****

本科毕业设计（论文）

**基于局部视觉的路径规划算法研究**

|  |  |
| --- | --- |
| **学院**  **专业**  **学生姓名**  **学生学号**  **指导教师**  **提交日期** | **自动化科学与工程学院** |
| **自动化专业** |
| **徐海鑫** |
| **201830040420** |
| **罗家祥** |
| **年 月 日** |

# 摘 要

（标题：小二号，黑体，居中，单倍行距，段前、段后各0.5行，两字中间空2字符）

（摘要正文共400—600个字；小四号，宋体，1.5倍行距，段首行空两个汉字）

炔烃和叠氮化合物的点击化学反应，有着快速、百分百原子利用率、产物高选择性等众多优点，被誉为点击化学中的精华。基于此反应拓展而来的点击聚合反应，迅速在高分子材料领域获得了了广泛关注和应用。

……

我们还尝试了采用不同单体，在最优条件下进行反应，均获得了高分子产物。表明了该反应体系的普适性。

（此处隔一行）

关键词：计算机视觉；强化学习；路径规划

（“关键词”：小四号，黑体；关键词3—5个：小四号，宋体；关键词之间用分号隔开；最后一个关键词不打标点符号）

另起页：外文摘要范例；英文摘要和关键词应该是中文摘要和关键词的翻译）

# Abstract

（标题：小二号，Times New Roman字体，居中，单倍行距，段前、段后各0.5行）

（正文：小四号，Times New Roman字体，1.5倍行距，两端对齐）

Artificial Neuron Network (ANN) simulates human being’s brain function and build the network structure. Convolutional Neural Network (CNN) have many advantage, such as ……

(2) This paper introduces the common pretreatment method of image, such as collecting image, normalization, graying and binarization. And apply these to the handwritten numeral recognition experiment and handwritten numerals writer recognition experiments.

**Keywords**: Writer recognition；Convolutional Neural Network；Handwritten character recognition

(“Keywords”：Times New Roman字体，小四号，加粗，居左）（关键词：Times New Roman字体，小四号）

（另起页：目录范例）

# 目 录

（标题：小二号，黑体，居中，两字之间空2字符，目录为电脑自动生成）

（各章标题、结论、参考文献、致谢：黑体，四号；其余：宋体，小四号，行距1.5倍）

# 绪论

（各章标题：黑体，小二号，居中，单倍行距，段前、段后各0.5行；章节序号与标题之间空一字符）

## 引言

（各节一级标题：黑体，小三号，居左，单倍行距，段前、段后各0.5行）

（正文：1.5倍行距；中文：宋体，小四号，每段首行空2个汉字；字母和阿拉伯数字：Times New Roman字体，小四号）

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%。

## 研究背景

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。

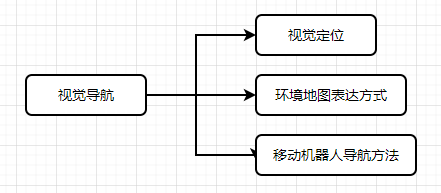


图1 视觉导航关键技术

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。

（正文中引用文献序号用小4号Times New Roman体、以上角标形式置于方括号中）

## 研究现状

笔迹获取的方式有两种，所以鉴别方式也分为离线鉴别和在线鉴别[2,3] （此处引用连续多篇文献，序号用逗号隔开）。在线鉴别是采用专用的数字板来实时收集书写信号。由文献[4-7] （此处参考文献为文中直接说明，其序号应该与正文排齐）可知，因为信号是实时采集的，所以能采集的数据不仅包括笔迹序列，而且可以采集到书写时的加速度、压力、速度等丰富有用的动态信息。

## 论文结构

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。

# 强化学习与深度学习基础知识

（各章标题：黑体，小二号，居中，单倍行距，段前、段后各0.5行；章节序号与标题之间空一字符）

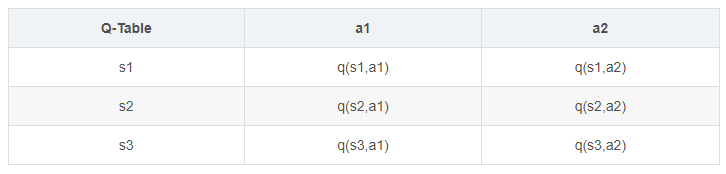
## 强化学习

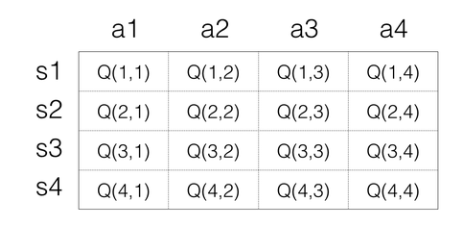
强化学习是机器学习的一个分支，与其他机器学习的方式不同的是，强化学习的训练的过程中，需要智能体与环境交互，智能体执行一次动作，会获得一个观测值，并且得到一个环境给予智能体的奖励，智能体通过获得的奖励值，通过梯度下降算法或其他优化方法，不断增大获取到的奖励值，从而学习到一种能够很好完成当前任务的策略。强化学习根据方法的不同可以分成基于值的和基于策略的两类方法。而随着深度学习的不断发展，现在许多新提出的强化学习算法已经和深度学习相结合，因此能够适用于许多连续状态空间以及连续动作空间的复杂任务。

本文主要涉及的强化学习方法有两种，A2C和DQN，前者是基于策略的强化学习方法，为后者是基于值的强化学习方法。

### DQN

Q-Learning方法是最经典的强化学习方法之一，通过构建一个Q表，来返回每个状态下智能体执行不同动作而的价值，根据智能体与环境互动获得的奖励不断的更新Q表，从而使得智能体能够根据Q表执行动作时获得更大的奖励。





Q表的形式如上图所示，当智能体处于某一状态Si时，根据Q表会返回一个执行不同动作对应的价值函数q(Si,ai)，q值越大代表执行此动作能获得的收益越大，最开始Q表是随机初始化的，所以执行的动作具有很大的随机性，执行这些动作会获得一个奖励，智能体根据这些奖励，采用时间差分法TD（融合了蒙特卡洛和动态规划）更新Q表，更新的公式如公式？



根据Q-Learning算法的表现形式我们不难看出，如果环境的状态空间是一个连续值，那么我们很难在Q表中将每一个状态都体现出来，这就极大的限制了强化学习的适用范围，只能适用于一些状态空间离散且比较小的情况。

而DQN就是在Q-Learning的基础上，融入了深度学习算法，这就使智能体实现的Q值不再是一个表格，而是一个网络，调节网络的参数，我们就可以获得一个连续的Q函数，因为融入了深度学习网络，所以可以轻易地通过网络得到任何一个连续状态输入对应的输出，从而将强化学习应用的范围拓宽。

DQN算法一共实现了两个Q网络，一个是目标值网络，负责生成训练过程中的目标，每训练N步，会将当前值网络的参数复制给目标值网络，另一个是当前值网络，负责给出当前状态对应的动作，同时也是每一步训练对应的网络。

网络训练的损失函数是



其中ω和ω-分别为当前值网络和目标值网络的参数。

### A2C

## 卷积神经网络

日常生活中，我们可以见到许多需要使用神经网络处理图片的场景，由于我们日常中看到的图片分辨率比较高，所以如果使用全连接神经网络模型构建深度学习网络是很不现实的，主要有以下三个缺点：

1. 将图像展开为向量会丢失图像的空间信息；
2. 会引入很多的参数，导致训练过程十分困难；
3. 大量的参数会使网络过拟合。

因此，卷积神经网络应运而生。卷积神经网络在 首次被提出，主要用于处理图片信息，与常见的全连接神经网络不同，卷积神经网络各层神经元以三维形式排列，因此有宽度、高度和深度三个参数，宽度与高度比较好理解，而这里的深度并不是指整个网络的层数，而是

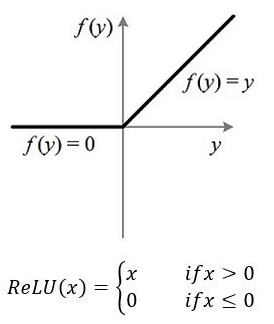
卷积神经网络主要有五部分组成：输入层，卷积层，ReLU层、池化层和全连接层，其中输入层与全连接层由于之前的全连接神经网络结构一直，这里主要介绍卷积层，ReLU层和池化层。

### 卷积层

卷积层使整个卷积神经网络的核心部分，承担了整个网络中的大部分运算量。卷积层的主要作用使作为滤波器使用，可以提取出图像的特征信息，可以被看作神经元的一个输出，有效的降低参数数量。

每个神经元只与输入数据的一部分连接，我们称这一部分为感受野，卷积核在输入数据上以步长进行滑动，并与滑动到的数据进行点积。

### ReLU层



ReLU层并没有像全连接层或者卷积层那样需要初始化以及改变的参数，而是始终执行一个同样的操作，将输入数据中小于等于0的部分变为0，而大于0的部分保持不变。在CNN中适应ReLU作为激活函数，会使神经元的一部分输出变成0，使网络更具有稀疏性，减少参数间的依存性，降低过拟合的可能。因为在卷积层后面一般会跟一个ReLU激活函数，所以我们一般会将它与卷积层看作同一层。

### 池化层

池化层的主要功能是降采样。有多种不同形式的池化函数，如最大池化（Max pooling），平均池化（average pooling），重叠池化（OverlappingPooling）和空金字塔池化（Spatial Pyramid Pooling）等。其中最大池化最为常见，

## LSTM

LSTM是一种特殊的RNN，主要用于应对长序列，降低产生梯度消失或者梯度爆炸的可能性。

图表, 瀑布图

描述已自动生成

## 本章小结

# 路径规划算法



## 输入输出层的设计

……

……

## 本章小结

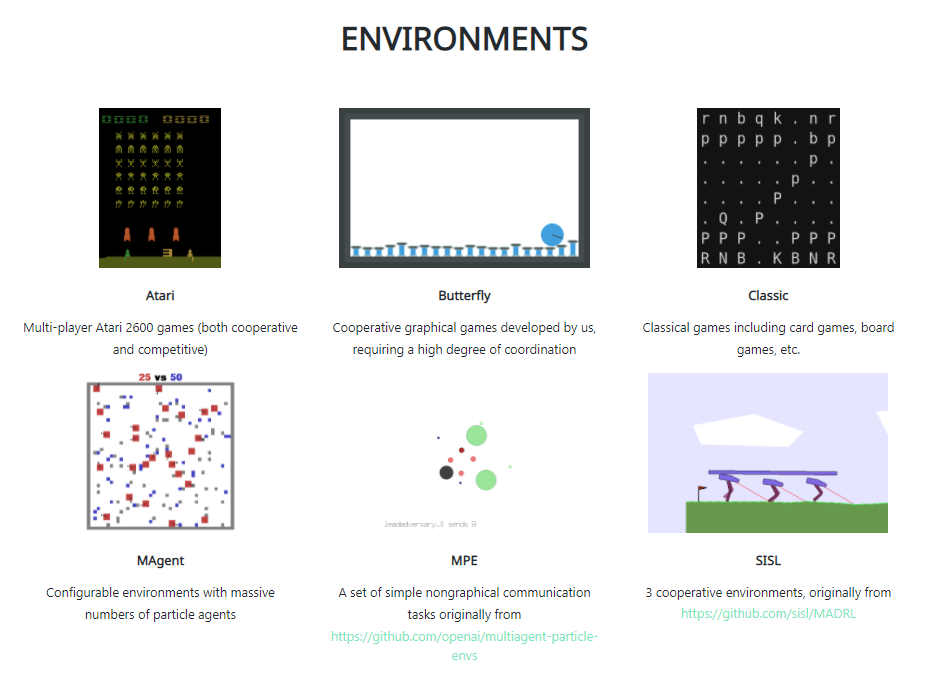
# 仿真实验

（各章标题：黑体，小二号，居中，单倍行距，段前、段后各0.5行；章节序号与标题之间空一字符）



## PettingZoo

PettingZoo是加速多agent强化学习的研究平台，类似于OpenAI的Gym库对单agent强化学习的作用，是一个在Python API下的不同多agent环境库。其中提供了很多用于训练强化学习的环境，尤其适用于多智能体的强化学习。



如图所示为PettingZoo中提供的几类主要的模型，除此之外，它还提供了一些第三方平台的训练环境，不仅如此，你还可以根据它的官方文档自己创建适合你的强化学习环境，给强化学习方向学习者提供了很多便利。

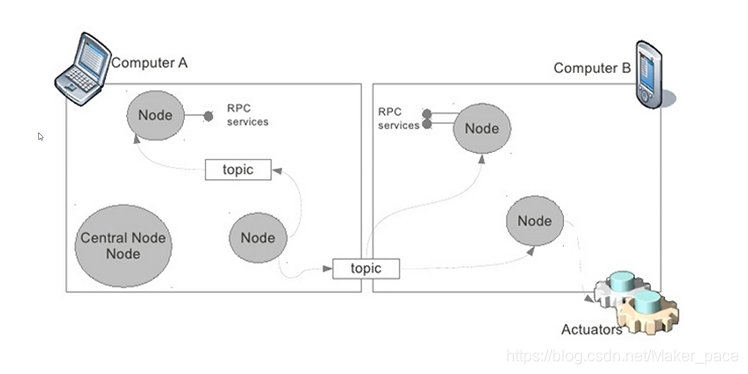
## ROS & Gazebo

### Ros简介

ROS是一个跨平台的机器人操作系统，采用松耦合分布式软件框架，原型源自[斯坦福大学](https://baike.baidu.com/item/%E6%96%AF%E5%9D%A6%E7%A6%8F%E5%A4%A7%E5%AD%A6/278716)，ROS被解释为“ROS = Plumbing + Tools + Capabilities + Ecosystem”，即ROS是通讯机制、工具软件包、机器人高层技能以及机器人生态系统的集合体。

ROS 的主要目标是为机器人研究和开发提供代码复用的支持。ROS是一个分布式的进程（也就是“节点”）框架，这些进程被封装在易于被分享和发布的程序包和功能包中。ROS也支持一种类似于代码储存库的联合系统，这个系统也可以实现工程的协作及发布。这个设计可以使一个工程的开发和实现从文件系统到用户接口完全独立决策（不受ROS限制）。同时，所有的工程都可以被ROS的基础工具整合在一起。

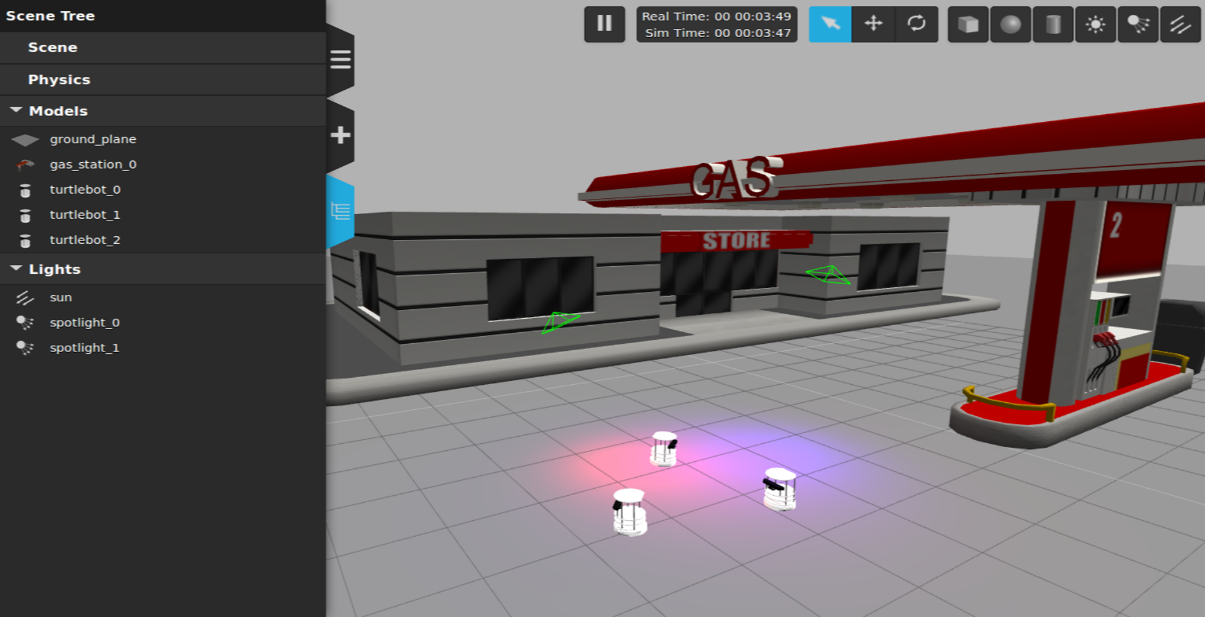
ROS主要包括节点、节点管理器、话题和服务、参数和文件系统等等。



每个node就是一个节点，每个节点就是机器人系统中完成具体功能（图像识别、传感器驱动等）一个进程，时一个能够独立运行的可执行文件，相当于windows的一个exe文件，节点之间不是孤立的，编程语言也不是固定的，不同节点可以分布式运行到不同的主机上。节点可以通过节点管理器进行统一管理。不同节点间通过话题与服务进行通信。

### Gazebo简介

Gazebo是一个机器人仿真工具，可以进行机器人的建模，运动学及动力学仿真等工作。Gazebo 提供了在复杂的室内和室外环境中准确有效地模拟机器人群体的能力。强大的物理引擎、高质量的图形以及方便的程序和图形界面触手可及。同时Gazebo 是免费的，拥有一个充满活力的社区。

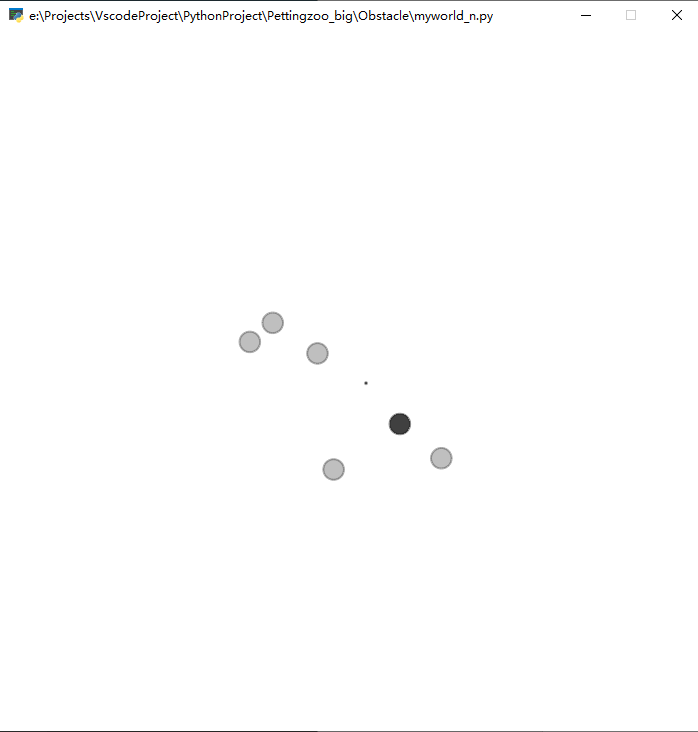


## 环境搭建

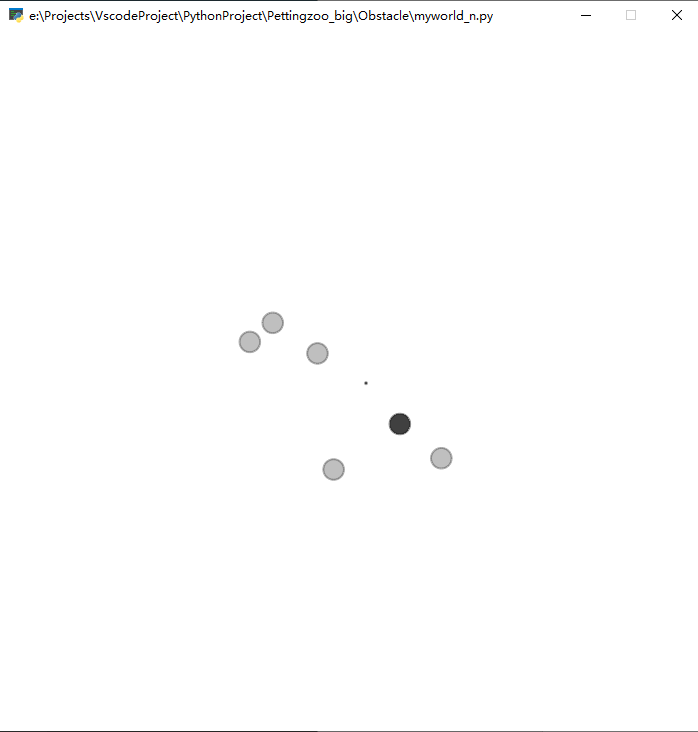
### PettingZoo环境搭建

PettingZoo中有很多已经搭建好的用于强化学习的环境，这里采用提供的Multi Particle Environments（MPE）环境，再其中的simple环境基础上进行更改。

首先是无障碍情况下，实现一个运动智能体与一个目标点的环境。其中环境的action为discrete，并且共有5个离散的动作。环境返回的观测值是一个形状为[4]数组，前两个值为智能体当前速度，后两个值为智能体与目标点横纵坐标之差。



而在有障碍情况下，其中环境的action仍为5个离散的动作。环境返回的观测值是一个形状为[4+2\*Num\_Obstacle]的数组，前四个值分别为为智能体当前速度和智能体与目标点横纵坐标之差，而Num\_Obstacle代表环境中Obstacle的个数，2\*Num\_Obstacle个状态分别为智能体与这些障碍物间横纵坐标之差



### Gazebo环境搭建

# 总结与展望

## 论文工作总结

## 工作展望

# 参考文献

1. 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.
2. S. Kobayashi, "Polychaete Like Omnidirectional Propulsion Robot in Water." Journal of the Robotics Society of Japan 33.1(2015):21-24.
3. 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
4. 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.
5. 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.
6. 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.
7. 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.
8. J. Borenstein, H. R. Everett and L. Feng, "Navigating mobile robots: systems and techniques." Wellesley: A K Peters Led, 1996.
9. 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.
10. 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.
11. C. Watkins, J. Christopher, and P. Dayan, "Q-learning." Machine Learning 8.3(1992):279-292..
12. G. A. Rummery and M. Niranjan, "On-Line Q-Learning Using Connectionist Systems." Technical Report (1994).
13. 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.
14. 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.
15. 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.
16. D. Muse and S. Wermter . "Actor-Critic Learning for Platform-Independent Robot Navigation." Cognitive Computation 1.3(2009):203-220.
17. 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.
18. 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.