

基于单目视觉与改进势场法的避障算法

纪宇航,蔡文靖,刘鑫,王礼贺
(中电科光电科技有限公司,北京 100015)

摘要:针对无人平台在未知环境缺乏全局障碍物位置信息难以进行实时避障的问题,提出一种融合深度神经网络和改进人工势场的动态避障方法。首先利用 YOLOv5s 轻量目标检测网络和轻量化深度估计网络构建障碍感知模块,探测障碍物位置和深度;其次使用目标框和等效深度描述周围障碍物的三维信息;然后将自身平台投影到图像平面形成核心区,并根据障碍物的等效深度网格与核心区的位置关系计算核心区在像平面虚拟势场受力方向、所需要的偏航角和线速度;最后控制系统接收到信号后引导无人平台转向或制动,使核心区内部深度大于设定的安全距离完成动态避障。实验基于单目可见光和红外进行避障测试,结果表明视觉感知模块可以准确探测常见障碍物的位置和深度,核心区能够直观地反映无人平台与障碍物的位置关系,方法仅依赖单目传感器便可实现有效避障,相比于传统方法成本低廉、部署灵活,为无人平台在日间、夜间的未知环境进行避障提供了新思路。

关键词:深度学习;动态避障;人工势场

中图分类号:TN29;TP274 **文献标识码:**A **DOI:**10.3969/j.issn.1001-5078.2024.06.021

Obstacle avoidance algorithm based on monocular vision and improved potential field

Ji Yu-hang, Cai Wen-jing, Liu Xin, Wang Li-he
(CETC Electro-Optics Technology Co. Ltd., Beijing 100015, China)

Abstract: In unknown environments without global obstacle location information, real-time avoidance is a challenging task for unmanned platforms. To address this issue, a method that fuses deep neural networks (DNNs) with an improved artificial potential field (APF) algorithm is proposed in this paper. Firstly, YOLOv5s and a lightweight depth estimation model is used to construct an obstacle perception module to detect the location and depth of obstacles. Then, the target frame and equivalent depth are utilized to describe the three-dimensional information of the surrounding obstacles. Subsequently, the platform is projected onto the image plane, and the core area is calculated based on the positional relationship between the obstacle's equivalent depth grid and the core area to compute the direction of the core area in terms of the force on the virtual potential field in the image plane, the required yaw angle, and the linear velocity. Finally, the control system receives the signals and guides the unmanned platform to steer or brake, ensuring that the internal depth of the core area is greater than the safety distance. The experiment is based on monocular visible light and infrared for obstacle avoidance test, and the results demonstrate that the perception module can accurately detect the location and depth of common obstacles and the core area can intuitively reflect the positional relationship between the unmanned platform and surrounding obstacles. Compared to traditional methods, the proposed method relies on a monocular sensor alone to effectively avoid obstacles, achieving lower cost and flexible deployment,

which provides a new idea for unmanned platforms to avoid obstacles in unknown environments during daytime and nighttime.

Keywords: deep learning; dynamic avoidance; artificial potential field

1 引言

无人系统能够完成侦查监视、火力打击和排雷布障等军事任务,在现代化战争中的作用日益重要,正朝信息化、智能化方向迅猛发展^[1-2]。避障算法通过感知周围运动障碍物,提供实时环境信息,进而传递至决策控制系统进行避障操作,有效保障了无人系统行进过程中的安全性^[3],因此无人平台避障算法具有重要的研究意义。

传统避障方法可分为基于局部路径规划与基于局部防碰撞 2 类^[4],前者方法假设全局障碍物位置已知,随后规划出一条由起始点到目标点的可达路径,如遗传^[5]、智能群^[6]和 A^*/D^* ^[7-8]等启发算法,适用于规避静态障碍物;后者方法不依赖于全局信息,能够利用传感器探测障碍物进行实时避障,常见方法有基于制导率、基于速度障碍法^[9]和基于人工势场法^[10]等。基于光流的方法^[11-13]利用光流追踪像素在连续帧间检测运动障碍物,再结合路径规划算法实现避障。虽然在特定场景下可以完成避障任务,但在稳定性和适用性方面易受环境影响,且计算过程较为复杂。近年来,随着深度学习理论的发展,使用深度神经网络(Deep Neural Network, DNN)自适应提取图像深度特征在目标检测、分割和深度预测等视觉任务中取得较好表现,并在无人系统避障方面取得一定研究进展。文献[14]依据目标边框和距离将图像纵切为不同深度的区域,利用运动约束和评价函数筛选出可航行区域,进而规划运动,实现了单目动态避障。文献[15]提出了一种基于双目立体视觉的避障算法,该方法通过在双目中匹配、检测出同一障碍物中心点坐标计算视差,利用三角形关系解出目标障碍物中心深度实现避障,但该方法难以判断出距离障碍物的最近深度。文献[16]基于语义分割结果分析 6 个预设区域的可通行性,进而指导机器人完成避障。文献[17]建立模型预测动态障碍物质心可达区域,进而借助多步椭圆包络场和新 sigmoid 场引导机器人完成避障。基于深度强化学习的避障方法,如深度 Q 学习网络^[18-19]等,是一种建立在奖惩机制之上根据已发动作指令的反馈学习避障决策的方法,使无人系统逐渐具备

全局规划与路径寻优的能力。然而需要根据实际任务和传感器数据设计合理的奖励函数,受限于样本规模和学习效率,目前大多数研究主要集中于仿真环境。此外,基于主动式传感器的方法^[20-24],如利用 Lidar、RGBD 深度相机、超声波等作为避障算法的源数据,虽然能够提供精确地障碍物距离信息,但因探测范围有限或成本较高而应用场景受限。

单目相机具备便捷、低成本和隐蔽性强等优势。基于此,本文融合 DNN 和人工势场设计了一种高效的单目实时避障算法。通过利用 YOLOv5s 和 Lite-Mono-8m 两种轻量化卷积神经网络实现障碍目标的位置和距离探测,然后基于视觉感知结果制定避障策略。先将无人平台几何体投影到成像平面与视觉感知结果统一在同一参考框架中,然后引入改进的人工势场确定其与障碍物之间的位置关系,从而指导无人平台进行实时避障。本文基于可见光和红外图像进行测试,实验表明,该算法仅利用单目相机便可准确探测障碍物的位置和距离,具备良好的避障效果,能够满足实际应用需求,可作为通用方法为无人平台提供一种新思路。

2 实时避障算法

本文基于 YOLOv5s 与 Lite-Mono-8m 深度估计网络设计了一种实时运行的局部防碰撞算法。整体思路是将无人车行进过程中的周围环境映射至视觉成像空间,在像平面描述障碍物相对于无人平台的位置和深度三维属性,再结合人工势场思想构造关于障碍物像素面积与深度的斥力函数,将其推出投影后的无人平台几何核心区实现避障。总体技术路线如图 1 所示,主要包括了 3 个模块:视觉感知模块、避障决策模块和运动控制模块。

视觉感知模块包括了轻量的目标检测网络 YOLOv5s 和深度估计网络 Lite-Mono-8m,能够在移动设备端实时探测障碍物的位置和距离,进一步进行融合计算,获取障碍物目标的等效深度格网。避障决策模块主要包含 3 部分:①先将无人车投影到像平面构成核心区;②构造斥力场计算无人车受到的斥力合力;③综合分析小车受力情况,指导释放制动或转向避障决策。控制模块负责对潜在造成危险

的障碍物进行预警,同时根据小车运动模型执行避障动作。

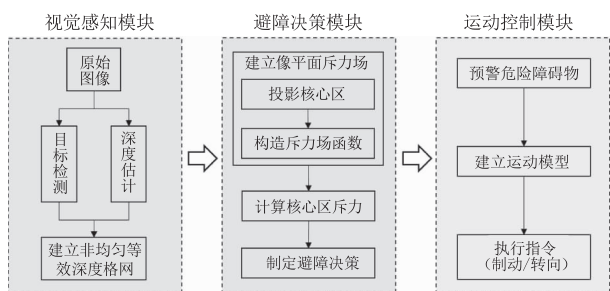


图 1 总体技术路线图

Fig. 1 Technical framework

2.1 视觉感知

2.1.1 YOLOv5s 目标检测算法

YOLOv5 (You Only Look Once) 是性能较为优越的单阶段目标检测算法,其发布版本包括 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x 等一系列算法,支持根据不同任务在精度与速度之间进行平衡,从而选择最适方法。考虑到避障算法的实时性要求,本文选取模型结构小、推理速度快的 YOLOv5s 实现对行驶环境下,行人、车辆等典型避障目标的实时检测。

不同于 RCNN 系列等两阶段检测算法,“先预测位置,再判断类别”,YOLOv5s 先将整张图像划分为等大的网格,并假设目标按照某种分布规律出现在不同的网格中,每个网格预测目标出现的概率、类别和尺寸,再通过标签进行监督训练,回归纹理与分布规律之间的函数,达到一次性预测出全部目标的目的。

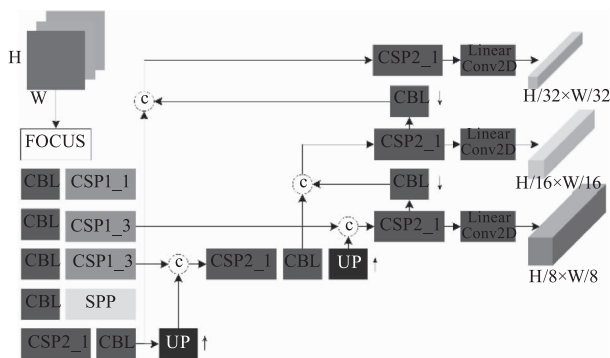


图 2 YOLOv5s 结构图

Fig. 2 YOLOv5s model

如图 2 所示,YOLOv5s 主要包括输入端、特征能够提取网络、特征融合网络和输出端 4 部分:

(1) 在输入端输入型为 $H \times W \times 3$ 的彩色图像,

采用 Mosaic 方式对 4 张图像进行随机缩放、剪裁和堆叠,进而提升网络训练效果和泛化能力。

(2) 特征提取网络如图中结构所示, FOCUS 操作通过利用切片索引实现对原始图像的无参化降采样和最大程度地保留原始信息。借鉴 CSPNet, 设计了两种 CSP 结构作为处理图像特征的基本组件。CSP 旨在通过分割梯度流减少冗余计算量,提高训练参数的优化效果和计算效率,使得模型在保持同等或达到更高精度的条件下,极大程度减少参数量,令模型整体轻量化。原始图像共经过 5 次下采样,得到 $H/4 \times W/4$ 、 $H/8 \times W/8$ 、 $H/16 \times W/16$ 和 $H/32 \times W/32$ 尺度的特征图,并基于后三种尺度特征进行融合,预测出目标。

(3) 特征融合网络用于融合多尺度深度特征,使特征进一步向语义信息转化。为了增加信息的丰富性,该部分被设计为 FPN + PAN 的自底向上和自顶向下结构,首先通过上采样将小尺度特征上采样 2 次分别与对应尺度的骨干网络特征进行拼接构建 FPN 结构,随后将所得大尺度特征下采样 2 次分别与 FPN 对应尺度的特征进行拼接构成 PAN 结构,逐级传递、融合语义和位置信息。

(4) 输出端由线性激励的二维卷积预测目标的位置与类别,将输出张量和标签代入 GIOU 函数计算损失,反求梯度进行参数变量的自更新。最后更具输出张量解出每个网络的预测框,并利用 NMS 算法过滤冗余和低置信度结果,保留最佳预测框,最终实现障碍目标的实时检测。

2.1.2 Lite-Mono-8m 深度估计算法

单目深度估计能够从单张图像中测量出物体到相机的距离,根据训练方式的不同可以分为有监督方法、无监督方法和半监督方法^[25]。Lite-Mono (Lightweight Monocular, LM)^[26] 共有 LM、LM-small、LM-tiny 和 LM-8m 四个版本, Lite-Mono 系列是一种结合卷积神经网络 (Convolutional Neural Networks, CNN) 和视觉 Transformer 架构的无监督深度估计网络,具有轻量、高效的编码-解码结构。

如图 3 所示,编码器共有三个 stage, 每个 stage 又由下采样,数量不等的连续膨胀卷积 (Consecutive Dilated Convolutions, CDC) 模块和 1 个局部-全局特征交互 (Local-Global Features Interaction, LGFI) 模块构成。网络先采用 CDC 在不同的尺度下捕捉图

像局部特征,膨胀卷积能够扩大感受野,有助于提取局部特征。然后利用 LGFI 计算多头注意力并将全局信息进行特征编码,将 XCiT^[27] 注意力计算空间域的方式更改为仅计算通道域,达到减小计算量的目的。解码器由上采样和卷积构成,通过利用二次线性插值拼接不同 stage 的特征,最后用卷积预测头输出 1/4、1/2 和全分辨率的逆深度图。

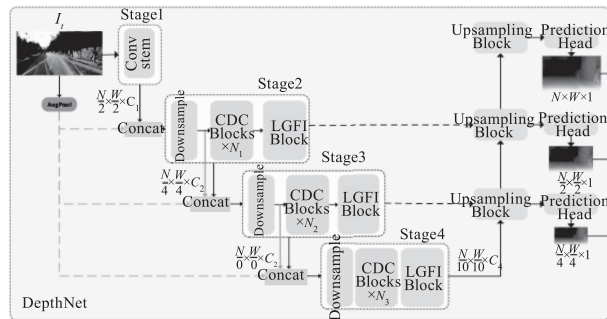


图3 Lite-Mono 结构图

Fig. 3 Lite-Mono model

考虑到有监督方法可为模型提供准确的监督信号而更易获取精度较高的深度,因此本文使用雷达提供的深度真值训练模型,而不用视频流或双目图像对进行无监督训练。所以本文添加了多项式损失函数,如公式:

$$LOSS = \lambda_1 L_1 + \lambda_2 L_2 + \lambda_3 L_{SILog} \quad (1)$$

其中,损失函数是 L1 损失、L2 损失和 $SILog^{[28]}$ 损失的加权求和。L1 损失有利于获取边缘清晰的深度推理结果,L2 损失用于惩罚较大的误差估计结果,而加入 $SILog$ 损失函数则有利于保证场景内物体深度关系的不变性。

由于本文所使用的图像分辨率较大,因此在训练 Lite-Mono 网络时,与原始网络模型不同。本文借鉴 YOLOv5,先用 Focus 降低一半图像分辨率,然后在解码阶段,设置上采样的倍率为 4,使得模型能够恢复出全分辨的逆深度图。目的在于减少计算量的同时避免较大的精度损失。

2.1.3 等效深度格网

视觉感知模块的作用是利用传感器获悉障碍物的空间位置及其运动情况,YOLOv5 与深度网络实现了对障碍目标位置的探测和距离的估计,但两者结果相对独立,需要进一步融合。为了描述无人车的周围环境,本文将视野范围内的三维空间投影到像平面,建立格网表示障碍物的状态。若采用规则格网方法,则确定格网的最适尺寸会过度依赖工程

经验,且易引入不必要的计算量增加算法延时。因此,本文使用检测到的障碍物角点四个坐标 x_{\min} 、 x_{\max} 、 y_{\min} 和 y_{\max} 为边界确定网格的划分方式,区分出障碍区域。单目深度估计模块预测出像素级的密集深度,由格网边界确定出不同的深度集合 $D_{\text{depth}} = \{D_1, D_2, \dots, D_n\}$ 。由于避障过程中关心无人平台与障碍物间的最近距离,所以取每个集合的最小值得到等效深度集 $\bar{D}_{\text{depth}} = \{\min(D_i) \mid D_i \in D_{\text{depth}}, i = 1, 2, \dots, n\}$ 。

2.2 改进人工势场的避障决策

传统的人工势场法是在行进方向所确定的平面内,假设障碍物产生斥力场,目标位置产生引力场,然后在确定范围构造出关于距离的斥力、引力场函数,最后求解负梯度计算出合力,指导无人车产生加速度完成避障。本文将这一思想引入到像平面空间,为等效深度格网建立关于深度与面积的虚拟势场,产生水平方向的斥力作用于小车,达到将障碍目标所在的格网推出核心区的目的。

2.2.1 核心区投影

等效深度格网表示了障碍物的空间信息,但仍缺少无人车自身的位置。本文通过将无人车的实际位置投影到像平面构成核心区,从而为后续分析障碍物和无人车的空间关系提供参考依据。

如图 4 所示,首先假设安全距离为 d_{safe} ,即无人车采取规避动作的最小距离。现只需将无人车在该距离碰撞时的真实尺寸进行投影,可分为图中(a)、(b)、(c)三种情况:

(a) 行至 d_{safe} 处碰撞时,相机视锥下界、地面、小车前侧所在的重力线相交于同一点;

(b) 行至 d_{safe} 处碰撞时,视锥下界与地面的交点远于 d_{safe} 处;

(c) 行至 d_{safe} 处碰撞时,视锥下界与地面的交点近于 d_{safe} 处。

对于情况(b),核心区高度所对应的 d_{safe} 处实际高度为 $H'_{\text{car}} < H_{\text{car}}$,不能确保核心区外的目标不与小车发生碰撞,方法失效。因此,在整机系统的设计阶段应调整 d_{safe} 与相机纵向视场角 VFOV 避免此情况的发生。现考虑更为一般的情况(c),假设相机的等效焦距是 f ,为了确定核心区在图像中的高度和宽度,根据相机小孔成像模型的相似三角关系可得到:

$$h_{\text{pixel}} = \frac{fH_{\text{car}}}{d_{\text{safe}}} S_V^{-1} \quad (2)$$

$$w_{\text{pixel}} = \frac{fW_{\text{car}}}{d_{\text{safe}}} S_H^{-1} \quad (3)$$

式(2)、(3)中, h_{pixel} 、 w_{pixel} 是核心区的高和宽,单位为像素; H_{car} 、 W_{car} 是小车实际高度; d_{safe} 为安全距离; f 是相机焦距; S_V 、 S_H 分别代表垂直和水平两个方向上每个像素在 f 处(像平面内)的实际长度。不妨考虑视场角的影响因素,则有:

$$S_V = \frac{2f \tan \frac{\alpha_{\text{VFOV}}}{2}}{H_{\text{img}}} \quad (4)$$

$$S_H = \frac{2f \tan \frac{\beta_{\text{HFOV}}}{2}}{W_{\text{img}}} \quad (5)$$

式(4)、(5)中, α_{VFOV} 、 β_{HFOV} 分别表示垂直视场角和水平视场角; H_{img} 、 W_{img} 是图像高度和宽度,单位为像素。代入公式(2)、(3)得到:

$$h_{\text{pixel}} = \frac{H_{\text{car}} H_{\text{img}}}{2d_{\text{safe}} \tan \frac{\alpha_{\text{VFOV}}}{2}} \quad (6)$$

$$w_{\text{pixel}} = \frac{W_{\text{car}} W_{\text{img}}}{2d_{\text{safe}} \tan \frac{\beta_{\text{HFOV}}}{2}} \quad (7)$$

核心区底边与图像底边可能具有一定距离。当出现情况(a)时,该高度应为0;若出现情况(c)时,假设高度为 l_{pixel} ,单位为像素。此处需要先确定在垂直和水平两个方向上每个像素在 d_{safe} 处的实际长度,只需将公式(4)、(5)中的 f 替换为 d_{safe} 即可。 l_{pixel} 可写为:

$$l_{\text{pixel}} = \begin{cases} \frac{H_{\text{img}}}{2} \left(1 - \frac{H_{\text{car}}}{d_{\text{safe}} \tan \frac{\beta_{\text{HFOV}}}{2}} \right), & \text{情况(c)} \\ 0, & \text{情况(a)} \end{cases} \quad (8)$$

核心区可以由公式(6)、(7)、(8)确定。

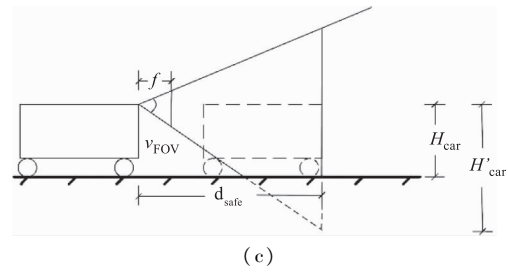
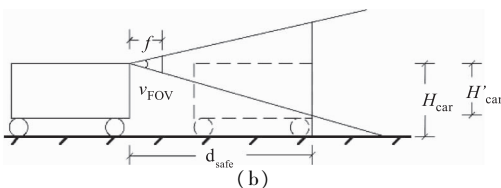
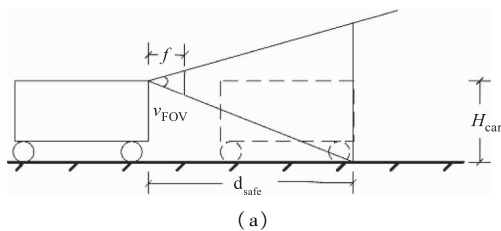


图4 避障安全距离示意图

Fig. 4 Safe distance for avoidance

2.2.2 像平面斥力场建立

传统人工势场法如图5所示,障碍物、小车与目标点处于行进方向的平面内。目标点产生对小车的引力,方向始终指向目标点,大小与距离成正比。障碍物产生方向指向小车,大小与距离成反比的斥力。

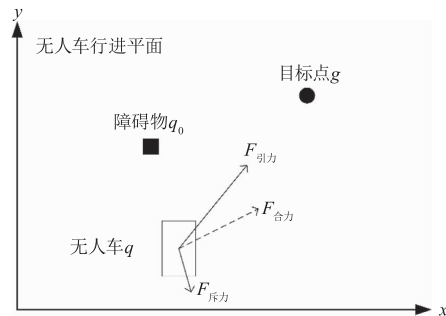


图5 人工势场示意图

Fig. 5 The principle of APF

目标点产生对小车的引力势场:

$$U_{\text{att}}(q) = \frac{1}{2} \eta \rho^2(q, q_g) \quad (9)$$

式中, η 为引力增益系数; $\rho(q, q_g)$ 大小为小车与目标点的距离,方向指向目标点,对势场求负梯度得到引力函数:

$$F_{\text{att}}(q) = -\nabla U_{\text{att}}(q) = -\eta \rho(q, q_g) \quad (10)$$

图中障碍物点产生斥力场:

$$U_{\text{rep}}(q) = \begin{cases} \frac{1}{2} k \left(\frac{1}{\rho_{\text{c}(q, q_0)}} - \frac{1}{\rho_0} \right)^2, & 0 \leq \rho_{(q, q_0)} \leq \rho_0 \\ 0, & \rho_{(q, q_0)} \geq \rho_0 \end{cases} \quad (11)$$

式中, k 为斥力系数; $\rho_{(q, q_0)}$ 的大小为小车与障碍物的距离,方向指向小车; ρ_0 为常数表示障碍物的最大作用范围,则相应斥力函数为:

$$F_{\text{rep}}(q) = \begin{cases} k \left(\frac{1}{\rho_{(q, q_0)}} - \frac{1}{\rho_0} \right) \frac{1}{\rho_{(q, q_0)}^2}, & 0 \leq \rho_{(q, q_0)} \leq \rho_0 \\ 0, & \rho_{(q, q_0)} \geq \rho_0 \end{cases} \quad (12)$$

受人工势场启发,本文将这一思想引入像空间,在成像平面内根据障碍物深度和像素面积建立虚拟斥力势场,使进入核心区域的目标在一定距离范围内对小车产生水平斥力,在合力作用下规划车辆运动方向。为简化条件,不考虑目标点的影响,将公式(11)中的距离要素替换为深度在像平面的定积分 $V(x,y) = \iint d(x,y) dx dy$,方向水平向左或者向右。 $d(x,y)$ 表示深度在像平面的密度,采取等效深度进行替代,可写为 $V(x,y) = \bar{d}s$,则有:

$$F_{\text{rep}}(V) = \begin{cases} k \left(\frac{1}{\bar{d}s} - \frac{1}{d_{\text{safe}}s} \right) \frac{1}{(\bar{d}s)^2}; 0 \leq \bar{d} \leq d_{\text{safe}} \text{ 且 } s > 0 \\ 0; \bar{d} \geq d_{\text{safe}} \text{ 或 } s \leq 0 \end{cases} \quad (13)$$

公式(13)中, \bar{d} 为格网区域的最小深度; s 为格网与核心区的交并比,由此能够确定斥力的大小,若在核心区中心左侧,则斥力方向水平向右,反之亦然。当格网等效深度进入 d_{safe} 范围内且 $s \geq 0$,此时无人车将采取避障措施。

2.2.3 避障策略分析

无人车根据核心区的受力情况采取不同的避障措施,主要包括3种情形:

(a)第一种是水平斥力的合力大小不为零且方向向左,此时应控制无人车向左转向实施避障;

(b)第二种是水平斥力的合力大小不为零且方向向右,此时应控制无人车向右转向实施避障;

(c)第三种是水平斥力的合力大小为零,此时应考虑核心区格网深度小于 d_{safe} 的面积 s ,若 s 不为零,无人车应采取制动,若 s 为零无人车应保持当前运动状态,无需避障。

2.3 运动控制

无人车底盘具有稳定的控制系统,可根据决策指令实现偏航角及速率的控制,运动模型如图6所示。当采取避障动作时,需要由避障算法实时计算出小车产生偏航角 θ ,以及小车的速率。

无人车转向与合力方向一致,在确定转向后,偏航角 θ 可由核心区内障碍物与核心区边缘的最大距离 ΔL 推算得到:

$$\theta = \arctan \frac{2\Delta L \tan \frac{\beta_{\text{HFOV}}}{2}}{W_{\text{img}}} \quad (14)$$

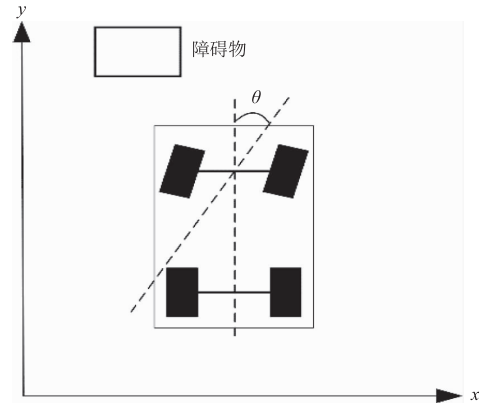


图6 小车转向示意图

Fig. 6 Unmanned vehicle steering

线速度由制动力和水平斥力共同决定也可以根据无人车速度设定固定值,直到障碍物目标消失在视野,然后解除避障回归全局规划航线。避障算法的主要逻辑如表1所示。

表1 避障算法主逻辑

Tab. 1 The primary implementation of algorithm

- | |
|---|
| S1. 输入相机单目图像; |
| S2. 由 YOLOv5s 检测出所有目标框的 4 个角点坐标; |
| S3. 由 LiteMono-8m 网络预测出像素级的密集深度图; |
| S4. 建立核心区; |
| S5. 以目标框角点坐标为索引在深度图上计算框内的最小深度 \bar{d} 作为等效深度; |
| S6. 计算所有目标框与核心区的交并比; |
| S7. 筛选符合交并比大于 0 且等效深度小于安全距离的目标框; |
| S8. 根据斥力函数计算为符合条件的目标框计算水平斥力 F_{rep} , 方向向右为正、向左为负; |
| S9. 计算所需的实时航偏角 θ ; |
| S10. 计算所需的实时线速度。 |

3 实验结果与分析

3.1 平台和数据集

本文的实验平台为自研无人车平台,搭载自主研发的可见光相机和红外相机各 1 个、图达通激光雷达 1 个、北斗星通 GNSS/IMU 系统 1 套和独立的北微传感 IMU 系统 1 套。无人车的部分参数如表 2 所示,可见光和红外是中心视场对齐的,图像分辨率为 2 倍关系,分别为 1280×1024 和 640×512 ,激光雷达的有效扫描距离为 120 m,深度真值的距离范围更大。在实验前已经对该平台进行了严格的硬件时间同步和各传感器的内外参标定,以确保不同传感器数据具有良好的时空一致性。

软件环境方面,为了便于数据的采集和管理,本文在 Ubuntu 1604 系统中搭建了 ROS 环境,经过解析 ROS 数据包,可以获取具有相同时间戳的图像和

点云数据,再利用标定文件将同步后的点云分别转换到可见光和红外的相机坐标系,得到一组对应的可见光图像、红外图像和真值深度图。选择 Pytorch 和 Tensorflow2 深度学习框架搭建 YOLOv5 和 Lite-Mono 的网络,在 24GB 显存的 RTX 3090GPU 上实现模型训练和推理。

实验场地为某工业园区,园区占地面积约为 30 万平方米,在不同的路段采集数据构建数据集,根据路段的不同划分出训练集和测试集。该数据集可作为深度估计模型的训练数据,包括可见光、红外和雷达深度图,其中训练集有 78406 张,测试集有 11720 张。对于目标检测算法,在 COCO、VOC 等大型公开数据的预训练模型能够满足使用需求,直接利用预训练的目标检测模型避免了从头构建数据集时大量的人工标注工作。

表 2 无人平台部分参数

Tab. 2 Partial specification of platform

无人车平台	
长×宽×高/mm ³	980×745×1200
最高速度/(m·s ⁻¹)	1.5
可见光相机	
焦距/mm	8.54
视场角/(°) ²	44×35.17
图像分辨率	1280×1024
红外相机	
焦距/mm	14.30
图像分辨率	640×512
激光雷达	
有效距离/m	120
角分辨率/(°) ²	0.18×0.24
扫描频率/(f·s ⁻¹)	10

3.2 视觉感知结果

3.2.1 YOLOv5s 目标检测

表 3 是 YOLOv5 的 5 种模型在 COCO2017 验证集的测试情况。考虑到 YOLOv5s 模型在参数量方面较小,仅为 3.2 m,运行效率较高,对于检测常见障碍物目标,其检测精度能够满足实际需求,故选取 YOLOv5s 作为视觉模块的目标检测算法。

设置 YOLOv5s 的置信度阈值为 0.25,交并比阈值为 0.45,图 7 是该算法在本文构建的数据集上运行的结果。可以看出,模型能够从可见光图像中较为准确地检测出行人、车辆等常见的障碍

物。图 8 是模型特征的可视化结果,模型先从图像中提取物体轮廓、颜色等底层信息,随着神经网络层数加深,模型将特征进一步提取更为抽象地高层次特征,最后在模型顶端输出语义信息。本文也在红外图像上进行了实验,由于红外是基于温差进行成像,红外图像也具有物体的形状、轮廓等信息,在特征上与可见光图像具有一定相似性。所以对于容易给无人车行进造成安全威胁的近处障碍物,YOLOv5s 也可以进行准确探测,结果如图 9 所示。

表 3 YOLOv5 系列模型的具体性能

Tab. 3 Performance of YOLOv5 models

模型	尺寸	mAP	参数量/M
YOLOv5n	1280	54.4	3.2
YOLOv5s	1280	63.7	12.6
YOLOv5m	1280	69.3	35.7
YOLOv5l	1280	71.3	76.8
YOLOv5x	1280	72.7	140.7



图 7 可见光图像障碍物检测结果

Fig. 7 The detected results of obstacle in RGB images



图 8 模型特征可视化

Fig. 8 Feature visualization from model

由此可见,采用在 COCO2017 公开数据集上进行预训练的 YOLOv5s 能够满足本文需求,且避免了从头构建数据集、迭代调参训练等大量工作,可为后续避障模块提供障碍物的位置信息。



图9 红外图像障碍物检测结果

Fig. 9 The detected results of obstacle in infrared images

3.2.2 Lite-Mono-8m 单目深度估计

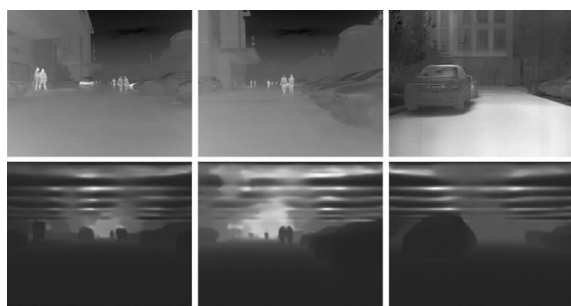
在自定义数据集上展开单目深度估计实验,将原始无监督的 Lite-Mono 更改为有监督的学习方式,使用激光雷达点为模型提供准确的监督信号,有利于获取较好的深度估值。训练时随机对数据进行增强,以扩充训练样本,缓解模型的过拟合。确保每轮训练都有 40% 的原数据,其余数据的亮度、对比度和饱和度在一定限度内随机变化,并进行随机上下、左右翻转。采用初始学习率为 0.0001,其余参数默

认的 Adam 优化器训练模型。分别使用可见光和红外图像训练模型,并在测试集上计算 80 m 内和 10 m 内的绝对相对误差来衡量各模型的性能优劣。表 4 是 Lite-Mono 各版本模型的测试结果。其中 Lite-Mono-8m 的参数量最大,但同时具备最优的精度,对于可见光图像,在测试集上 80 m 之内的平均绝对相对误差为 11.93%。常见障碍物的距离预测结果见图 10(a)。实验也测试了红外数据,由于缺少丰富的颜色和纹理信息,红外深度估计的误差相对较大,在 80 m 内的误差为 14.50%,推理效果见图 10(b)。为了验证模型用于避障的可行性,本文还测试了 10 m 以内的误差,结果表明 Lite-Mono-8m 的近距离误差为 9.29%,在红外状态下为 10.7%,均能满足无人平台避障的实际需求。

表 4 Lite-Mono 系列模型及测试情况

Tab. 4 Performance of Lite-Mono on test dataset

模型	数据源	图像分辨率	参数量/M	AbsRel. (< 10 m)	AbsRel. (< 80 m)
LM	可见光	1280 × 1024	3.08	10.35 %	13.89 %
LM-small	可见光	1280 × 1024	2.48	10.09 %	13.70 %
LM-tiny	可见光	1280 × 1024	2.15	10.85 %	15.21 %
LM-8m	可见光	1280 × 1024	8.78	9.29 %	11.93 %
LM-8m	红外	640 × 512	8.78	10.7 %	14.50 %



(a)可见光图像深度图



(b)红外图像深度图

图 10 Lite-Mono-8m 推理结果

Fig. 10 The depth results of Lite-Mono-8m

3.3 避障结果

根据视觉感知模块的计算结果可以实时分析路况条件,进而根据不同情况制定出合理的避障策略。首先需要将相对独立的检测结果和预测的深度进行融合,目标检测模型提供了障碍物的四个角点坐标,取目标同格网中的最近深度等效替代整个格网像素的深度。然后参考公式(6)、(7)、(8)将无人平台几何体投影到像平面构成核心区。表 5 对应了不同安全距离下像平面核心区的尺寸。随着安全距离的增大,核心区尺寸在不断减小,安全距离为 2m 时, l_{pixel} 的计算值为负数,说明此时符合图 4 中的情况(b)。设置不同的安全距离,图 11 是相应的无人车核心区投影结果,中心的红色矩形框内部代表投影过后的核心区。图 12 是本文方法的避障效果图,展示了可见光图像和红外图像的直观避障结果,同时合力方向和偏航角由公式(13)、(14)计算得出,无人车线速度设定为最大速率 1.5 m/s 匀速行进。图 12 中的(a)、(b)和(i)~(l)图像是运动的无人车

相对于静止的人和车辆由算法给出的转向方向和角度,通过对比较为连贯的两幅画面可以看出,此时无人车能够正确地进行自主转向,成功躲避具有潜在威胁的障碍物。由此可见,本文提出的避障方法能实时地计算出正确的偏转方向和角度,在控制系统接收到势场力方向、偏航角和线速度等信号后控制无人车转向或制动,进而完成实时避障。

表 5 不同安全距离对应的核心区尺寸

Tab.5 Core area dimensions corresponding to different safety distances

d_{safe}/m	w_{pixel}/pix	h_{pixel}/pix	l_{pixel}/pix	情况
2	594	969	-248	(b)
3	396	646	5	(c)
5	238	388	208	(c)
7	170	277	295	(c)
10	119	194	360	(c)

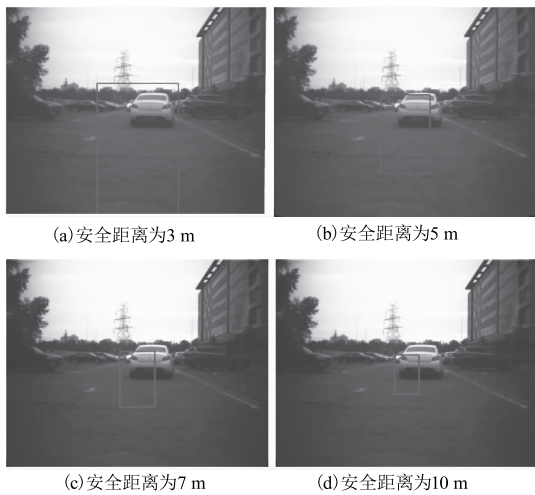


图 11 核心区投影示意图

Fig. 11 Schematic diagram of the core area projection

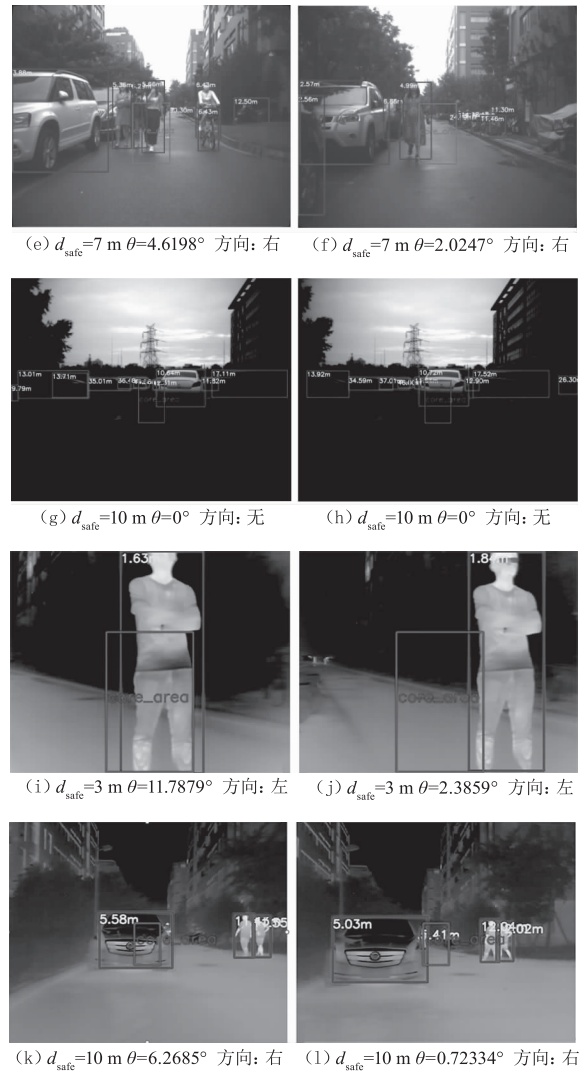
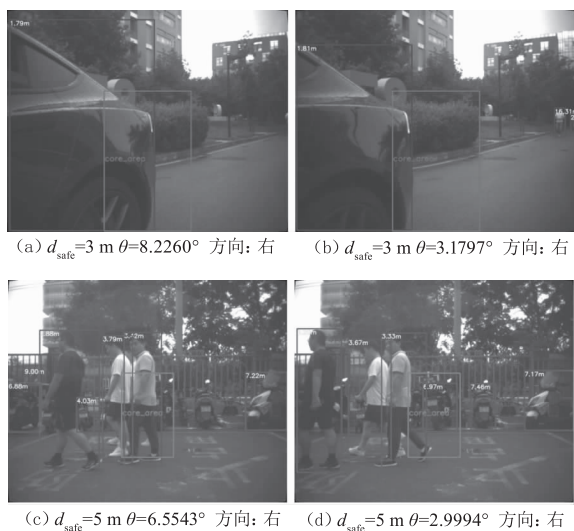


图 12 避障效果图

Fig. 12 Obstacle avoidance rendering

4 结 语

避障是无人平台实现自主导航的前提,确保了行进过程中的安全性。以往的避障方法通常是通过将自身放入三维场景进行问题建模,需要借助昂贵的主动传感器,算法条件苛刻,影响鲁棒性。本文将三维空间的障碍物和平台统一到二维成像平面,给出了平台投影到像平面核心区的具体过程,用等效深度格网描述障碍物的位置和距离,并通过判断与核心区的关系计算势场力和偏航角引导无人车完成实时避障,这一模式与人眼视觉避障更加贴合。实验表明,本文视觉感知模块能够准确探测障碍物的位置和深度信息,提出的避障方法可以在常见场景中成功避障。在未来工作中,会进一步采用视觉密集型网络替换现有的感知模块,实现一网多用、提高运行效率的同时优化感知精度,进而增强系统的稳定性。

参考文献:

- [1] Meng Hong, Zhu Sen. The development and future trends of unmanned ground systems [J]. *Acta Armamentarii*, 2014, 35(S1): 1-7. (in Chinese)
孟红, 朱森. 地面无人系统的发展及未来趋势[J]. *兵工学报*, 2014, 35(S1): 1-7.
- [2] Wu Tong, Fu Zhongli. 3D reconstruction technology and military application [J]. *National Defense Technology*, 2015, 36(1): 31-34. (in Chinese)
吴彤, 傅中力. 三维重建技术及其军事应用[J]. *国防科技*, 2015, 36(1): 31-34.
- [3] Xu Bei, He Songbai, Zheng Zhi. Military unmanned vehicle safety technical analysis [J]. *Scientific and Technological Innovation*, 2018, (29): 28-29. (in Chinese)
徐蓓, 何松柏, 郑智. 无人驾驶军用车辆安全技术分析[J]. *科学技术创新*, 2018, (29): 28-29.
- [4] Mujumdar A, Padhi R. Evolving philosophies on autonomous obstacle/collision avoidance of unmanned aerial vehicles [J]. *Journal of Aerospace Computing Information & Communication*, 2011, 8(2): 17-41.
- [5] S Yan, F Pan. Research on route planning of AUV based on genetic algorithms [C]//2019 IEEE International Conference on Unmanned Systems and Artificial Intelligence (ICUSAI), Xi'an, China, 2019: 184-187.
- [6] Lim H S, Fan S, Chin C K H, et al. Constrained path planning of autonomous underwater vehicle using selectively-hybridized particle swarm optimization algorithms [C]//IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles. 2020.
- [7] Szczerba R J, Galkowski P. Robust algorithm for real-time route planning [J]. *IEEE Transactions on Aerospace & Electronic Systems*, 2000, 36(3): 869-878.
- [8] Ganapathy V, Yun S C, Chien T W. Enhanced D* lite algorithm for autonomous mobile robot [J]. *International Journal of Applied Science & Technology*, 2011, 1(1): 58-74.
- [9] Durand N. Constant speed optimal reciprocal collision avoidance [J]. *Transportation Research Part C: Emerging Technologies*, 2018, 96(NOV.): 366-379.
- [10] X Yang, W Yang, H Zhang, et al. A new method for robot path planning based artificial potential field [C]//2016 IEEE 11th Conference on Industrial Electronics and Applications (ICIEA), Hefei, China, 2016: 1294-1299.
- [11] Sun Xiaofeng, Gong Jinliang, Zhang Yanfei. Improved obstacle avoidance of optical flow method and aided navigation algorithm [J]. *Laser Journal*, 2019, 40(10): 16-20. (in Chinese)
孙晓峰, 宫金良, 张彦斐. 改进光流法的避障与辅助导航算法 [J]. *激光杂志*, 2019, 40(10): 16-20.
- [12] Fu Qiang, Wang Jin, Gong Yue, et al. Obstacle avoidance of flapping-wing air vehicles based on optical flow and fuzzy control [J]. *Transactions of Nanjing University of Aeronautics and Astronautics*, 2021, 38(2): 206-215.
- [13] Yu Chaofan, Sun Jianhui. A real-time obstacle avoidance system for multi-rotor unmanned aerial vehicle based on optical flow sensor [J]. *Computer Applications and Software*, 2018, 35(1): 206-210. (in Chinese)
余超凡, 孙建辉. 基于光流传感器的旋翼无人机实时避障系统 [J]. *计算机应用与软件*, 2018, 35(1): 206-210.
- [14] Zhang Xiangzhu, Zhang Lijia, Song Yifan, et al. Obstacle avoidance algorithm for unmanned aerial vehicle vision based on deep learning [J]. *Journal of South China University of Technology: Natural Science Edition*, 2022, 50(1): 101-108, 131. (in Chinese)
张香竹, 张立家, 宋逸凡, 等. 基于深度学习的无人机单目视觉避障算法 [J]. *华南理工大学学报: 自然科学版*, 2022, 50(1): 101-108, 131.
- [15] Wei Jiansheng, Pan Shuguo, Tian Guangzhao, et al. Design and experiments of the binocular visual obstacle perception system for agricultural vehicles [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021, 37(9): 55-63. (in Chinese)
魏建胜, 潘树国, 田光兆, 等. 农业车辆双目视觉障碍物感知系统设计与试验 [J]. *农业工程学报*, 2021, 37(9): 55-63.
- [16] Zhao Xiaoyong, Chen Qin Zhu, Zheng Hongyan. Research and application of obstacle avoidance method for inspection robot based on road scene understanding [J]. *Microelectronics & Computer*, 2022, 39(4): 118-127. (in Chinese)
赵小勇, 陈钦柱, 郑鸿彦, 等. 基于道路场景理解的巡检机器人避障方法研究与应用 [J]. *微电子学与计算机*, 2022, 39(4): 118-127.
- [17] Peng Fan, Xie Yongfang, Chen Xiaofang, et al. Robot real-time obstacle avoidance algorithm based on prediction of obstacle reachable area [J]. *Journal of Northeastern University: Natural Science*, 2022, 43(9): 1225-1233. (in Chinese)
彭帆, 谢永芳, 陈晓方, 等. 基于障碍物可达区域预测

- 的机器人实时避障算法[J]. 东北大学学报:自然科学版, 2022, 43(9): 1225 - 1233.
- [18] B Li, Y Wu. Path planning for UAV ground target tracking via deep reinforcement learning[J]. IEEE Access, 2020, (8): 29064 - 29074
- [19] Yan C, Xiang X, Wang C. Towards real-time path planning through deep reinforcement learning for a UAV in dynamic environments[J]. J. Intell. Robot Syst. 2020, 98: 297 - 309.
- [20] Nieuwenhuisen M, Droschel D, et al. Autonomous navigation for micro aerial vehicles in complex gnss-denied environments[J]. J. Intell. Robot. Syst. , 2016, 84(1 - 4): 199 - 216.
- [21] Iacono, M. , Sgorbissa, A. . Path following and obstacle avoidance for an autonomous UAV using depth camera [J]. Robot. Auton. Syst. , 2018, 106: 38 - 46.
- [22] Cui, J. Q. , Lai, S. , Dong, X. , et al. Autonomous navigation of UAV in foliage environment [J]. J. Intell. Robot. Syst. , 2016, 84(1 - 4): 259 - 276.
- [23] Liu Yanju, Zou Bo, Song Jianhui, et al. Design of an obstacle avoidance system of unmanned vehicles with single-line laser radar [J]. Journal of Shenyang Ligong University, 2017, 36(2): 12 - 16. (in Chinese)
刘砚菊, 邹波, 宋建辉, 等. 一种基于单线激光雷达的无人车避障系统设计 [J]. 沈阳理工大学学报, 2017, 36(2): 12 - 16.
- [24] Tan Baocheng, Yang Cheng. Dynamic obstacles detection by laser radars [J]. Journal of Xi'an Technological University, 2015, 35(3): 205 - 209. (in Chinese)
谭宝成, 杨成. 激光雷达动态障碍物检测 [J]. 西安工业大学学报, 2015, 35(3): 205 - 209.
- [25] Jiang Junjun, LI Zhenyu, LIU Xianming. Deep learning based monocular depth estimation: a survey [J]. Chinese Journal of Computers, 2022, 45 (6): 1276 - 1307. (in Chinese)
江俊君, 李震宇, 刘贤明. 基于深度学习的单目深度估计方法综述 [J]. 计算机学报, 2022, 45 (6): 1276 - 1307.
- [26] Zhang N, Nex F, Vosselman G, et al. Lite-Mono: a lightweight CNN and transformer architecture for self-supervised monocular depth estimation [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, 2023: 18537 - 18546.
- [27] Ali A, Touvron H, Caron M, et al. Xcit: Cross-covariance image transformers [J]. Advances in Neural Information Processing Systems, 2021, 34: 20014 - 20027.
- [28] Eigen D, Puhrsch C, Fergus R. Depth map prediction from a single image using a multi-scale deep network [J]. Advances in Neural Information Processing Systems, 2014, 27: 1 - 9.