

边缘智能驱动的高能效无人机自主导航算法研究

林椿珉, 曾烈康, 陈旭

(中山大学计算机学院, 广东 广州 510006)

摘要: 近年来,无人机的自主导航技术在多个行业中受到了广泛的关注,相比于传统的导航技术,采用图像感知的深度学习方法具有很好的泛化能力并且不受全球定位系统(GPS, global positioning system)信号的影响,被证明是一种具有前景的自主导航方法。然而,深度学习的推断需要较大功耗,这对于能耗资源十分有限的无人机来说是一项挑战。针对该问题,基于边缘智能理论,将强化学习技术引入无人机端侧的推断过程中,根据无人机所处的环境复杂度实时感知信息,动态配置卷积神经网络的结构参数,使得无人机在保持稳定导航的同时,尽可能地减少计算功耗开销,实现无人机高可靠、低时延与高能效的自主导航飞行能力。该算法在仿真环境和现实环境中分别进行了验证,实验结果表明,相比于对比算法,所提的基于强化学习动态配置算法能够让无人机花费更少的计算能耗开销具有更长的飞行距离与更高的成功率。

关键词: 无人机; 边缘智能; 深度学习; 强化学习; 自主导航

中图分类号: TP399

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2021.00229

Research on power efficient autonomous UAV navigation algorithm: an edge intelligence driven approach

LIN Chunmin, ZENG Liekang, CHEN Xu

School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China

Abstract: Autonomous drone navigation has received growing attention in the recent community. Compared with traditional navigation approaches which rely on location-based services highly, deep learning based visual methods have showed superior performance in self-adaption and generalization, which are a promising solution for autonomous navigation. Running the resource-hungry deep learning execution in the resource-constrained unmanned aerial vehicle (UAV), however, significant challenges were presented in power efficiency. To tackle this challenge, following the idea of edge intelligence, a deep reinforcement learning approach was introduced to dynamically configure the computational scale of the deep learning model on UAV and hence realize the autonomous navigation with low latency and high energy efficiency. Evaluations based on both simulation and real prototype experiments show that the proposed approach has the less energy consumption, longer navigation trail and higher obstacle avoidance rate.

Key words: UAV, edge intelligence, deep learning, reinforcement learning, autonomous navigation

1 引言

近年来,无人机的自主导航能力受到了机器人领域的广泛关注,自主导航无人机的易部署、敏捷

性强和机动性高等优点,让它在许多行业中得到了广泛应用,例如,消防巡检、精准农业、快递配送和公共安全巡视等^[1-4]。这些应用的执行通常依赖于图像识别等计算密集型任务,传统部署方式是利用

收稿日期: 2020-12-29; 修回日期: 2021-03-20

通信作者: 陈旭, chenxu35@mail.sysu.edu.cn

基金项目: 国家自然科学基金资助项目(No.U20A20159, No.61972432)

Foundation Items: The National Natural Science Foundation of China (No.U20A20159, No.61972432)

云计算基础设施,将终端侧的计算任务卸载至远程云数据中心执行,而近年来兴起的边缘计算概念则提供了一种全新的解决方案。相较于集中式算力的云计算范式,边缘计算将计算资源移动到靠近数据源的网络边缘侧,为终端应用的执行提供网络时延更低的计算环境,进一步提高服务的实时性、稳定性和可靠性。通过将边缘计算与人工智能有机结合,边缘智能能够进一步赋能端侧应用,让计算变得更加智能高效。本文基于边缘智能的观点,提出了一种全新的无人机自主导航框架,在赋予无人机边缘算力的同时,利用深度强化学习等人工智能算法,综合优化无人机运行功耗和计算时延等目标,实现了自主导航任务在无人机上的自适应执行。

为了实现在低空环境中的自主导航,无人机端侧计算需要解决感知、控制与定位这 3 个基本的问题。解决这些问题的传统方法,首先使用 GPS 与惯性传感器进行定位与姿态估计,然后根据到达的目标点计算控制指令。这种方法依赖于 GPS 信号强度,在有高楼遮挡或者室内等环境时,会出现自动导航精度低或者完全失效的情况。单目相机等摄像头传感器由于其质量轻、体积小、不依赖于 GPS 信号等优点,适用于载重有限的无人机平台。因此基于视觉传感器的方法被越来越多地运用在无人机自主导航中。其中单目视觉同步定位与建图(SLAM, simultaneous localization and mapping)的方法虽然能够在未知环境中实现定位与导航^[5-6],但是在动态场景中提取特征点的难度大,并且无人机的运动模糊图像会有鲁棒性的问题,导航可靠性较差。

采用视觉感知技术的另一种方法是使用深度学习模型进行导航^[7-9]。基于监督学习的卷积神经网络模型能够从大量图像中学习特征信息,模拟出人的操作过程,具有识别精度高、鲁棒性好和时延低等优点,在无人机自主导航领域具有广阔的前景。然而,深度卷积神经网络模型的实时计算需要较大的计算功率开销,若在载重能力与功耗资源有限的无人机上运行该类任务,其续航能力将会受到极大的限制。为了让无人机能耗利用率最大化,文献[10]设计了一种面向服务架构的云计算平台,通过将无人机的计算任务卸载到云数据中心的 Kubernetes 集群,缓解无人机的计算和存储压力。然而,由无人机移动性所导致的带宽波动会对该方法的性能产生巨大影响,难以保证稳定的性能表现。文献[11]设计了专用于微型无人机自主导航卷积神经网络

计算的系统级芯片(SoC, system on chip),提供了较高性能的低功耗自主飞行。然而,基于专用硬件的系统设计可扩展性较差,且在实际部署中需要集成额外的传感器元件,极大地提高了开发难度。面向自主导航任务的无人机计算场景如图 1 所示。

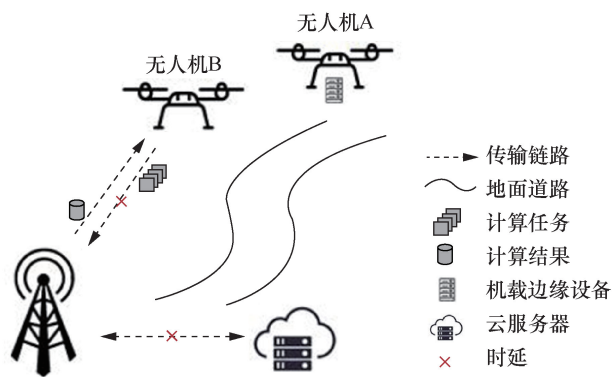


图 1 面向自主导航任务的无人机计算场景

嵌入式高性能芯片的发展让边缘端设备的算力获得了极大的提升,使机载设备拥有了实时处理计算密集型任务的能力,越来越多的人工智能应用迁移到边缘进行。边缘计算与人工智能的结合催生了边缘智能技术^[12-14],它能够使边缘设备拥有执行智能算法的能力。目前,针对边缘智能所依赖的 4 个特性为:低功耗性、低时延性、高准确率性和隐私安全性,学术界在边缘智能的研究方向主要包括模型优化、任务资源分配、边缘联邦智能、云边端协同^[15-19]。其中,模型优化从边缘赋能智能的角度出发,主要解决了计算准确性与实时性、能量消耗之间的矛盾。本文将智能算法下沉到边缘设备端(如图 1 中无人机 A)进行部署,结合无人机数据的本地性与智能算法的决策能力,解决资源受限问题和实时处理问题。

为了实现高可靠、低时延和高能效的无人机自主导航飞行,本文将无人机导航计算任务与边缘智能计算任务卸载到无人机搭载的边缘计算设备中,避免了无人机动态飞行过程中网络带宽波动所带来的时延影响,从而满足实时导航任务对低时延的要求^[15,20]。针对无人机能够携带计算能耗资源的有限性,本文提出了一种基于深度强化学习方法的环境感知导航和模型动态配置策略。该策略能够结合无人机所处的环境状态信息,通过强化学习动态获取最优卷积神经网络配置,实现在无人机平台上基于卷积神经网络模型的低时延与高能效的自主导航。本文的主要贡献可概括如下。

1) 设计了一种轻量型的多输入卷积神经网络无人机自主导航框架, 采用多尺度感知以及多种规模数据共同训练的方式, 让模型在具有较高的精度与泛化能力的同时, 能够具备多尺度图像环境特征信息感知与计算功率动态配置的能力。

2) 提出了一种基于强化学习的决策算法, 可根据无人机所处的环境状态信息选出适用于当前状态的功率最优的卷积神经网络计算配置。在保证无人机完成可靠地自主飞行的同时, 提高了无人机可用能源的利用效率, 实现低时延与高能效的无人机自主导航。

3) 分别搭建了仿真环境下和现实场景中的原型系统, 对本文方法进行了全面的性能验证, 将所提算法与卷积神经网络模型传统静态计算方式、基于阈值选取的算法进行对比。仿真的实验结果表明, 本文提出的算法与两种对比算法相比, 连续飞行的距离分别延长了 5% 和 13%, 飞行成功率分别提高了 3% 和 9%。现实场景的实验表明, 本文的强化学习配置算法与两种对比算法相比, 分别使无人机的自主导航任务续航时间提高了 15% 和 7.5%。

2 无人机自主导航框架

本文基于神经网络搭建无人机自主导航模型, 网络中加入残差块提升网络训练性能, 采用仿真环境中采集的数据集训练模型的环境感知能力。该模型的目标在于根据无人机前置摄像头反馈的图像推断出偏航角度以及碰撞概率, 其计算结果将被转化为指引无人机自主导航飞行的指令, 让无人机能够沿着所处路径飞行, 并且能够及时避免与障碍物发生碰撞。模型的主体框架结合了残差块和空间域金字塔池化层, 使其在保证准确度的同时, 具有多

规模尺度特征信息的提取能力, 让模型具有动态配置能力。本文采用在虚拟现实仿真环境中采集的多种规模数据共同训练方式让导航模型能够适配不同规模大小的输入。部署到现实环境中时, 将仿真环境数据集训练的模型作为预训练模型, 然后在现实环境中收集少量数据集进行迁移学习, 提高模型训练过程中的安全性与可靠性。

2.1 多输入卷积神经网络

采用深度学习进行无人机自主导航时, 需要设计卷积神经网络模型对环境图像特征与语义进行感知与推断。卷积神经网络当前存在多种卷积神经网络框架可供选择, 如使用重复卷积核的网络 VGG (visual geometry group)^[21]、含合并连接的网络 GoogLeNet^[22]、残差网络 ResNet^[23]等, 根据计算时延、卷积网络性能、训练样本集等因素, 本文选取残差网络作为导航系统的主框架, 使得模型能够在较少层数下获得较好的训练效果。

卷积神经网络结构如图 2 所示, 不考虑池化、批归一化等操作, ResNet 模型共由 18 个网络层组成。ResNet 主体部分包含 3 个残差块, 每个残差块中包含两个卷积层与一个跳跃卷积层。3 个残差块中卷积核大小都为 3×3 , 输出通道数分别为 32、64 和 128。残差块内部的数据流可以采用如式(1)所示形式来表示。

$$y = F(x) + x \quad (1)$$

其中, x 为模块的输入, y 为模块的输出, F 为对 x 的残差映射, 即当残差 $F(x) = 0$ 时, 残差块只做恒等映射。

在卷积神经网络的推断中, 对于不同规模的图像输入所需的计算功耗不同, 虽然规模小的输入能

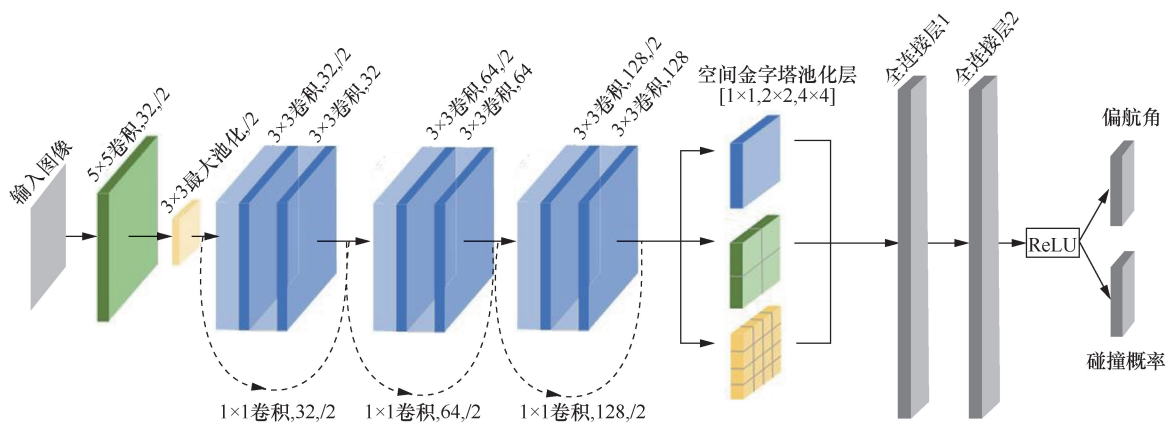


图 2 卷积神经网络结构

够带来低功耗的优势,但是它所具备的环境感知能力较弱。为此,本文采用空间金字塔池化(SPP, spatial pyramid pooling)技术^[24]实现多尺度感知,并且采用多种规模数据共同训练方式让导航模型能够适配不同规模大小的输入,使其能够结合强化学习动态配置策略在环境感知力与计算功耗之间进行折中。空间金字塔池化结构如图3所示。

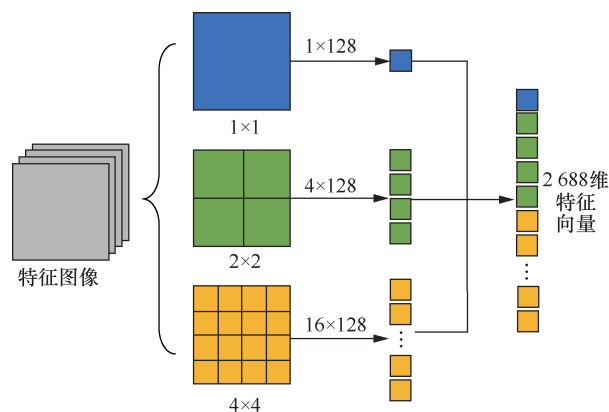


图3 空间金字塔池化结构

空间金字塔池化是一种对卷积输出从不同尺度提取特征并且进行融合的技术,通过将不同维度空间的局部特征映射到固定大小的特征向量,提高模型的识别精度。空间金字塔池化首先将特征图像划分为 4×4 、 2×2 和 1×1 刻度的网格窗口,如图2所示,然后在网格中的窗口采取最大池化操作,提取相应维度的特征,然后将3个特征向量进行线性融合得到固定维度的特征向量,该融合特征包含了不同层级的特征信息,保留了不同尺度的形状、纹理和高级语义等特征,在提高精度及鲁棒性的同时让模型能够适配不同功耗大小的计算。

将空间金字塔池化层融合之后的特征向量与全连接层相连来计算最后的结果。模型全连接结构包含两层全连接层,每一个全连接层的通道数(即空间金字塔池化融合特征数)为2 688。经过线性整流函数(ReLU, rectified linear unit)之后,两个任务停止共享网络参数,分为两个不同的全连接层分支,第一个全连接层输出无人机偏航角,第二个全连接层采用Sigmoid函数将输出映射到(0,1)表示输出碰撞的概率。

2.2 数据集采集与模型训练

监督式深度学习模型的训练需要大量的数据集,然而,在现实环境中收集训练数据耗时且缺乏安全性。因此,本文采用微软公司基于虚幻引擎开

发的模拟仿真器AirSim^[25]进行训练数据的收集与监督学习训练。AirSim能够提供高度还原现实的环境,模拟无人机的动力和传感,还提供了飞行图像与日志的记录。在控制无人机飞行的过程中,仿真器可以保存无人机前向图像、惯性测量单元(IMU, inertial measurement unit)、速度、高度、坐标和姿态等数据,本文选取前向图像以及对应的姿态偏航角数据对无人机的方向控制进行训练。另外,出现障碍物或者即将发生碰撞时的前向图像会被标记为1,否则标记为0,以此训练碰撞概率的推断。总共收集的数据集包含32 000张图像,并且均匀分布在大小为 32×32 、 112×112 、 224×224 和 320×320 的图像中。

偏航角度的计算是一个回归问题,使用均方差(MSE, mean square error)作为误差函数,而碰撞概率是一个二分类问题,使用二元交叉熵(BCE, binary cross entropy)作为误差函数进行训练。但是,在使用多个损失函数进行联合优化的过程中会出现严重的收敛不均问题,因为每个损失函数产生的梯度的量级大小差别很大。更具体地说,在训练过程中对这两种损失函数不施加权重,会导致网络参数严重偏向大量级损失函数的方向进行收敛。因此在训练过程中需要考虑对两个损失函数施加动态权重,让两个梯度维持在一个可比的水平。因此本文在训练过程中随着轮次动态改变二元交叉熵的权重系数,而均方差的权重保持为1,如式(2)所示。

$$L_{\text{tot}} = L_{\text{MSE}} + \max(0, 1 - e^{-d(\text{epoch} - \text{epoch}_0)})L_{\text{BCE}} \quad (2)$$

在训练过程中将衰减值 d 设置为0.1,将 epoch_0 设置为10,使用初始学习率为0.001的Adam优化器进行优化并且每步指数衰减等于 10^{-5} 。为了减少训练过程中切换图像规模所花费的开销,在一个规模的图像数据集上训练完一个完整的轮次之后,系统会存储当前轮次的模型权重并切换到另一个规模大小的图像数据集开始下一个轮次的训练,并依次迭代。

3 强化学习动态配置

在面向云服务器的计算场景中,无人机将计算任务传输到远程云计算服务器中进行计算,计算完成后,云服务器再将结果返回给无人机(如图1中无人机B)。由于无人机匮乏的机载能源和其执行任务的性能要求,任务的卸载往往需要考虑能源消耗和计算时延这两个方面。导航任务的输入视频流传

输能耗过大, 无法避免对无人机的任务执行时间的缩短; 而无人机的导航任务处理时延会考虑传输图像的编解码时延、将任务传输到地面服务器的传输时延、当服务器繁忙时的排队时延和任务在服务器上的计算时延。对于实时性要求极高的无人机自主导航任务, 最小化时延的优先级高于最小化无人机的能耗, 因此本文将智能算法下沉到边缘设备端(如图 1 中无人机 A) 进行部署, 利用强化学习对卷积神经网络模型进行动态优化, 最大化无人机任务执行效用。

本文所采用的强化学习策略根据无人机所处的环境状态信息选出计算功率效用最优的卷积神经网络配置。强化学习通过智能体无人机与环境进行交互, 在每一个时间块 t , 智能体观测状态 s_t , 并且选择一个动作 a_t 。执行完动作之后, 智能体会收获奖励 r_t , 环境状态转换成 s_{t+1} 。训练的目标在于最大化预期累积奖励 $E\left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$, 其中 $\gamma \in (0, 1]$ 是未来奖励的衰减因子。

强化学习动态配置框架如图 4 所示, 图 4 描述了本文设计利用强化学习配置卷积网络计算的训练过程。强化学习的算法可以分为三大类: 基于策略梯度的强化学习、基于值函数的强化学习和基于 actor-critic 框架的强化学习。在基于策略梯度的强化学习中, 最优策略是通过直接操纵策略来计算的, 而基于价值的函数通过找到最优值函数来隐式地找到最优策略。基于策略梯度的强化学习在高维和随机的连续动作空间以及学习随机策略方面非常有效, 而基于值函数的强化学习在样品效率和稳定性方面表现出色。基于 actor-critic 框架的强化学习结合了基于策略和基于值函数的强化学习的优点, 是一类具有较好性能及收敛保证的强化学习方法, 因此本文所提方法主要基于 actor-critic 框架的强化学习算法架构。

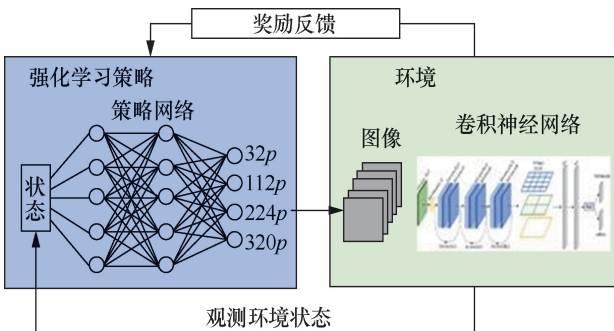


图 4 强化学习动态配置框架

3.1 系统模型

强化学习模型的目标在于让无人机在保持自主导航可靠性的情况下, 消耗尽可能少的计算功耗, 实现高可靠的无人机自主导航。无人机与环境的交互过程可以描述为马尔可夫决策过程 (MDP, Markov decision process), 其中状态空间、动作空间与奖励函数描述如下。

1) 环境状态

无人机飞行过程中周期性地对环境状态进行分析, 每次分析与配置的时间长度称为一个配置窗口, 每个配置窗口被分为 w 个小段, 包含了连续分布在 T 时间间隔的前向图像集合 (默认情况下 $T = 4\text{s}$), 飞行配置时间线如图 5 所示。在同一个配置窗口时间内无人机飞行所处的环境往往具有一致性, 因此对第一段时间 (默认情况下 $t = 1\text{s}$) 进行状态采集, 并根据策略网络选择动作进行配置, 并且在剩余的 $T - t$ 时间保持该配置。

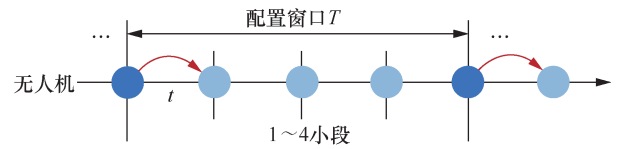


图 5 飞行配置时间线

经过时间段 t 之后, 将观测环境状态 $s_t = (x_t, o_t, \vec{a}_t)$ 输入策略网络中。其中, x_t 是 t 时间内的平均碰撞概率, $x_t \in (0, 1]$, o_t 是 t 时间内的平均环境复杂度, \vec{a}_t 前 k 次动作组成的集合向量, 即 $\vec{a}_t = (a_{t-k+1}, \dots, a_{t-1}, a_t)$, k 的大小表征了无人机的环境记忆能力。

环境复杂度^[26-28]是环境状态的重要组成, 它表征了图像中物体的线状特征的复杂度, 如树木、道路、湖泊、电线杆和房屋, 以及相应的渲染纹理边缘如软阴影、镜面反射、漫反射等。使用空间信息 (SI, spatial information) 来表示图像复杂度, 首先对图像进行灰度处理, 使用水平和垂直两个模块的索贝尔算子^[29]分别对灰度图像进行卷积, 索贝尔算子把图像中每个像素的上下左右 4 个域的灰度值加权差, 从而计算该像素的信息强度为

$$s_h = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} A s_v = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} A \quad (3)$$

其中, A 为原始灰度图像, 每个像素点的信息强度

取 s_h 与 s_v 中的最大值作为该像素点的输出值, 即 $s_r = \max(s_h, s_v)$ 。使用各个像素点输出值的平均值作为该图像的复杂度值, 可以使得当前环境不同规模大小的图像所测得的复杂度值大致相同, 并且不同环境的测量值差异明显, 可根据式(4)计算。

$$SI = \frac{1}{N} \sum_1^N s_r \quad (4)$$

其中, N 为图像中的像素点个数, 可由图像规模直接计算。不同环境复杂度灰度图像如图 6 所示, 分别为环境复杂度为 27.56、50.76、82.13 的灰度图像示例。

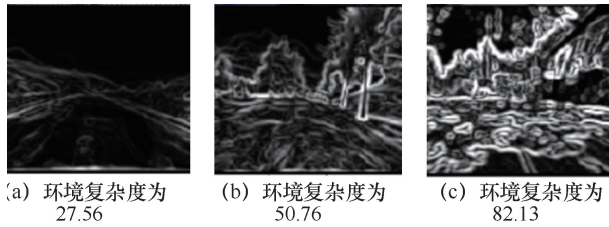


图 6 不同环境复杂度灰度图像

2) 动作空间

无人机获取环境状态 s_t 后, 会根据策略执行动作 a_t , 每一个动作对应不同的输入图像规模, 具体的规模集合包含 $\{32p, 112p, 224p, 320p\}$ 4 个动作, 不同的动作会产生不同的计算功率与环境感知力, 学习到的策略应该根据无人机所处的环境状态信息选出适用于当前状态的功率最优的卷积神经网络计算配置。

3) 奖励函数

策略选择动作的目标在于最大化从环境中获取的累计奖励, 因此奖励函数应使得无人机根据当前环境状态选择计算功率尽可能小的配置, 并且不能影响飞行导航的可靠性。即时奖励可分为两个部分: 环境匹配奖励部分与导航性能奖励部分。

$$r_\varepsilon = -\alpha(C_\varepsilon - P_\varepsilon)^2 - \beta E_\varepsilon \quad (5)$$

其中, r_ε 为一个配置窗口时间 T 结束之后无人机可以得到的奖励值, C_ε 为 $T-t$ 时间块的平均环境复杂度, P_ε 为 $T-t$ 时间块的平均计算功率, 在实际计算时, 计算功率将被缩放至与环境复杂度可比的大小。 E_ε 是碰撞概率分布的信息熵, 它通过测量散度来量化推断结果的不确定性, 散度越小推断结果越可靠。对于 E_ε 可由式(6)计算获得。

$$E_\varepsilon = -\frac{1}{\log q} \sum_{y \in C} f(y|x, w) \log(f(y|x, w)) \quad (6)$$

其中, x 为配置大小的图像帧输入, w 为神经网络权重, 障碍碰撞输出 $y \in C = \{0, 1, 2\}$, 神经网络的输出表示为 $f(y|x, \theta) \in [0, 1]$ 。

3.2 模型训练

本文采取如图 7 所示的 actor-critic 算法框架进行模型训练。对于每一个状态 s_t , 无人机根据训练的策略 $\pi(s_t, a_t) \rightarrow [0, 1]$ 选择动作, 在 actor-critic 算法中, 策略 π 由包含可调参数 θ 的策略神经网络表示, 即 $\pi_\theta(s_t, a_t)$, 动作执行之后会从环境中获取奖励, 训练策略神经网络的目标在于最大化累积回报的期望。本文采用策略梯度算法^[30]优化策略网络, 策略梯度的关键思想在于通过观察执行策略的轨迹来估计预期总奖励的梯度。策略参数 θ 所对应的累积折现奖励的梯度可以描述为

$$\nabla_\theta \mathbb{E}_{\pi_\theta} \left[\sum_{t=0}^{\infty} \gamma^t r_\varepsilon \right] = \mathbb{E}_{\pi_\theta} \left[\nabla_\theta \log \pi_\theta(s, a) A^{\pi_\theta}(s, a) \right] \quad (7)$$

其中, $A^{\pi_\theta}(s, a)$ 是优势函数表示在状态 s 下执行动作 a 所获得的累积奖励相比于基于策略的动作所获得的奖励的差异, 表明执行当前动作是否对策略优化具有增益。在实际运用中, 使用动作执行轨迹的采样, 并利用经验计算的优势 $A(s_t, a_t)$ 作为 $A^{\pi_\theta}(s, a)$ 的无偏估计, 每次根据如式(8)所示策略梯度更新策略网络的参数。

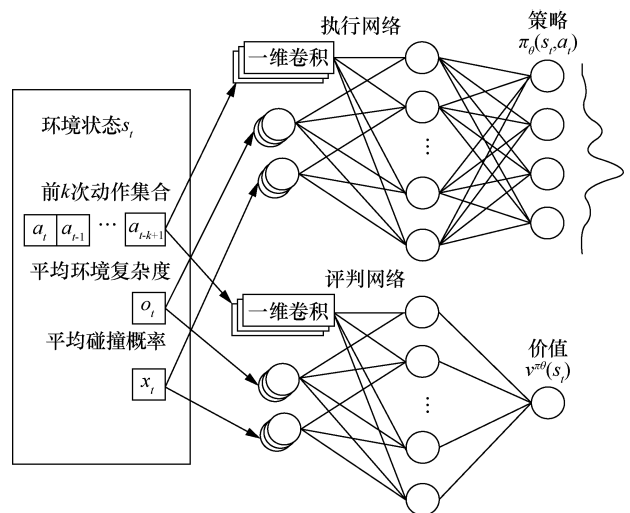


图 7 actor-critic 算法框架

$$\theta \leftarrow \theta + \alpha \sum_t \nabla_\theta \log \pi_\theta(s_t, a_t) A(s_t, a_t) + \beta \nabla_\theta H(\pi_\theta(\cdot | s_t)) \quad (8)$$

其中, α 为学习率, $\nabla_{\theta} \log \pi_{\theta}(s_t, a_t)$ 表示策略更新参数的方向, 更新的幅度取决于当前状态 s_t 下执行动作 a_t 所取得的优势函数大小, 因此梯度会向能获得更多奖励的方向上更新参数。在参数的更新规则中加入熵正则项可以让策略网络在训练过程中充分探索动作空间, 以收敛到一个更好的策略。其中 $H(\cdot)$ 是策略网络输出的动作概率分布的熵, 能够让参数 θ 往熵增的方向更新, 参数 β 在训练开始阶段设为较大的值 (鼓励探索), 并在训练过程中逐渐减小以强调提高奖励值。

为了能够计算一个给定的经验的优势 $A(s_t, a_t)$, 需要对 actor-critic 算法框架中的状态价值函数 $v^{\pi_{\theta}}(s)$ 进行估计, $v^{\pi_{\theta}}(s)$ 表示从状态 s 出发采用策略 π_{θ} 所能获得的奖励累积期望, 评判网络的作用就在于从经验观测到的奖励中学习 $v^{\pi_{\theta}}(s)$ 的估计, 可以使用标准时序差分法来训练评判网络参数 θ_v 。

$$\theta_v \leftarrow \theta_v - \alpha' \sum_i \nabla_{\theta_v} (r_i + \gamma V^{\pi_{\theta}}(s_{i+1}; \theta_v) - V^{\pi_{\theta}}(s_i; \theta_v))^2 \quad (9)$$

其中, $V^{\pi_{\theta}}(\cdot; \theta_v)$ 是由评判网络输出的 $v^{\pi_{\theta}}(\cdot)$ 的估计, α' 是评判网络的学习率, 对于执行一次动作的经验 (s_t, a_t, r_t, s_{t+1}) 其优势 $A(s_t, a_t)$ 可以估计为 $r_t + \gamma V^{\pi_{\theta}}(s_{t+1}; \theta_v) - V^{\pi_{\theta}}(s_t; \theta_v)$ 。

为了生成动态配置算法, 将前 $k=5$ 次的动作集合向量输入包含 128 核的一维卷积层中, 每一个卷积核大小为 4, 步长为 1。第一层的结果将被聚合成隐藏层的输入, 隐藏层包含 128 个神经元通过 Softmax 激活函数函数输出 4 个动作的概率分布。评判网络采用相同的神经网络结构, 但是使用线性神经元作为最后的输出 (无激活函数)。在训练过程中, 衰减因子 γ 设置为 0.99, 即当前的动作将影响未来 100 个步骤, actor 网络和 critic 网络的学习率分别设置为 10^{-4} 和 10^{-3} 。另外, 熵权重因子 β 在整个迭代过程从 1 到 0.1 逐渐衰减。在实验过程中这些超参数保持不变。

强化学习训练过程在 AirSim 仿真环境中进行, 无人机根据 actor 网络的输出选择动作, 然后从环境中获取奖励更新网络参数。当无人机在环境中与障碍物发生碰撞、超出当前路径范围或者到达路径终点时结束当前训练轮次。出于兼容性考虑, 采用基于 TensorFlow 后端的 Keras 深度学习库实现算法框架, 同时用于训练和实验阶段。计算平台使用的 CPU 是英特尔酷睿 i7-9850H, GPU 使用 GeForce RTX 2080Ti, 内存为 32 GB。算法训练收敛曲线如

图 8 所示。实验中, 当训练进行到 2 000 轮次的时候, 模型开始收敛到一个较好的水平。

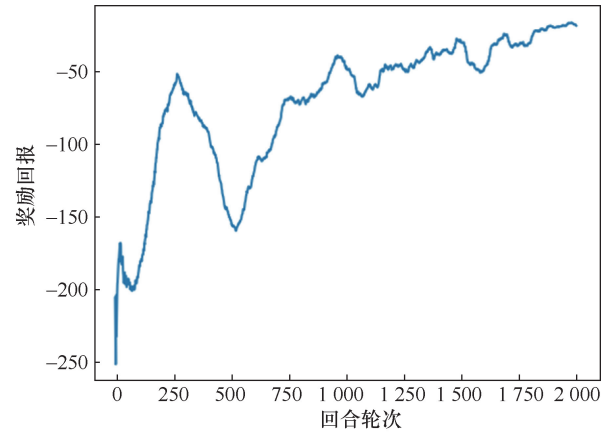


图 8 算法训练收敛曲线

4 仿真实验分析

本节先在仿真环境下对所提的模型动态优化算法进行验证与分析, 并与传统的两种算法进行对比, 验证了所提算法的有效性。结合无人机输入图像数据的本地性与搭载设备的强计算能力, 系统将自主导航卷积模型与强化学习智能算法下层到边缘设备终端部署, 在边缘终端侧利用智能模型实时优化模型效用, 解决自主导航任务实时性敏感的问题, 同时提高无人机的续航能力。为了模拟低时延的工作环境, 仿真实验利用机器人操作系统 (ROS, robot operating system) 构建软件模型运行在高性能服务器中, 通过进程端口与仿真器通信, 利用进程间的通信模拟硬件串口通信, 使用预先测量的输入图像尺寸与功率之间的映射关系来估计能耗的模拟。

本文所使用的对比算法包括以下两种设置。

- 1) 根据当前环境复杂度区间选择卷积网络输入大小, 将环境复杂度区间从小到大划分为 4 个区间: [0,30]、(30,60]、(60,90]、(90,255], 每一个区间按照大小分别对应一个动作, 无人机根据当前环境复杂度对应的区间阈值选择动作^[31];
- 2) 传统静态配置算法, 在飞行过程中, 无论环境复杂度如何, 始终保持输入图像为 224×224 大小不变。

算法所使用的仿真环境是 AirSimNH^[25], 它对应着一个小城镇的虚拟地图。无人机在环境中沿着所处道路前行, AirSim 仿真训练环境如图 9 所示。使用模型推断的碰撞概率 p_i 控制无人机的飞行速度 v_k , 让碰撞概率与速度近似成反比, 采用低通过滤

的方式计算出连续光滑的速度值 v_k 为

$$v_k = (1 - \alpha)v_{k-1} + \alpha(1 - p_t)V_{\max} \quad (10)$$

其中, $0 \leq \alpha \leq 1$ 。类似的, 将推断的偏航角 s_k 映射为无人机 z 轴的旋转角度 θ_k , 将 s_k 从 $[-1, 1]$ 的范围转换到范围为 $[-90^\circ, 90^\circ]$ 的偏航角度, 并且对其低通滤波平滑处理如式(11)所示。

$$\theta_k = (1 - \beta)\theta_{k-1} + \beta \frac{\pi}{2} s_k \quad (11)$$

在实际部署中, 将 α 设为 0.7 并将 β 设为 0.5, V_{\max} 根据实际飞行环境进行调整。上述参数的选择根据经验, 以牺牲一定可靠性作为代价, 提高无人机的平稳性。

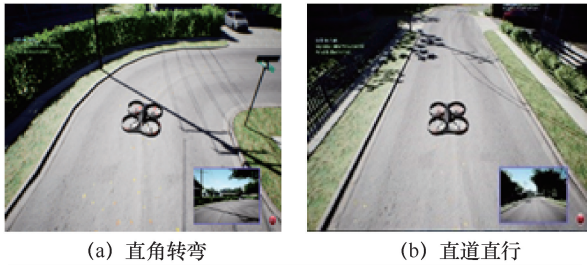


图 9 AirSim 仿真训练环境

如图 9 所示, 实验区域包含 4 个直角转弯以及 4 个直道, 在回合进行过程中可以获取地图上路的坐标信息, 返回的坐标包括道路拐弯和交叉处的斜线坐标, 这些坐标会参与距离的计算。每一段为长度 20 m 并且具有不同的环境复杂度的道路。实验采用以下指标对比不同的算法性能: 距离、成功率和平均功率。距离是连续飞行长度, 如果错过交叉点、发生碰撞或到达终点, 距离测量将停止。成功率是实验过程中到达终点回合数占总回合数的比率。功率是在仿真飞行中模拟的能量消耗与飞行时间的比率。使用预先测量的输入尺寸与功率之间的映射关系来估计能耗的模拟, 并且记录飞行时间, 以此计算平均功率。所有的测量结果在 6 次实验中取平均值获得。3 种算法飞行指标数据如表 1 所示。

表 1 3 种算法飞行指标数据

算法	距离/m	成功率	功率/W
静态配置	73.6	89%	6.1
阈值配置	67.2	83%	4.2
强化学习配置	77.6	92%	3.9

为了保证对比算法处于最佳的推断条件, 所有导航模型的训练使用相同的数据与训练轮次。由表 1

可知, 本文的强化学习算法(强化学习配置)在距离、成功率和功率 3 个指标中都优于其他对比算法。在路程为 80 m 的实验场景中, 本文提出的强化学习配置方法下无人机平均飞行距离为 77.6 m, 飞行的成功率为 92%, 并且消耗的平均功率仅为 3.9 W。与其他两种对比方法相比, 本文提出的强化学习配置方法能够使连续飞行的距离延长 5%和 13%, 飞行成功率提高了 3%和 9%, 消耗的平均功率节省了 36%和 7%。综合比较而言, 本文方法在距离、成功率和功率方面取得有效的综合性能提升, 并且在强化学习配置与静态配置两种算法所消耗的平均功率比较结果中可以印证强化学习计算所带来的能耗开销小于动态配置所带来的能效增益。

5 系统原型验证

为了进一步验证算法的性能, 在现实环境中实现了一套原型系统, 无人机硬件架构如图 10 所示, 并在其上部署了本文提出的算法。让无人机沿着 2 m 宽的轨迹飞行, 飞行过程中保持 2 m 定高, 最大速度为 1.5 m/s, 导航模型推断频率保持为 15 Hz。无人机使用 DJI F540 机架, 飞行控制器使用运行 PX4 LPE 固件的 Pixhawk, 摄像头使用 MyntEye 相机, 在图像采集时只使用左侧摄像头进行图像输入。为了能够提供足够的计算性能, 无人机机载计算设备采用 Nvidia Jetson TX2 核心版, 表 2 给出了其硬件规格信息。电池采用 5200-Mah 容量的 3s 锂电池, 电池不仅给无人机供电还通过 UBEC (ultra battery elimination circuit) 电调给 TX2 计算设备供电。Nvidia Jetson TX2 规格如表 2 所示。

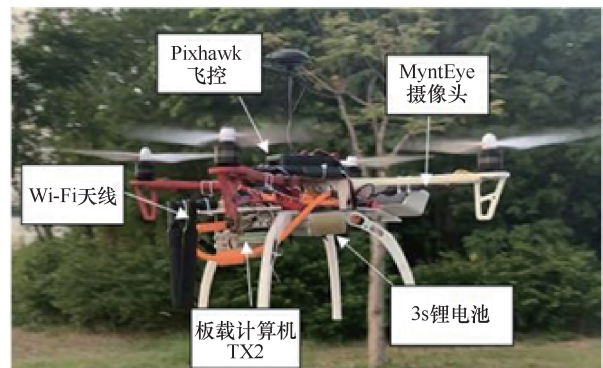


图 10 无人机硬件架构

软件框架基于 ROS 搭建, 包含 4 个 ROS 节点进行协同工作: 1) 深度摄像头节点通过 MyntEye API 获取实时的视频图像进行发布; 2) 视觉导航节点接

收来自深度摄像头节点进行卷积计算, 推理出偏航角与碰撞概率; 3) 指令转化节点将视觉导航节点的控制信息转化为底层指令发送给无人机飞控模块, 控制无人机飞行; 4) 动态配置节点获取当前的环境状态信息, 执行卷积神经网络的配置动作, 调整视频帧与对应输入大小配置。数据及指令通过微型航空器连接 (MavLink, micro air vehicle link) 协议进行传输。

表 2 Nvidia Jetson TX2 规格

CPU	2.00 GHz 四核 ARM Cortex-A572.00 GHz 双核 Denver 2	
内存	1 600 MHz 8 GB LPDDR4	
GPU	Pascal 架构, 256 个 CUDA 核心	
功耗/W	理想状态	5
	满载状态	15
	平均	9.5

在执行自主导航任务之前, 所有 ROS 节点以及摄像头通过 SSH (secure shell) 远程连接启动。先让无人机从自稳模式调整至起飞位置, 然后切换至板载模式让无人机执行 TX2 发送的上层指令。

在实验中对无人机连续自主导航飞行的持续时间进行比较, 评估计算功耗性能。实验前, 测试了无人机搭载 TX2 机载计算板 (150 g)、MyntEye 摄像头 (142 g) 以及 5 200 mAh 容量的 3s 锂电池设备的情况下, 采用不同算法进行自主导航时的续航时间, 当无人机电量耗尽时停止时间测量。为了确保公平性, 无人机在每一次飞行前, 电池电量都是充满状态。3 种算法续航时间如表 3 所示。

表 3 3 种算法续航时间

算法	静态配置	阈值配置	强化学习配置
续航时间/s	821	879	945

由表 3 可知, 使用本文提出的基于强化学习的动态配置算法 (强化学习配置) 相对于传统静态配置方法, 在使用计算设备执行任务时可显著降低设备的计算开销, 延长续航时间。具体而言, 静态配置算法与阈值配置算法续航时间均不足 900 s, 而本文提出的强化学习配置方法续航时间长达 945 s, 与对比算法续航时间相比分别提高了 15% 和 7.5%, 有效扩展了无人机自主导航任务的运用范围。

6 结束语

本文研究了一种基于边缘智能驱动的无人机

自主导航飞行方法。本文提出了一种依据无人机所处环境状态对卷积神经网络进行动态配置的深度强化学习算法, 有效提高了无人机自主导航的准确性, 降低无人机的计算功耗, 延长了续航时间。基于仿真模拟和原型验证的多重实验结果表明, 相对于传统采用静态配置的算法以及基于阈值配置的算法, 本文提出的算法最高可使无人机自主导航飞行成功率提升 9%, 使真实续航距离提升 15%。

目前, 本文所提框架模型的推理计算依赖于无人机的本地计算设备, 虽然能够满足计算实时性, 但是没能综合考虑使用边缘服务器进行协同计算。对于未来的工作, 可以考虑采用模型分割等方式, 将无人机本地机载设备与边缘服务器进行协同, 采用分布式执行自主导航神经网络的架构, 在避免传输视频流所带来的传输能耗的同时减轻无人机本地设备的计算压力, 进一步提高系统的能效。本文所提算法具备一定的普遍性, 不仅可用于自主导航任务, 还能扩展到其他应用领域。

参考文献:

- [1] FAESSLER M, FONTANA F, FORSTER C, et al. Autonomous, vision-based flight and live dense 3D mapping with a quadrotor micro aerial vehicle[J]. *Journal of Field Robotics*, 2016, 33(4): 431-450.
- [2] SCHERER S, REHDER J, ACHAR S, et al. River mapping from a flying robot: state estimation, river detection, and obstacle mapping[J]. *Autonomous Robots*, 2012, 33(1/2): 189-214.
- [3] PALOSSI D, GOMEZ A, DRASKOVIC S, et al. Extending the lifetime of nano-blimps via dynamic motor control[J]. *Journal of Signal Processing Systems*, 2019, 91(3/4): 339-361.
- [4] ZHOU Z, CHEN X, LI E, et al. Edge intelligence: paving the last mile of artificial intelligence with edge computing[J]. *Proceedings of the IEEE*, 2019, 107(8): 1738-1762.
- [5] SHEN S J, MULGAONKAR Y, MICHAEL N, et al. Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV[C]//2014 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2014: 4974-4981.
- [6] WEISS S, SCARAMUZZA D, SIEGWART R. Monocular-SLAM-based navigation for autonomous micro helicopters in GPS-denied environments[J]. *Journal of Field Robotics*, 2011, 28(6): 854-874.
- [7] ROSS S, MELIK-BARKHUDAROV N, SHANKAR K S, et al. Learning monocular reactive UAV control in cluttered natural environments[C]//2013 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2013: 1765-1772.
- [8] GIUSTI A, GUZZI J, CIREŞAN D C, et al. A machine learning approach to visual perception of forest trails for mobile robots[J]. *IEEE Robotics and Automation Letters*, 2016, 1(2): 661-667.
- [9] SMOLYANSKIY N, KAMENEV A, SMITH J, et al. Toward low-flying autonomous MAV trail navigation using deep neural networks for environmental awareness[C]//2017 IEEE/RSJ International

- Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2017: 4241-4247.
- [10] MEHROOZ G, EBEID E, SCHNEIDER-KAMP P. System design of an open-source cloud-based framework for Internet of drones application[C]//2019 22nd Euromicro Conference on Digital System Design (DSD). Piscataway: IEEE Press, 2019: 572-579.
- [11] PALOSSI D, LOQUERCIO A, CONTI F, et al. A 64-mW DNN-based visual navigation engine for autonomous nano-drones[J]. IEEE Internet of Things Journal, 2019, 6(5): 8357-8371.
- [12] 李肯立, 刘楚波. 边缘智能: 现状和展望[J]. 大数据, 2019, 5(3): 69-75.
LI K L, LIU C B. Edge intelligence: state-of-the-art and expectations[J]. Big Data Research, 2019, 5(3): 69-75.
- [13] 莫梓嘉, 高志鹏, 苗东. 边缘智能: 人工智能向边缘分布式拓展的新触角[J]. 数据与计算发展前沿, 2020, 2(4): 16-27.
MO Z J, GAO Z P, MIAO D. Edge intelligence: a new exploration for artificial intelligence expanding to edge[J]. Frontiers of Data & Computing, 2020, 2(4): 16-27.
- [14] ZHANG X Z, WANG Y F, LU S D, et al. OpenEI: an open framework for edge intelligence[C]//2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS). Piscataway: IEEE Press, 2019: 1840-1851.
- [15] FRAGKOS G, KEMP N, TSIROPOULOU E E, et al. Artificial intelligence empowered UAVs data offloading in mobile edge computing[C]//ICC 2020-2020 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2020: 1-7.
- [16] 张星洲, 鲁思迪, 施巍松. 边缘智能中的协同计算技术研究[J]. 人工智能, 2019, 6(5): 55-67.
ZHANG X Z, LU S D, SHI W S. Research on collaborative computing technology in edge intelligence[J]. Artificial Intelligence, 2019, 6(5): 55-67.
- [17] 刘通, 方璐, 高洪皓. 边缘计算中任务卸载研究综述[J]. 计算机科学, 2021, 48(1): 11-15.
LIU T, FANG L, GAO H H. Survey of task offloading in edge computing[J]. Computer Science, 2021, 48(1): 11-15.
- [18] MCMAHAN B, RAMAGE D. Federated learning: collaborative machine learning without centralized training data[J]. Google Research Blog, 2017: 3.
- [19] FANG B Y, ZENG X, ZHANG M. NestDNN: resource-aware multi-tenant on-device deep learning for continuous mobile vision[C]//MobiCom'18: Proceedings of the 24th Annual International Conference on Mobile Computing and Networking. [S.l.:s.n.], 2018: 115-127.
- [20] 董超, 沈赞, 屈毓铨. 基于无人机的边缘智能计算研究综述[J]. 智能科学与技术学报, 2020, 2(3): 227-239.
DONG C, SHEN Y, QU Y B. A survey of UAV-based edge intelligent computing[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(3): 227-239.
- [21] SIMONYAN, KAREN, ZISSERMAN A, et al. Very deep convolutional networks for large-scale image recognition[J]. arXiv:1409.1556, 2014.
- [22] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2015: 1-9.
- [23] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 770-778.
- [24] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [25] SHAN S, DEY D, LOVETT C, et al. Airsim: high-fidelity visual and physical simulation for autonomous vehicles[J]. Field and Service Robotics, 2018: 621-635.
- [26] YU H H, WINKLER S. Image complexity and spatial information[C]//2013 5th International Workshop on Quality of Multimedia Experience (QoMEX). Piscataway: IEEE Press, 2013: 12-17.
- [27] PERLIO J, HYVARINEN A. Modelling image complexity by independent component analysis, with application to content-based image retrieval[J]. International Conference Artificial Neural Networks, 2009: 704-714.
- [28] ROMERO J, MACHADO P, CARBALLAL A, et al. Using complexity estimates in aesthetic image classification[J]. Journal of Mathematics and the Arts, 2012, 6(2/3): 125-136.
- [29] KAZAKOVA N, MARGALA M, DURDLE N G. Sobel edge detection processor for a real-time volume rendering system[C]//2004 IEEE International Symposium on Circuits and Systems. Piscataway: IEEE Press, 2004.
- [30] AWATE Y P. Policy-gradient based actor-critic algorithms[C]//2009 WRI Global Congress on Intelligent Systems. Piscataway: IEEE Press, 2009: 505-509.
- [31] PALACIN J, PALLEJA T, VALGANON I, et al. Measuring coverage performances of a floor cleaning mobile robot using a vision system[C]//Proceedings of the 2005 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2005: 4236-4241.

[作者简介]



林椿珉 (1997-), 男, 中山大学计算机学院硕士生, 主要研究方向为无人机自动驾驶、边缘计算、边缘智能等。



曾烈康 (1996-), 男, 中山大学计算机学院博士生, 主要研究方向为移动边缘计算、协同智能计算、边缘智能等。



陈旭 (1986-), 男, 博士, 中山大学计算机学院教授、先进网络与计算系统研究所所长、数字家庭互动应用国家地方联合工程实验室副主任, 主要研究方向为边缘计算、边缘智能、智能物联网。