****

**本科毕业设计（论文）开题报告**

|  |  |
| --- | --- |
| **论文题目** | 基于局部视觉信息的路径规划 |

|  |  |
| --- | --- |
| **学 院：** | 自动化科学与工程学院 |
| **专 业：** | 自动化 |
| **学生姓名：** | 徐海鑫 |
| **学生学号：** | 201830040420 |
| **指导教师：** | 罗家祥 |

|  |  |  |
| --- | --- | --- |
| 1. **课题背景及意义（含国内外研究现状综述）**   近年来，机器人产业发展迅速，尤其是扫地机器人的成功问世，更是将移动机器人带入了千家万户。伴随着这些产品的成功落地，机器人的智能化程度也越来越高，其中，作为机器人自主导航的核心技术，路径规划问题得到了广泛的关注。  移动机器人（Mobile Robot）指的是能够自主或按照预定程序改变机器人所在位置和姿态的一类机械装置。早在1956年，美国斯坦福研究所就已经开始了移动机器人的研究，并于1972年成功研制出了世界上第一条移动机器人Shakey，斥资75万美金，全面应用了人工智能技术，并配备了三角测距仪、电子摄像机等多种传感器。近年来，移动机器人产业更是发展迅速，应用场景也不断扩展，自动耕种机器人[1]，海底探测机器人[2]，工厂流水线作业机器人等新型机器人层出不穷，机器人逐渐在人类社会中起到了不可磨灭的作用，尤其是在工业领域。2018-2020年间，全球工业机器人的销售量和销售额逐年增长， 2020年，工业移动机器人出货量70602台，销售23.588亿美元，分别同比增长42%和25%。  为了让机器人在各种各样的环境那种出色的完成任务，必须让机器人拥有感知周围环境并且进行路径规划的能力，这就需要用到各种各样的传感器。移动机器人常用的传感器有激光雷达、相机、IMU等等，而前两种由于具有较良好的鲁棒性和可跟随性，占据了大部分的市场份额。  激光雷达具有快速精确的特点，早在2005年，就已经被研究的比较透彻，并且框架也已确定，是当今市场上最稳定且最主流的传感器，但是激光雷达本身也存在许多缺点，例如成本高，体积大等等，所以市场上急需一种可以替代其的产品，而相机正是其中之一。虽然基于相机的导航定位仍在研发当中，却已经展现出其许多优势。实验证明，相较于传统的激光雷达，视觉传感器受磁场或传感器的干扰更小，同时也是最接近人类习惯的一种方式，采样速率也远高于激光雷达，由此可以看出，基于视觉传感器的导航定位技术具有着良好的发展前景。    图1 视觉导航关键技术  随着计算机视觉的火热发展，将视觉算法应用于机器人导航定位的技术也愈发成熟。视觉导航主要有三个关键技术点[3]。视觉定位是通过移动机器人身上安装的视觉传感器对周围的环境进行捕获，在经过一系列的图像处理之后，获得当前机器人的位姿，基于环境模型的定位方法主要有两种：相对定位和绝对定位。相对定位又称航迹推算，即已知初始位姿，通过计算相邻两帧的相对位姿，推算出运动轨迹。因为当前位姿是历史各时刻相对位姿的累加，所以容易造成较大的累计误差。视觉里程计由Matthies 等人[4]提出，采用特征值提取和特征匹配等方法估计原始模型。赵黎明[5]通过改进SIFT 算法，提高了实时性，减小了实际偏差。而绝对定位方法则不需要知道机器人的初始位姿，只需要通过摄像头提取环境特征，然后与地图中的关键图像库进行匹配，确定当前机器人的位置。Royer 等[6]提出了一种基于单目视觉训练的导航方法，该方法先建立完整的环境地图，然后利用建立好的地图导航。环境地图的表达方法主要有栅格地图[7]、几何地图、拓扑地图以及综合几种地图优势的混合地图。  基于视觉的室内导航技术是通过视觉传感器对周围环境进行捕获，进行障碍物与非障碍物的识别，规划出一条可行的路径，从而实现机器人的导航。常见的移动机器人视觉导航方法有：基于环境信息的地图模型匹配[8]，同步定位与地图构造[9]以及不依赖地图导航。  路径规划就是使目标对象在最小的时间或距离代价下，在规定区域范围内找到一条从起点到终点的安全无碰撞路径。目前，路径规划的方法主要有常规方法和强化学习方法，其中常规方法主要有传统方法、图形学方法及智能仿生学方法等，而随着对移动机器人智能化需求的上升，基于强化学习的路径规划方法逐渐进入大众视野并且取得了显著的效果[10]。  强化学习的方法主要有基于值的以及基于策略的，随后，又发展出了基于值和策略相结合的强化学习方法。基于值的强化学习方法主要有TD, Q-Learning[11]和SARSA[12]方法，主要适用于离散动作空间，目标是通过最大化每个状态的值函数来得到最优策略；基于策略的强化学习方法主要有策略梯度（Policy Gradient，PG）[13]、模仿学习（Imitation Learning，IL）[14]等，可用于连续的动作空间，而基于值和策略相结合的Actor-Critic[15]方法则融合了两者的优点，比传统的策略梯度法效率高，同时可应用于连续动作空间，有较为广阔的应用空间。  MUSE等[16]成功将Actor-Critic方法应用于移动机器人路径规划任务中；LACHEKHAB 等[17]提出一种模糊Actor-Critic 学习算法（Fuzzy Actor-Critic Learning Algorithm，FACL），基于该算法实现机器人从起点到终点的路径规划，并在模糊规则中基于概率选择机器人的下一行动，可以在已知环境地图的情况下较好地完成路径规划的任务。SHAO等[18]提出一种基于广义优势估计（Generalized Advantage Estimator，GAE）的Actor-Critic 算法，使智能体可以从多个过程中进行学习以节约训练时间，之智能体可以从多个过程中进行学习，从而减少训练时间，并利用GAE 估计优势函数，从而提高了估计精度。 | | |
| 1. **课题研究主要内容及研究基础**   课题研究内容  本课题主要研究基于局部视觉信息的机器人路径规划方法，总结现有路径规划方法的优缺点，尤其是有障碍物情况下的路径规划，并设计基于强化学习的路径规划方法，首先在pettingzoo环境中进行简单仿真，然后在GAZEBO环境中搭建与实际环境较为接近的智能车路及规划算法实验平台，先进行无障碍物的简单训练，然后总结归纳训练结果，优化训练数据，再在环境中加入障碍物，训练智能车在有障碍物环境下进行路径规划，不断改进训练数据与特征提取方法以及奖励方式，优化训练效果，分析算法的有效性。  研究基础   1. Actor-Critic方法   (1)AC方法  Actor-Critic先更新价值函数V，再基于价值函数更新动作策略。策略梯度的方法只能执行完一个回合之后更新，因为要计算再当前动作之后所用的利益的累计，而Actor-Critic可以单步更新，用Q\_Learning方法可以单步更新价值，然后借助V来计算梯度，更新策略。  (2)A2C方法  AC方法中需要估计两个网络：Q 网络和V 网络，估测不准的风险就变成两倍。事实上在演员-评论员方法里面，我们可以只估测网络V，可以用V 的值来表示Q 的值。  有个初始的演员去跟环境做互动，先收集资料。拿这些资料去估计价值函数，可以用时序差分或蒙特卡罗来估计价值函数。接下来，你再基于价值函数，套用下式去更新π。    图2 A2C方法流程  (3)A3C方法  异步优势演员- 评论员算法方法就是同时开很多个进程（worker），当第一个进程做完想要把参数传回去的时候，本来它要的参数是θ1，等它要把梯度传回去的时候。可能别人已经把原来的参数覆盖掉，变成θ2 了。但是没有关系，它一样会把这个梯度就覆盖过去就是了。这个就是异步优势演员- 评论员算法。    图3 A3C方法流程   1. PER机制   PER优先经验回放机制的思路来源与优先清理(Prioritized sweeping),它以更高的频率回放对学习过程更有帮助的样本，从而加快模型的收敛速度。如何界定某个样本是否对模型的学习有更大的帮助呢？这里使用TD-error作为评判指标，TD-error越大，则认为该样本具有更多值得学习的经验，所以应当以更大的概率重放该样本。但是每次训练后并不会更新每一个样本的TD-error,这就导致有些样本的TD-error本身很小，一直无法被采样，TD-error也得不到更新。最后，PER会在很小的范围内重放，导致过拟合。  为了解决上述问题，采用了结合纯贪心采样和和均匀分布采样的随机采样方法，确保在样本的优先级中采样的概率是单调的，同时即使对于最低优先级的样本也要保证非零概率。定义采样的概率为：  对于优先级的定义，存在两种变体：  （1）第一种变体：proportional prioritization(比例优先级)  其中为TD-error,而则是为了防止样本的TD-error为零时不再被回放。  （2）第二种变体：rank-based prioritization(基于排名的优先级)  这里定性地考虑priority，没有定量地考虑priority。实现类似分层抽样，事先将排名段分为几个等概率区间，再在各个等概率区间里面均匀采样，其中在采样的过程中采用了退火偏差处理。 | | |
| 1. **研究（或调研）方案和思路（技术路线）**   研究方案   1. 首先进行现存的基于强化学习的路径规划算法的论文进行复现，并针对复现结果提出相关的改进措施， 2. 在python中实现改进的算法，并对改进算法的效果进行分析 3. 应用强化学习算法进行移动机器人的实时定位，在仿真平台进行仿真，获取提出算法在遇到位置障碍物情况下进行路径规划的效果。   研究思路   1. 查阅有关强化学习的文献，理解其中算法基本的数理原理 2. 根据A2C算法原理实现移动机器人的导航 3. 对实现的基于强化学习的移动机器人导航方法进行测试与评价分析，确保实现的算法切实可行 4. 根据测试结果进行进一步优化 | | |
| 1. **论文框架结构**   本论文共五章，各章节安排如下：  第一章 绪论：首先介绍了移动机器人及其相关技术的研究背景及意义。然后，调研了目前主流的基于视觉的移动机器人导航定位技术，并且研究了主要的移动机器人路径规划方法。最后，给出了本文的研究内容及各章节安排。  第二章 介绍了本文所用到的计算机视觉及强化学习相关知识。  第三章 针对在已知部分地图，在未知区域随机出现障碍物的情况，提出了一种基于强化学习的路径规划算法。  第四章 在GYM平台对提出的算法进行仿真，并在GAZEBO平台对进行虚拟环境的构建，并且进行基于提出的强化学习避障路径规划算法的移动机器人在虚拟环境下运行的仿真。  第五章 总结与展望：对本文的全部研究内容进行了总结归纳，并提出了有待提高的部分，最后对接下来的研究工作进行了构想与展望。 | | |
| 1. **参考文献** 2. A. English, P. Ross, D. Ball and P. Corke, "Vision based guidance for robot navigation in agriculture," 2014 IEEE International Conference on Robotics and Automation (ICRA), 2014, pp. 1693-1698, doi: 10.1109/ICRA.2014.6907079. 3. S. Kobayashi, "Polychaete Like Omnidirectional Propulsion Robot in Water." Journal of the Robotics Society of Japan 33.1(2015):21-24. 4. 马凯,林义忠.移动机器人视觉导航技术综述[J].物流科技,2020,43(10):39-41+46.DOI:10.13714/j.cnki.1002-3100.2020.10.010 5. L. Matthies and S. Shafer, "Error modeling in stereo navigation," in IEEE Journal on Robotics and Automation, vol. 3, no. 3, pp. 239-248, June 1987, doi: 10.1109/JRA.1987.1087097. 6. 赵黎明. 基于单目视觉的港口AGV 自主导航关键技术研究[D]. 厦门：集美大学（硕士学位论文），2015. 7. E. Royer, M. Lhuillier, M. Dhome, et al, "Monocular Vision for Mobile Robot Localization and Autonomous Navigation," International Journal of Computer Vision, 2007,74(3):237-260. 8. H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," Proceedings. 1985 IEEE International Conference on Robotics and Automation, 1985, pp. 116-121, doi: 10.1109/ROBOT.1985.1087316. 9. J. Borenstein, H. R. Everett and L. Feng, "Navigating mobile robots: systems and techniques." Wellesley: A K Peters Led, 1996. 10. R. Smith, M. Self and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," Proceedings. 1987 IEEE International Conference on Robotics and Automation, 1987, pp. 850-850, doi: 10.1109/ROBOT.1987.1087846. 11. 闫皎洁,张锲石,胡希平.基于强化学习的路径规划技术综述[J].计算机工程,2021,47(10):16-25.DOI:10.19678/j.issn.1000-3428.0060683. 12. C. Watkins, J. Christopher, and P. Dayan, "Q-learning." Machine Learning 8.3(1992):279-292.. 13. G. A. Rummery and M. Niranjan, "On-Line Q-Learning Using Connectionist Systems." Technical Report (1994). 14. 刘建伟，高峰，罗雄麟. 基于值函数和策略梯度的深度强化学习综述［J］. 计算机学报，2019，42（6）：1406-1438. 15. Q. Wang, D. Xu and L. Shi, "A review on robot learning and controlling: imitation learning and human-computer interaction," 2013 25th Chinese Control and Decision Conference (CCDC), 2013, pp. 2834-2838, doi: 10.1109/CCDC.2013.6561428. 16. I. Grondman, L. Busoniu, G. A. D. Lopes and R. Babuska, "A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 42, no. 6, pp. 1291-1307, Nov. 2012, doi: 10.1109/TSMCC.2012.2218595. 17. D. Muse and S. Wermter . "Actor-Critic Learning for Platform-Independent Robot Navigation." Cognitive Computation 1.3(2009):203-220. 18. F. Lachekhab and M. Tadjine, "Goal seeking of mobile robot using fuzzy actor critic learning algorithm," 2015 7th International Conference on Modelling, Identification and Control (ICMIC), 2015, pp. 1-6, doi: 10.1109/ICMIC.2015.7409370. 19. K. Shao, D. Zhao, Y. Zhu and Q. Zhang, "Visual Navigation with Actor-Critic Deep Reinforcement Learning," 2018 International Joint Conference on Neural Networks (IJCNN), 2018, pp. 1-6, doi: 10.1109/IJCNN.2018.8489185. | | |
| 1. **工作进度安排** | | |
| **序号** | **设计（论文）各阶段任务** | **时间安排** |
| 1 | 调研基于局部视觉信息的机器人路径规划方法 | 1.1-3.1 |
| 2 | 总结现有路径规划方法的优缺点，尤其是有障碍物情况下的路径规划 | 3.1-3.15 |
| 3 | 进行外文文献翻译 | 3.15-3.20 |
| 4 | 设计基于强化学习的路径规划方法，对未遇到过的障碍物的避障策略进行学习 | 3.20-4.20 |
| 5 | 利用Python对算法进行仿真，分析算法的有效性 | 4.20-5.20 |
| 6 | 进行毕业论文的撰写 | 5.20-5.31 |