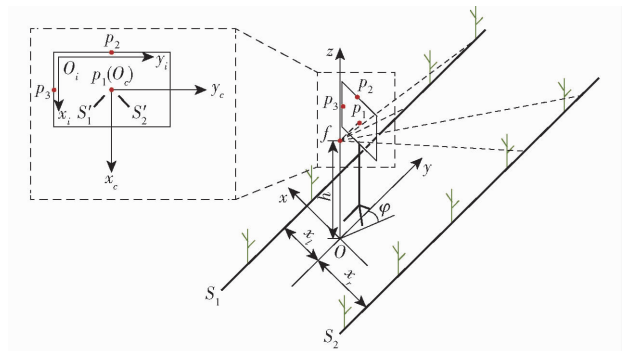


图 8 姿态-道路几何成像模型示意图

Fig. 8 Schematic of pose-road geometry imaging model

图 8 为一般果园场景示意图,左右树行边界 S_1 、 S_2 相互平行,由两树行所构成的空间为可通行区域,基于基本坐标系 $Oxyz$ 建立成像模型,光心 f 距地面高度为 h ,与左右边界的距离分别为 x_l 和 x_r ,偏航角为 φ ,图像坐标系为 $O_c x_c y_c$,像素坐标系为 $O_i x_i y_i$ 。基于果园道路边界的平行性和共线性,建立姿态-道路成像模型的方法如图 9 所示。

根据图 9 可知,建模过程可分为以下步骤:①由像平面关键点求解相机像平面方程。②基于针孔相机成像原理,计算边界线在像平面中的映射方程。③分析方程斜率和消失点表达式,逆解航向信息和横向位移。④基于地面点和像平面的映射关系,逆解果树的相对位置。其中,步骤①、②推导了道路边界在

边界方程、消失点和树干坐标的计算流程如下:首先,根据图 6 所示的边界方程和消失点计算方法,对道路掩码进行轮廓提取、寻找凸包和霍夫变换操作,获取描述边缘轮廓的直线簇,基于设定的斜率和长度阈值条件提取符合要求的边界方程,联立两侧边界线方程提取消失点坐标。其次,根据图 7 所示的树干像素坐标计算方法,提取果树掩码贴近道路边缘的坐标,位于消失点左侧的果树掩码,提取右下方坐标;位于消失右侧的果树掩码,提取左下方坐标;基于树干像素坐标计算距离边界线的距离,若该值大于阈值,则标记为邻行果树,不进行位置估计。

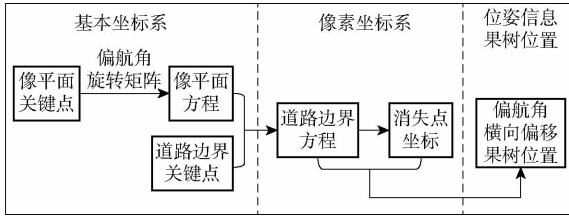


图9 位姿-道路几何成像建模方法

Fig.9 Pose-road geometry imaging modeling method

像平面中的映射方程,在本节展开分析;步骤③、④分析了根据边界方程计算位姿和果树位置。

1.3.1 求解像平面方程

计算无偏航信息时的像平面方程:在相平面任意选取3个关键点 $p_i (i=1,2,3)$,文中坐标均表示为列向量或多维列向量,用于描述相机像平面在 $Oxyz$ 坐标系下的平面方程。

$$p_i = \begin{bmatrix} 0 & 0 & C_w \\ C_F & C_F & C_F \\ H & C_H + H & H \end{bmatrix} \quad (1)$$

式中 C_F ——镜头焦距

C_w ——像平面宽度(相机感光元件宽度)

C_H ——像平面高度(相机感光元件高度)

H ——相机安装高度(焦点与地面距离)

当存在偏航角 φ 时,关键点 p_i 经偏航角旋转矩阵 t 变换为 p'_i

$$p'_i = tp_i = \begin{bmatrix} -C_F \sin \varphi & -C_F \sin \varphi & C_w \cos \varphi - C_F \sin \varphi \\ C_F \cos \varphi & C_F \cos \varphi & C_F \cos \varphi + C_w \sin \varphi \\ H & C_H + H & H \end{bmatrix} \quad (2)$$

其中

$$t = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

式中 p'_i ——旋转变换的关键点

由关键点 p'_i 可确定像平面法向量 n

$$n = p'_2 p'_1 \times p'_3 p'_1 = \begin{bmatrix} C_H C_w \sin \varphi \\ -C_H C_w \cos \varphi \\ 0 \end{bmatrix} \quad (4)$$

再与任一关键点(本文选取 p'_1 点)构成平面点法式方程

$$\begin{aligned} n(1)(x - p'_1(1)) + \\ n(2)(y - p'_1(2)) + n(3)(z - p'_1(3)) &= 0 \\ \Downarrow \\ x \sin \varphi - y \cos \varphi + C_F &= 0 \end{aligned} \quad (5)$$

1.3.2 计算树行边界在像平面中的映射方程

任取树行边界 S_1 两点 r_1, r_2 (边界 S_2 同理),焦

点 f 与 r_1, r_2 构成空间直线方程 l_1 和 l_2 。

$$f = \begin{bmatrix} 0 \\ 0 \\ H \end{bmatrix} \quad (6)$$

$$i=1 \quad i=2$$

$$r_i = \begin{bmatrix} x_i & x_i \\ y_i & y_i \\ 0 & 0 \end{bmatrix} \quad (7)$$

$$l_1: \begin{cases} x = x_i \beta_1 \\ y = y_i \beta_1 \\ z = -H(\beta_1 - 1) \end{cases} \quad (8)$$

$$l_2: \begin{cases} x = x_i \beta_2 \\ y = y_i \beta_2 \\ z = -H(\beta_2 - 1) \end{cases} \quad (9)$$

式中 x_i ——关键点 r 的 x 轴坐标

y_1, y_2 ——关键点 r 的 y 轴坐标

β_1, β_2 ——直线参数方程的参数

联立像平面方程(式(5))与直线方程 l_1 和 l_2 , 解得交点坐标为

$$a_i = \begin{bmatrix} \frac{C_F x_i}{y_1 \cos \varphi - x_i \sin \varphi} & \frac{C_F x_i}{y_2 \cos \varphi - x_i \sin \varphi} \\ \frac{C_F y_1}{y_1 \cos \varphi - x_i \sin \varphi} & \frac{C_F y_2}{y_2 \cos \varphi - x_i \sin \varphi} \\ H - \frac{C_F H}{y_1 \cos \varphi - x_i \sin \varphi} & H - \frac{C_F H}{y_2 \cos \varphi - x_i \sin \varphi} \end{bmatrix} \quad (10)$$

式中 a_i ——直线参数方程与像平面的交点坐标 ($Oxyz$ 坐标系)

将交点的世界坐标转换至图像坐标系,即计算 a_i 在向量 $p'_2 p'_1$ 和向量 $p'_3 p'_1$ 上的投影 a'_i 为

$$a'_i = \begin{bmatrix} \frac{p'_2 p'_1 \cdot p'_1 a_i}{|p'_2 p'_1|} & \frac{p'_2 p'_1 \cdot p'_1 a_i}{|p'_2 p'_1|} \\ \frac{p'_3 p'_1 \cdot p'_1 a_i}{|p'_3 p'_1|} & \frac{p'_3 p'_1 \cdot p'_1 a_i}{|p'_3 p'_1|} \end{bmatrix} =$$

$$i=1 \quad i=2$$

$$\begin{bmatrix} -\frac{C_F(y_1 \sin \varphi + x_i \cos \varphi)}{y_1 \cos \varphi - x_i \sin \varphi} & -\frac{C_F(y_2 \sin \varphi + x_i \cos \varphi)}{y_2 \cos \varphi - x_i \sin \varphi} \\ \frac{C_F H}{y_1 \cos \varphi - x_i \sin \varphi} & \frac{C_F H}{y_2 \cos \varphi - x_i \sin \varphi} \end{bmatrix} \quad (11)$$

式中 a'_i ——直线参数方程与像平面的交点坐标 ($O_c x_c y_c$ 坐标系)

确定 l_1 在 $O_c x_c y_c$ 坐标系内的斜率 k_1 、截距 b_1 和

S_1 在图像坐标系 $O_c x_c y_c$ 的方程 S'_1, S'_2 的计算方法同理。其中, $a'_1(1), a'_1(2)$ 为 $\mathbf{a}'_1$ 向量的第 1 个元素和第 2 个元素, $a'_2(1), a'_2(2)$ 同理。

$$k_l = \frac{a'_2(1) - a'_1(1)}{a'_2(2) - a'_1(2)} = -\frac{x_l}{H \cos \varphi} \quad (12)$$

$$b_l = a'_1(1) - k_l a'_1(2) = -C_F \tan \varphi \quad (13)$$

$$S'_1: x_c = k_l y_c + b_l = -\frac{x_l}{H \cos \varphi} y_c - C_F \tan \varphi \quad (14)$$

$$S'_2: x_c = -\frac{x_r}{H \cos \varphi} y_c - C_F \tan \varphi \quad (15)$$

式中 k_l ——左边界直线方程斜率($O_c x_c y_c$ 坐标系)

b_l ——左边界直线方程截距($O_c x_c y_c$ 坐标系)

在存在偏航角和横向偏移的条件下, 推导两侧树行边界在图像坐标系下的映射方程, 进一步计算消失点坐标, 估计偏航角和横向偏移。也可利用该映射关系, 由像素坐标估计与两侧果树及相机的相对位置。

1.4 位姿估计与果树定位方法

基于建立的位姿-道路几何成像模型, 首先, 根据树行边界线方程计算消失点坐标, 消失点是平行直线经过透视投影后的交点, 基于该点的像素坐标, 推导了偏航角的计算方法。其次, 由像素坐标系下的边界方程, 可解得与左右树行的距离, 横向偏移是当前位置与道路中线的横向距离, 基于与两侧边界的距离, 推导了道路宽度和横向偏移的计算方法。最后, 由树行边界点的三维坐标与像素坐标的映射关系, 推导了果树相对位置的计算方法。

1.4.1 偏航角与横向偏移计算方法

首先, 计算消失点坐标。联立图像坐标系下的边界方程 S'_1 和 S'_2 , 解得消失点(两直线交点)坐标, 即联立式(14)、(15)求解交点坐标, 可得消失点坐标 X_c 和 Y_c 为

$$\begin{cases} X_c = -C_F \tan \varphi \\ Y_c = 0 \end{cases} \quad (16)$$

其次, 推导偏航角计算方法。由式(16)可知, 消失点 x_c 轴坐标 X_c 为偏航角 φ 的一元函数, 因此, 可通过求解消失点坐标 X_c 方程可得偏航角 φ 。在图像坐标系中, 消失点的坐标无法直接获取, 可通过图像-像素坐标系的相互映射关系间接计算

$$X_c = \frac{2X_i - I_x}{I_x} C_w \quad (17)$$

式中 X_i ——消失点 x_i 轴坐标($O_i x_i y_i$ 坐标系)

I_x ——图像横轴分辨率

基于以上分析可知, 通过消失点的像素坐标可计算偏航角估计值 $\hat{\varphi}$, 计算方法为

$$\hat{\varphi} = -\arctan \frac{(2X_i - I_x) C_w}{I_x C_F} \quad (18)$$

再次, 计算与边界 S_1 和 S_2 的距离。以左边界线为例, 由 S'_1 表达式可知, 在图像坐标系和像素坐标系中, 边界线斜率是关于偏航角 φ 和边界距离 x_l 的二元函数, 偏航角 φ 可由式(18)解出, 因此, 基于图像坐标系中边界线斜率可计算距离 x_l 和 x_r 。基于以上分析, 与边界 S_1, S_2 的距离 x_l, x_r 估计值计算式为

$$\hat{x}_l = -\frac{k_l}{H \cos \hat{\varphi}} \quad (19)$$

$$\hat{x}_r = -\frac{k_r}{H \cos \hat{\varphi}} \quad (20)$$

式中 $\hat{x}_l$ ——左边界距离估计值

$\hat{x}_r$ ——右边界距离估计值

最后, 计算道路宽度和横向偏移。由于 x_l, x_r 为距离左右边界的距离, 二者绝对值之和为道路宽度 W , 二者绝对值之差的一半为横向偏移 λ , 估计值计算式为

$$\hat{W} = |\hat{x}_l| + |\hat{x}_r| \quad (21)$$

$$\hat{\lambda} = \frac{|\hat{x}_l| - |\hat{x}_r|}{2} \quad (22)$$

式中 $\hat{W}$ ——道路宽度估计值

$\hat{\lambda}$ ——横向位移估计值

1.4.2 果树相对位置计算方法

由图 8 可知, 在基本坐标系中, S_1 和 S_2 为树行和可通行区域的交界, 因此位于边界 S_1 和 S_2 的果树 x 轴坐标分别为 x_l 和 x_r , 只需求解 y 轴坐标即可估计果树的相对位置, 因为果树的像素坐标可以从图像中获得, 可利用式(11)所示的映射关系求解 y 轴坐标。因为安装高度 h 由测量得出, 一般带有测量误差, 为避免引入二次误差, 采用 y_c 轴坐标求解 y , 将 $\mathbf{a}'_i$ 的 y_c 轴坐标(式(11))一般化, 由式(17)和

$$x_c = -\frac{C_F (y \sin \varphi + x_i \cos \varphi)}{y \cos \varphi - x_i \sin \varphi} \quad (23)$$

可得果树的 y_c 轴与 y 轴坐标的映射关系, 即 y 轴坐标的计算方法为

$$\hat{y} = \begin{cases} \frac{\hat{x}_l [(x_i - I_x) C_w \tan \hat{\varphi} - C_F I_x]}{(x_i - I_x) C_w - C_F I_x \tan \hat{\varphi}} & (x_i \leq X_i) \\ \frac{\hat{x}_r [(x_i - I_x) C_w \tan \hat{\varphi} - C_F I_x]}{(x_i - I_x) C_w - C_F I_x \tan \hat{\varphi}} & (x_i > X_i) \end{cases} \quad (24)$$

式中 $\hat{y}$ ——果树 y 轴坐标($Oxyz$ 坐标系)

1.4.3 算法流程

基于对偏航角、横向位移和位姿-道路几何成像模型的分析, 可由图像中的消失点、树行边界方程估计当前位姿, 计算偏航角和横向偏移。由树干的像素坐标计算果树与相机的相对位姿。若将单目相机刚性连接于农业机械装备, 可通过图像中的关键点完成位姿估计和果树定位, 计算流程如图 10 所示。

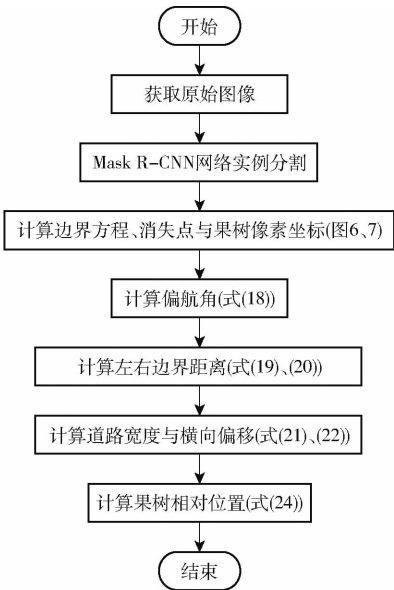


图 10 位姿估计与果树定位方法流程图
Fig. 10 Pose estimation and fruit tree positioning method

2 实验与结果分析

为了验证本文提出位姿估计和果树定位方法的准确性和鲁棒性,在训练、测试实例分割网络 Mask R - CNN 的基础上,进一步计算消失点与边界方程,完成了偏航角、横向偏移、果树相对位置计算实验,并验证本文所提方法的有效性。

2.1 实验条件

通过安装在三角架的单目相机抓取果园图像,偏航角由惯性测量单元的加速度计和陀螺仪融合滤波得出,通过卷尺测量得到横向偏移与果树相对位置。三脚架与 InvenSense MPU6050 型惯性测量单元刚性连接,图像采集设备为 Nikon D90 型相机,镜头焦距为 35 mm,相机水平视场角为 37.2°,感光元件尺寸为 23.6 mm × 15.8 mm,相机与三脚架螺纹连接,安装高度为 0.95 m。

数据采集地位于山西省太原市尖草坪区向阳店苹果种植园,该种植园属于半结构化场景,平均行距为 5 m,受光照、水肥、土壤等因素影响,果树生长具有较大随机性,其生长状况和株距也存在一定差异性。该苹果种植园中,果树树干与地面的颜色相似度较高,且可通行区域界线相对模糊,增加了道路与树干信息的提取难度,数据集包含多种拍摄角度、多种拍摄位置的图像数据,具有较强的代表性。

2.2 模型训练与验证

从相机中随机选取 1 200 幅图像制作数据集,其中 720 幅用于训练,240 幅用于验证,240 幅用于测试,采用 Labelme 工具标注树干和道路可通行区

域。为提升训练和预测速度,图像预处理时,将图像长边的分辨率统一转换为 1 280 像素。为提高模型鲁棒性,抑制过拟合现象,本文训练时加入随机扰动扩充数据量,例如随机调整对比度、饱和度、亮度等。训练平台为 Intel i7 - 6870K CPU, Nvidia Geforce GTX1080Ti × 4 GPU, 32 GB RAM。

为使区域建议网络生成的预测框更贴近目标最小外接矩,基于 K-means 方法聚类标注框的尺寸,为抑制不同尺寸的标注框对聚类误差的影响,本文选取交并比度量类别的相似性。聚类中心数量越多,与标注框的交并比越大,但计算量也相应增加;聚类中心数量较少时,数据处理量随之减少,与标注框的交并比也随之降低。聚类中心数量对应的平均交并比如图 11 所示。

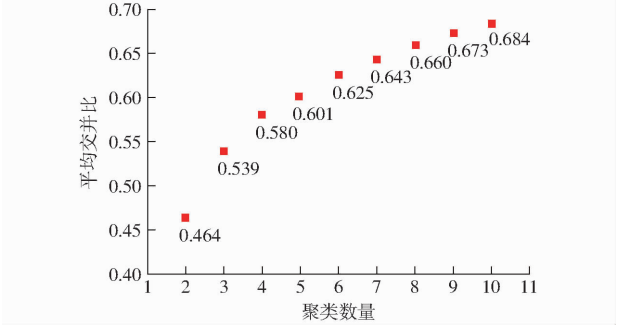


图 11 平均交并比随聚类数量变化散点图
Fig. 11 Scatter plot of average IoU and number of clusters

由图 11 可知,聚类数量从 2 类增加为 3 类时,平均交并比提升显著。由于数据处理量随着预测框的增加而明显增加,聚类数量等于 3 时的综合性能较高,此时聚类中心的高宽比分别为 0.7、3 和 8,并根据该比例调整区域建议网络生成的预测框。

在训练过程中,优化方法选择 Adam 自适应学习率优化器,其中 GPU 集群单次迭代样本数为 4,初始学习率为 0.005,学习率衰减系数为 0.1,最小学习率为 1×10^{-6} ,总计训练 100 周期,训练时的平均损失值和验证集的平均精确度如图 12 所示。

由图 12a 可知,训练 76 周期时,边框回归平均精确度为 0.564,分割平均精确度为 0.559,且损失函数值较低,模型具有较强的分割和识别能力。此外,本文对比了以 ResNet50 为主干网络的 Mask R - CNN 模型,其损失函数曲线与平均精确度曲线如图 12b 所示,从图中可知其损失函数收敛速度、模型性能均弱于 Res2Net,详细数据如表 1 所示。

由表 1 可知,Mask R - CNN(Res2Net50)的边框回归与分割具有更高精确度,分别提升 4.5% 和 4%,推理速度提升 20%;此外,基于该主干网络的 Mask R - CNN 对中小目标的检测和分割效果提升

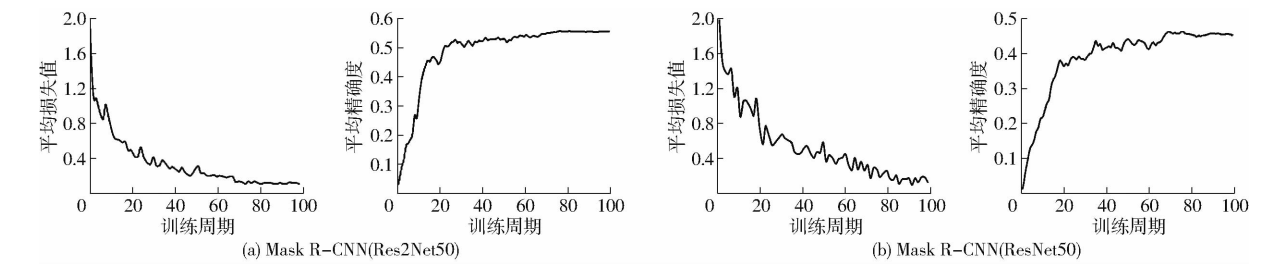


图 12 训练过程曲线
Fig. 12 Training process curves

表 1 2 种主干网络的 Mask R - CNN 模型性能对比															
Tab.1 Comparison of Mask R - CNN model performance of two backbones															
网络结构	boxmAP	maskmAP	boxAP			boxAR			maskAP			maskAR			计算 时间/ ms
			大目标	中等 目标	小目标	大目标	中等 目标	小目标	大目标	中等 目标	小目标	大目标	中等 目标	小目标	
Mask R - CNN (ResNet50)	0. 519	0. 519	0. 668	0. 290	0. 095	0. 719	0. 383	0. 139	0. 699	0. 295	0. 065	0. 728	0. 382	0. 112	138
Mask R - CNN (Res2Net50)	0. 564	0. 559	0. 689	0. 368	0. 124	0. 743	0. 438	0. 165	0. 716	0. 397	0. 113	0. 742	0. 437	0. 156	110

注：boxmAP 为边界框平均精确度均值，maskmAP 为掩码平均精确度均值，boxAP 为边界框平均精确度，boxAR 为边界框平均召回率，maskAP 为掩码平均精确度，boxAP 为掩码平均召回率。

显著,受成像原理、拍摄角度、拍摄距离与单目成像原理等因素的影响,图像中的目标尺度具有较大差异,而 Res2Net 模型采用分组卷积方法,可有效提取多尺度目标的特征,在本文试验条件中表现出更高的适应性。

基于神经网络分割的道路与树干掩码,提取消失点、边界方程、果树坐标等更深层次的信息;对于

果树掩码,分类其所属位置:左行、右行和邻行,并对左右行的树干掩码提取贴近道路边界的坐标;对于道路可通行区域掩码,依次完成寻找凸包、霍夫变换、计算边界和消失点坐标操作,其中红色圆点表示消失点,绿色直线表示可通行区域的左右边界,实验结果如表 2 所示,其中,左上角为坐标原点,水平轴为 u , 竖直轴为 v 。

表 2 消失点与边界提取结果									
Tab. 2 Vanishing point and boundary extraction results									
项目	采集点 1	采集点 2	采集点 3	采集点 4	采集点 5	采集点 6	采集点 7	采集点 8	采集点 9
实例分割									
道路掩码									
消失点 边界线									
消失点坐标	[642,495]	[646,247]	[612,361]	[423,509]	[804,426]	[1 107,461]	[361,444]	[772,212]	[339,437]
左边界方程	$u=-2.2v+1\ 755.7$	$u=-1.8v+1\ 085.2$	$u=-2.9v+1\ 661.4$	$u=-2.2v+1\ 543.3$	$u=-1.9v+1\ 606.8$	$u=-2.9v+2\ 446.1$	$u=-1.8v+1\ 447.4$	$u=-1.9v+1\ 172.2$	$u=-2.7v+1\ 506.7$
右边界方程	$u=2.5v-613.2$	$u=3.1v-115.9$	$u=1.5v+76.6$	$u=2.6v-901.1$	$u=2.5v-255.3$	$u=1.4v+482.8$	$u=3.5v-1\ 191.3$	$u=2.7v+205.5$	$u=2.1v-557.1$

注:第 3 行的图像中,红色圆点为消失点,紫色边线为道路边界。

由表 2 可知,在不同拍摄位置和拍摄视角条件下,本文采用的网络模型可较稳定且准确地完成实例分割任务,采用的消失点和边界线提取方法具有较高的鲁棒性。存在偏航角和横向偏移时,本文所用方法仍能在不同位姿条件下,稳定地提取道路与果树掩码、消失点坐标、道路边界方程,为计算位姿与果树相对位置提供可靠基础。

2.3 偏航角与横向偏移估计

位姿估计是农业装备在半结构场景下自主导航的前提和基础,利用建立的道路成像几何模型,仅用单目相机即可实现偏航角与横向偏移的估计。基于消失点像素坐标、相机感光元件和镜头焦距可得当前航向信息,再由图像中的边界线方程和偏航角,估计与左右树行的距离,并估计行距与横向偏移。其

中,偏航角、横向偏移与道路宽度的误差计算方法分别为

$$E_{\varphi} = \frac{|\hat{\varphi} - \varphi|}{A_h} \times 100\% \tag{25}$$

$$E_{\lambda} = \frac{|\hat{\lambda} - \lambda|}{W} \times 100\% \tag{26}$$

$$E_w = \frac{|\hat{W} - W|}{W} \times 100\% \tag{27}$$

式中 E_{φ} ——偏航角估计误差
 E_{λ} ——横向偏移估计误差

E_w ——道路宽度估计误差
 A_h ——水平视场角,取 37.2°

一般工况下,智能农业装备可能位于道路中间,也可能位于中线左侧或右侧。同理,偏航角也存在3种情况,既可能无偏航信息,也可能向左或向右偏航。为验证模型的鲁棒性与可靠性,本文分析了农业装备工作时的9种位姿,即3种横向位置和偏航角的交叉组合。实验结果如表3所示。

由表3可知,本文方法在存在偏航角和横向偏

表 3 位姿估计及其误差
Tab.3 Pose estimation and its error

采集点	真实值			估计值			估计误差			计算 时间/s
	$\varphi/(^{\circ})$	λ/mm	W/mm	$\hat{\varphi}/(^{\circ})$	$\hat{\lambda}/\text{mm}$	$\hat{W}/\text{mm}$	$E_{\varphi}/\%$	$E_{\lambda}/\%$	$E_w/\%$	
1	0	0	5 000	0.060	152	5 037	0.16	3.05	0.75	0.059
2	0	860	5 000	0.181	687	5 102	0.49	3.47	2.03	0.058
3	0	-1 054	5 000	-0.845	-747	4 615	2.27	6.15	7.71	0.057
4	-6.0	0	5 000	-6.521	214	5 087	1.40	4.28	1.74	0.063
5	5.0	0	5 000	4.938	318	4 612	0.17	6.37	7.75	0.065
6	13.7	-613	5 000	13.821	-840	4 615	0.32	4.54	7.71	0.053
7	-14.5	677	5 000	-8.361	916	5 593	16.50	4.78	11.87	0.068
8	4.0	500	5 000	3.978	418	4 808	0.06	1.63	3.84	0.060
9	-10.8	-785	5 000	-9.010	-330	5 031	4.81	9.10	0.61	0.050

移时均能较准确地计算位姿,其中,偏航角平均误差为2.91%,横向偏移平均误差为4.82%,道路宽度平均误差为4.89%,其精度可满足农业装备导航需求。

在此基础上,研究所处位姿对算法精度的影响,将偏航角与横向偏移取绝对值进行Min-Max归一化,对于偏航角误差,与当前偏航角、横向位移的协方差分别为0.069和0.035,对于横向偏移误差,相应的协方差分别为0.032和0.026,表明本文方法的精度受所处位姿的影响较小,具有较高的鲁棒性和适应性。

2.4 果树相对位置估计

农业自主导航任务不仅需要准确估计当前位姿,为轨迹规划和自主运行提供基础,更重要的是,在获取当前位姿的基础上,确定果树的相对位置,为智能化采摘、喷药和施肥等精细化操作提供保障。本文根据边界线像素-基本坐标系的相互映射关系,即根据式(19)~(24)计算左右2行各4棵果树的相对位置,其误差计算方法为

$$E_{xl} = \frac{|\hat{x}_l - x_l|}{x_l} \times 100\% \tag{27}$$

$$E_{xr} = \frac{|\hat{x}_r - x_r|}{x_r} \times 100\% \tag{28}$$

$$E_y = \frac{|\hat{y} - y|}{y} \times 100\% \tag{29}$$

$$E_d = \frac{|\hat{d} - d|}{d} \times 100\% \tag{30}$$

式中 E_{xl} ——左侧果树横向位置(x 轴)估计误差
 E_{xr} ——右侧果树横向位置(x 轴)估计误差
 E_y ——纵向位置(y 轴)估计误差
 E_d ——株距估计误差
 d ——株距真实值
 $\hat{d}$ ——株距估计值

由于相机成像模型为射影变换,成像过程为三维信息向二维平面的映射,无法直接表示深度信息,当存在偏航角和横向位移时,加剧了果树的层叠现象,存在层叠现象的实验场景如图13所示。



图 13 存在层叠现象的实验数据

Fig.13 Experimental data with cascading phenomenon

由于图13的采集位置位于道路中线偏左,图像中左侧果树存在较明显的层叠现象。除距离较近的两棵果树层次结构较明显外,其余果树遮挡现象明显,增加了果树识别与定位难度。因此,本文在道路中线选取3处航向角为正前的实验点进行测试,分

别测量每棵果树距离相机的横向相对位置、纵向相对位置与株距,并与估计值对比,计算其误差,实验场景如图 14 所示。



图 14 果树位置估计实验场景

Fig. 14 Experimental site for fruit tree position estimation

图 14 中的 3 处实验点位于同一行间的不同纵向位置,实验点 1 距离道路尽头最远,实验点 2 次之,实验点 3 距离道路尽头最近,3 处实验点的位置示意图如图 15 所示。

本文在 3 处实验点测量 20 m 以内果树的相对

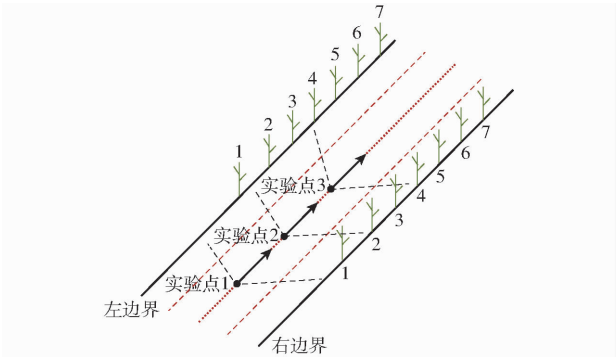


图 15 实验点位置示意图

Fig. 15 Schematic of experimental point location

位置以及与前株的株距,由图 15 所示,在实验点 1 分析编号为 1~4 的果树,在实验点 2 分析编号为 2~5 的果树,实验点 3 分析编号为 4~7 的果树,相应的实验数据如表 4~6 所示。

表 4 实验点 1 果树位置估计结果

Tab. 4 Estimation results of position of fruit trees in experimental site 1

果树 序号	横向位置(左行)			纵向位置(左行)			左行株距(前株)			果树 序号	横向位置(右行)			纵向位置(右行)			右行株距(前株)		
	x_l/m	$\hat{x}_l/m$	$E_{xl}/\%$	y/m	$\hat{y}/m$	$E_y/\%$	d/m	$\hat{d}/m$	$E_d/\%$		x_r	$\hat{x}_r$	$E_{xr}/\%$	y	$\hat{y}$	$E_y/\%$	d	$\hat{d}$	$E_d/\%$
4				20.78	20.92	0.67	2.95	2.90	1.69	4				17.33	16.69	3.69	1.87	1.76	5.88
3	2.50	2.68	7.20	17.83	18.02	1.07	4.58	5.21	13.76	3	2.50	2.54	1.60	15.46	14.93	3.43	2.73	2.50	8.42
2				13.25	12.81	3.32	2.75	2.28	17.09	2				12.73	12.43	2.36	3.93	3.33	15.27
1				10.50	10.53	0.29				1				8.80	9.10	3.41			

表 5 实验点 2 果树位置估计结果

Tab. 5 Estimation results of position of fruit trees in experimental site 2

果树 序号	横向位置(左行)			纵向位置(左行)			左行株距(前株)			果树 序号	横向位置(右行)			纵向位置(右行)			右行株距(前株)		
	x_l/m	$\hat{x}_l/m$	$E_{xl}/\%$	y/m	$\hat{y}/m$	$E_y/\%$	d/m	$\hat{d}/m$	$E_d/\%$		x_r	$\hat{x}_r$	$E_{xr}/\%$	y	$\hat{y}$	$E_y/\%$	d	$\hat{d}$	$E_d/\%$
5				18.64	18.26	2.04	2.71	2.82	4.06	5				16.81	17.48	3.99	3.61	3.89	7.76
4	2.50	2.37	5.20	15.93	15.44	3.08	2.95	2.53	14.24	4	2.50	2.67	6.80	13.20	13.59	2.95	1.87	1.83	2.14
3				12.98	12.91	0.54	4.58	4.60	0.44	3				11.33	11.76	3.80	2.73	2.46	9.89
2				8.40	8.31	1.07	2.75			2				8.60	9.30	8.14	3.93		

表 6 实验点 3 果树位置估计结果

Tab. 6 Estimation results of position of fruit trees in experimental site 3

果树 序号	横向位置(左行)			纵向位置(左行)			左行株距(前株)			果树 序号	横向位置(右行)			纵向位置(右行)			右行株距(前株)		
	x_l/m	$\hat{x}_l/m$	$E_{xl}/\%$	y/m	$\hat{y}/m$	$E_y/\%$	d/m	$\hat{d}/m$	$E_d/\%$		x_r	$\hat{x}_r$	$E_{xr}/\%$	y	$\hat{y}$	$E_y/\%$	d	$\hat{d}$	$E_d/\%$
7				17.04	16.51	3.11	3.98	3.88	2.51	7				16.41	16.05	2.19	1.51	1.37	9.27
6	2.50	2.51	0.40	13.06	12.63	3.29	2.71	2.62	3.32	6	2.50	2.46	1.60	14.90	14.68	1.48	3.69	3.56	3.52
5				10.35	10.01	3.29	2.95	2.41	18.31	5				11.21	11.12	0.80	3.61	3.30	8.59
4				7.40	7.60	2.70	4.58			4				7.60	7.82	2.89	1.87		

由表 4~6 可知,本文提出的果树相对位置估计方法,仅用单目相机即可较准确地估计果树位置,对最近探测距离到 20 m 以内的果树均能有效推测其位置,横向位置的平均误差为 3.80%,纵向位置的平均误差为 2.65%;此外,利用果树的相对位置可进一步估计株距,其平均误差为 8.12%。

3 讨论

(1)提出基于改进 Mask R-CNN 网络的道路与树干识别方法,与 Mask R-CNN(ResNet50)相比,边框回归精确率提升 4.5%,分割精确率提升 4.0%,推理速度提升 20%。基于道路掩码可进一步计算消失点坐标和边界方程,对采集位姿、边界

线的粗糙度具有较高的鲁棒性。基于提出的位姿-道路几何成像模型,利用消失点与边界方程,实现位姿估计与果树定位,偏航角、横向位移、果树位置的计算误差较低,且可提取不同位姿条件下的导航信息。

(2)误差主要来源于测量误差和相机系统误差。首先,用于测量偏航角的惯性导航单元存在累计误差,数据采集过程中的振动加剧了数据波动程度,因此,偏航角真值存在一定的误差。其次,本文采用相机并未经过标定,存在轻微畸变现象,因此,基于道路掩码提取的消失点坐标和边界方程也存在偏移,一定程度影响了位姿估计与果树定位精度。

(3)在路面平整度较低的种植园运行过程中,受道路不平度的影响,相机存在不同程度抖动问题,导致图像存在模糊的可能,可通过 Richardson-Lucy

算法抑制图像模糊。目前,果园作业仍对人工有一定的需求,可利用神经网络模型识别作业人员及其相对位置,基于数据帧差信息分析人员的运动趋势,进而实现更加智能化的行间自主导航。

4 结论

(1)提出基于改进 Mask R-CNN 网络的道路与树干识别方法。边框回归平均精确度为 0.564,目标分割平均精确度为 0.559,具有较强的抗干扰能力和更高的识别精度。

(2)提出基于位姿-道路几何成像模型的位姿估计与果树定位方法。偏航角、横向位移误差为 2.91% 和 4.82%,果树位置的横、纵向误差为 3.80% 和 2.65%,可提取不同位姿条件下的导航信息,具有较高的准确性和稳定性。

参 考 文 献

- [1] 国家统计局. 国际统计年鉴[M]. 北京:中国统计出版社,2021.
- [2] 国家统计局. 中国统计年鉴[EB/OL]. 2020. <http://www.stats.gov.cn/tjsj/ndsj/2020/indexch.htm>.
- [3] 张成涛,谭斌,吴刚,等. 基于达芬奇平台的联合收获机视觉导航系统路径识别[J/OL]. 农业机械学报,2012,43(增刊): 271-276.
ZHANG Chengtao, TAN Yu, WU Gang, et al. Visual navigation system path recognition algorithm based on DaVinci platform for combine harvester[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(Supp.): 271-276. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=2012s56. DOI:10.6041/j.issn.1000-1298.2012.S0.056. (in Chinese)
- [4] RADCLIFFE J, COX J, BULANON D M. Machine vision for orchard navigation[J]. Computers in Industry, 2018, 98:165-171.
- [5] SHARIFI M, CHEN X Q. A novel vision based row guidance approach for navigation of agricultural mobile robots in orchards[C]// IEEE International Conference on Automation, Robotics and Applications, 2015:251-255.
- [6] 韩振浩,李佳,苑严伟,等. 基于 U-Net 网络的果园视觉导航路径识别方法[J/OL]. 农业机械学报,2021,52(1):30-39.
HAN Zhenhao, LI Jia, YUAN Yanwei, et al. Path recognition of orchard visual navigation based on U-Net[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 30-39. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20210104&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2021.01.004. (in Chinese)
- [7] 关卓怀,陈科尹,丁幼春,等. 水稻收获作业视觉导航路径提取方法[J/OL]. 农业机械学报,2020,51(1):19-28.
GUAN Zhuohuai, CHEN Keyin, DING Youchun, et al. Visual navigation path extraction method in rice harvesting[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(1): 19-28. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20200102. DOI:10.6041/j.issn.1000-1298.2020.01.002. (in Chinese)
- [8] CHOI K H, HAN S K, HAN S H, et al. Morphology-based guidance line extraction for an autonomous weeding robot in paddy fields[J]. Computers and Electronics in Agriculture, 2015, 113:266-274.
- [9] MONTALVO M, PAJARES G, GUERRERO J M, et al. Automatic detection of crop rows in maize fields with high weeds pressure[J]. Expert Systems with Applications, 2012, 39(15):11889-11897.
- [10] 杨洋,张亚兰,苗伟,等. 基于卷积神经网络的玉米根茎精确识别与定位研究[J/OL]. 农业机械学报,2018,49(10):46-53.
YANG Yang, ZHANG Yalan, MIAO Wei, et al. Accurate identification and location of corn rhizome based on Faster R-CNN[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(10): 46-53. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20181006&flag=1. DOI:10.6041/j.issn.1000-1298.2018.10.006. (in Chinese)
- [11] 刘军,后士浩,张凯,等. 基于单目视觉车辆姿态角估计和逆透视变换的车距测量[J]. 农业工程学报,2018,34(13): 70-76.
LIU Jun, HOU Shihao, ZHANG Kai, et al. Vehicle distance measurement with implementation of vehicle attitude angle estimation and inverse perspective mapping based on monocular vision[J]. Transactions of the CSAE, 2018, 34(13): 70-76. (in Chinese)
- [12] 颜松,姚立健,曾松伟,等. 基于位姿状态的林区道路视觉导航[J]. 林业工程学报,2021,6(1):163-170.
YAN Song, YAO Lijian, ZENG Songwei, et al. Visual navigation of forest road based on pose state[J]. Journal of Forestry Engineering, 2021, 6(1):163-170. (in Chinese)