

上海大学出版社

2006年上海大学博士学位论文 28



基于增强学习和 ART2 神经网络的移动机器人 路径规划研究

- 作者：樊 建
- 专业：控制理论与控制工程
- 导师：吴耿锋



G643/23

上海大学出版社

001289206

2006年上海大学博士学位论文 28



基于增强学习和 ART2 神经网络的移动机器人 路径规划研究

- 作者：樊建
- 专业：控制理论与控制工程
- 导师：吴耿锋



贵阳学院图书馆



GYXY1289206

图书在版编目(CIP)数据

2006年上海大学博士学位论文.第1辑/博士学位论文编辑部编. —上海:上海大学出版社,2009.12

ISBN 978-7-81118-511-9

I. 2... II. 博... III. 博士—学位论文—汇编—上海市—2006 IV. G643.8

中国版本图书馆CIP数据核字(2009)第162521号

2006年上海大学博士学位论文
——第1辑

上海大学出版社出版发行

(上海市上大路99号 邮政编码 200444)

(<http://www.shangdapress.com> 发行热线 66135110)

出版人:姚铁军

*

南京展望文化发展有限公司排版

上海华业装潢印刷厂印刷 各地新华书店经销

开本 890×1240 1/32 印张 264.75 字数 7 376 千

2009年12月第1版 2009年12月第1次印刷

印数:1—400

ISBN 978-7-81118-511-9/G·513 定价:1000.00元(50册)

Shanghai University Doctoral Dissertation (2006)

上海大学

The Research of Mobile Robot Path Planning based on Reinforcement Learning and ART2 Neural Network

答辩委员会名单:

主任: 曹家麟 教授, 上海电力学院 200090

委员: Candidate: Fan Jian 华东师范大学 200082

Major: Control Theory and Control Engineering 1433

Supervisor: Wu GengFeng 同济大学 教授, 同济大学 200031

刘定田 教授, 上海大学 200031

导师: 吴耿峰 教授, 上海大学 200072

Shanghai University Press

• Shanghai •

答辩委员会上海大学的评语

本论文经答辩委员会全体委员审查,确认符合上海大学博士学位论文质量要求。

答辩委员会名单:

- | | | | |
|-----|-----|-----------|--------|
| 主任: | 曹家麟 | 教授,上海电力学院 | 200090 |
| 委员: | 黄国兴 | 教授,华东师范大学 | 200062 |
| | 胡运发 | 教授,复旦大学 | 200433 |
| | 陶树平 | 教授,同济大学 | 200331 |
| | 刘宗田 | 教授,上海大学 | 200031 |
| 导师: | 吴耿锋 | 教授,上海大学 | 200072 |

论文表明,作者具有坚实宽广的基础理论和深入的系统专业知识,有较强的独立从事科学研究的能力,在增强学习方法和路径规划方面做出了的一定的创造性

学大南工

评阅人名单:

- | | | |
|------------|------------|--------|
| 曹家麟 | 教授, 上海电力学院 | 200090 |
| 黄国兴 | 教授, 华东师范大学 | 200062 |
| 钱乐秋 | 教授, 复旦大学 | 200433 |

评议人名单:

- | | | |
|------------|------------|--------|
| 邵志清 | 教授, 华东理工大学 | 200237 |
| 穆淮扣 | 教授, 上海大学 | 200072 |

答辩委员会对论文的评语

移动机器人的路径规划和避碰撞研究是高技术领域课题,涉及多个学科的交叉,作者在增强学习和 ART2 神经网络的基础上对移动机器人的路径规划进行了研究,具有理论意义和应用价值。

论文的主要创新点:

提出了一种渐进式 Soft-Max 行为选择策略,并针对连续性任务提出了一种连续任务 On-Policy 增强学习方法和基于最先策略的增强学习方法;

提出了基于 RL-ART2 神经网络的移动机器人多障碍物避碰撞系统 RLART2-CAS,采用增强学习的方法使系统可以自主获得避碰撞规则;采用 RL-ART2 神经网络,引入增强学习评估和选择机制,解决了规则存储需要占用大量内存空间的问题,并自动评估避碰效果,使系统能够通过学习获得理想的避碰行为,完成路径规划。为了解决增强学习中传统的策略查询表(lookup-table)不适合存储大量的状态值和行为值的问题,在 ART2 神经网络中实现了增强学习和神经网络的集成,结合成 RL-ART2,为评估和选择已存储在 ART2 中的分类模式提供了一种解决方案。

论文表明,作者具有坚实宽广的基础理论和深入的系统专业知识,有较强的独立从事科学研究的能力,在增强学习方法和路径规划方面做出了的一定的创造性

成果。

论文结构合理,条理较清晰,表明作者有坚实宽广的基础理论和系统深入的专门知识,科学研究能力较强。

答辩委员会表决结果

答辩委员会经无记名投票,一致同意通过樊建同学的博士学位论文答辩并建议授予工学博士学位。

答辩委员会主席: **曹家麟**

2006年3月19日

摘 要

本文在增强学习和 ART2 神经网络的基础上,研究了移动机器人路径规划问题,并重点分析了路径规划中的避碰撞问题。

本文主要的工作和创新点有以下三个方面:

(1) 首先针对增强学习中的策略问题,提出了一种渐进式 Soft-Max 行为选择策略;即在学习初始时,因为对行为掌握的经验知识不够,学习主要通过探索以得到行为的经验知识;在学习的后期则因积累了足够多的经验知识,可以更好地利用行为,同时也可以通过某种突变以达到加快或减慢学习速度的目的。其次针对连续性任务,提出了一种连续任务 On-Policy 增强学习方法 S-Learning 以及基于最先策略的增强学习方法 FPRL (Foremost-Policy Reinforcement Learning),并在仿真实验中与传统 Q-Learning 方法进行对比,以说明这两种增强学习方法的可行性。

(2) 为了解决增强学习中传统的策略查询表 (lookup-table) 不适合存储大量的状态值和行为值的问题,本文采用 ART2 神经网络存储大量的状态空间和分类模式。面对大量的分类模式,手工评估和选择分类模式是非常困难的,为此本文在 ART2 神经网络中引入增强学习评估和选择机制,结合成 RL - ART2 (Reinforcement Learning based ART2 Neural Network) 神经网络,成功地解决了如何评估和选择已存储在 ART2 中的分类模式的问题。

(3) 提出了基于 RL - ART2 神经网络的移动机器人多障碍物避碰撞系统 RLART2 - CAS (RLART2-Collision Avoidance System). 为了解决多障碍物环境下, 难于手工获取避碰撞规则的难题, 本文采用增强学习的方法使系统可以自主获得避碰撞规则; 采用 RL - ART2 神经网络, 实现了增强学习和神经网络的集成, 解决了规则存储需要占用大量内存空间的问题, 并自动评估避碰效果, 使系统能够通过学习获得理想的避碰行为, 完成路径规划。

关键词 增强学习, ART2 神经网络, 移动机器人, 避碰撞, 路径规划

Abstract

Based on the reinforcement learning and ART2 neural network, the paper does the research on the path planning problem of mobile robot and analyzes the problem of collision avoidance of the path planning.

The main work and creative points of the paper are as follows:

(1) Firstly, the paper presents an action selecting policy named gradual Soft-Max to solve the problem of selecting policy in reinforcement learning. By using the gradual Soft-Max, the agent can explore to acquire more action experience in the beginning of learning and exploit the actions in the anaphase of learning due to accumulating enough action experience. It also can make a break to accelerate or slower the learning speed. Meanwhile, this paper presents a reinforcement learning method named S-Learning for continuing task and a learning method based on foremost-policy named FPRL (Foremost-Policy Reinforcement Learning).

(2) The paper uses ART2 to store abundant classify patterns and state space in order to solve the problem that the traditional lookup-table doesn't adapt to store great amounts of state and action values. We combine the mechanism of

selection and evaluation in reinforcement learning with the ART2 as RL - ART2 (Reinforcement Learning based ART2 Neural Network) to solve the difficulties of evaluating and selecting the stored classify patterns in ART2 manually.

(3) The paper finally presents a RL - ART2 based collision avoidance system of mobile robot (RLART2 - CAS). In order to solve the difficult problem of acquiring the collision avoidance rules by hand, the paper utilizes RL - ART2 to make system acquire the collision avoidance rules automatically. We also utilizes RL - ART2 neural network to solve the problem of rules storing and evaluate the collision avoidance result automatically. The RLART2 - CAS can perfectly accomplish the path planning by acquiring the ideal collision avoidance actions through reinforcement learning.

Key words reinforcement learning, ART2 neural network, mobile robot, collision avoidance, path planning

目 录

第一章 绪论	1
1.1 课题来源	1
1.2 课题研究的目的和意义	1
1.3 移动机器人国内外研究概况	2
1.4 论文的主要研究内容	5
第二章 增强学习研究	8
2.1 增强学习简介	8
2.2 最优搜索型强化学习算法	11
2.3 经验强化型增强学习算法	15
2.4 部分感知	16
2.5 函数估计	18
2.6 符号学习和增强学习偏差	20
2.7 增强学习应用	23
2.8 小结	24
第三章 ART2 自适应谐振神经网络理论	25
3.1 自适应谐振理论 ART	25
3.2 ART2 神经网络模型	26
3.3 ART2 神经网络的基本工作原理	27
3.4 ART2 神经网络学习算法	29
3.5 小结	30

第四章 移动机器人路径规划方法研究	32
4.1 传统路径规划方法	33
4.2 智能路径规划方法	34
4.3 小结	38
第五章 连续任务增强学习方法 S-learning 及基于最先策略的 增强学习方法 FPRL	39
5.1 增强学习行为选择策略	39
5.2 渐进式 Soft-Max 行为选择策略	41
5.3 一种连续任务 On-Policy 增强学习方法 S-Learning	42
5.4 基于最先策略的增强学习方法 FPRL	46
5.5 小结	47
第六章 基于增强学习的 ART2 神经网络 RL-ART2	48
6.1 基于增强学习的 ART2 神经网络 RL-ART2	49
6.2 RL-ART2 的学习算法	50
6.3 小结	52
第七章 基于 RL-ART2 神经网络的移动机器人多障碍物避 碰撞系统 RLART2-CAS	54
7.1 引言	54
7.2 基于规则的移动机器人避碰撞方法	55
7.3 基于 RL-ART2 的避碰撞系统 RLART2-CAS	60
7.4 小结	63
第八章 仿真实验分析	64
8.1 仿真平台及环境设置	64
8.2 增强学习评价仿真实验	66

8.3	RL-ART2 神经网络的避碰撞仿真实验	70
8.4	RL-ART2 神经网络路径规划仿真实验	72
8.5	小结	76
第九章 总结与展望		77
9.1	总结	77
9.2	进一步工作	78
参考文献		79
致谢		88

1.2 课题研究的目的是和意义

在 20 世纪 70 年代后期,机器人的研究得到迅速而广泛的发展,实现机器人的全智能化是人们梦寐以求的理想。目前的研究表明,要研制出在未知或复杂或变化环境下全自主式的机器人尚不可能,但随着技术的成熟,机器人将越来越多地代替人,人也将从相对较低层次的干预中解放出来,仅进行高层次的干预。在机器人智能化的过程中,人机交互也随之智能化,因而研究工作在交互方式下的智能机器人控制是走向“机器人全智能化”的一种明智和现实的选择。

传统的机器人是固定在某一工作位置,利用多关节手臂来完成任务的机器,通常,它笨重和复杂,只适用于完成某一特定任务或某一组特定任务,不能通过升级来完成新的任务,更无法完成要求具有移动能力的任务。进入 20 世纪 90 年代,由于具有一般功能的传统工业机器人的应用趋向饱和,而许多高级生产和特种应用则需要具有各种智能的机器人参与,因而便使智能机器人获得较为迅速的发展。同时,多传感器集成与融合技术在智能机器人上以研制高水平为

第一章 绪论

1.1 课题来源

本课题来源于上海市科学技术委员会重点科技攻关项目“基于 Internet 的智能控制技术的应用”，项目编号 015115042，成果登记号为 9312005Y0105，该项目已获 2005 年度上海市科技进步三等奖。

1.2 课题研究的目的和意义

在 20 世纪 70 年代后期，机器人的研究得到迅速而广泛的发展。实现机器的全智能化是人们梦寐以求的理想。目前的研究表明，要研制出在未知或复杂的变化环境下全自主式的机器尚不可能，但是随着技术的成熟，机器将越来越多地代替人，人也将从相对较低层次的干预中解放出来，仅进行高层次的干预。在机器智能化的过程中，人机交互也随之智能化，因而研制工作在交互方式下的智能机器控制是走向“机器全智能化”的一种明智和现实的选择。

传统的机器人是固定在某一工作位置，利用多关节手臂来完成任务的机器。通常，它笨重和复杂，只适用于完成某一特定任务或某一组特定任务，不能通过升级来完成新的任务，更无法完成要求具有移动能力的任务。进入 20 世纪 90 年代，由于具有一般功能的传统工业机器人的应用趋向饱和，而许多高级生产和特种应用则需要具有各种智能的机器人参与，因而促使智能机器人获得较为迅速的发展，同时，多传感器集成与融合技术在智能机器人上以研制高水平的环

境信息传感器、信息处理技术、高适应性的移动机器人控制技术和真实环境下的规划技术为标志,展开了对机器人等更高层次的研究.移动机器人是典型的自主式智能系统,是一种具有高度智能的自动化机器,目前移动机器人都已经被应用于探测遥远的星球和深海世界.

移动机器人是高技术领域中的一个热点,是人工智能、机器人学、控制理论和电子技术等多个技术学科交叉的产物,不仅有着潜在的十分诱人的应用价值和商业价值,而且对它的研究本身也是对智能化技术的挑战.移动机器人的发展为各种智能化技术和方案的研究提供了优良的试验平台.

1.3 移动机器人国内外研究概况

自主式移动机器人是一种具有高度自规划、自组织、自适应能力,适合于在复杂的非结构化环境中工作的机器人.自主式移动机器人的目标是在没有人的干预、无需对环境做任何规定和改变的条件下,有目的地移动和完成相应任务.

移动机器人从工作环境来分,可分为室内移动机器人和室外移动机器人;按移动方式来分:轮式移动机器人、步行移动机器人、蛇形机器人、履带式移动机器人、爬行机器人等;按控制体系结构来分:功能式(水平式)结构机器人、行为式(垂直式)结构机器人和混合式机器人;按功能和用途来分:医疗机器人、军用机器人、助残机器人、清洁机器人等;按作业空间来分:陆地移动机器人、水下机器人、无人飞机和空间机器人等.本文仅针对陆地移动机器人.

1.3.1 国外研究概况

移动机器人的研究始于 20 世纪 60 年代末期.它是高技术领域中的一个热点,是人工智能、机器人学、控制理论和电子技术等多个技术学科交叉的产物,不仅有着潜在的十分诱人的应用领域和商业价值,而且对它的研究本身也是对智能技术的挑战.在这方面,美、德、