

局部相对位姿空间内轮式月球车的避障策略学习

潘海宁¹, 崔平远^{1,2}, 居鹤华²

PAN Hai-ning¹, CUI Ping-yuan^{1,2}, JU He-hua²

1.北京工业大学 电子信息与控制工程学院,北京 100124

2.哈尔滨工业大学 深空探测基础研究中心,哈尔滨 150001

1.Department of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China

2.Deep Space Exploration Research Center, Harbin Institute of Technology, Harbin 150001, China

E-mail: panhaining@126.com

PAN Hai-ning, CUI Ping-yuan, JU He-hua. Wheeled lunar rover's obstacle avoidance strategy learning in local relative pose space. *Computer Engineering and Applications*, 2009, 45(7): 29-31.

Abstract: A learning approach for the lunar rover's obstacle avoidance is presented. Firstly the dynamics of the rover BH2 is modeled. The obstacle is avoided by following the sub-goals sequence in the vision range. The relative position vectors among the rover's body, the sub-goals and the obstacle are chosen to compose the learning states in order to ensure the strategy's robustness. The rotational torque, which simplifies the learning process, is chosen as the learning action. Experiment results show that the learnt strategy is adaptive to the changing environment.

Key words: lunar rover; obstacle avoidance; relative pose; machine learning

摘要:提出了一种轮式月球车的避障学习方法。首先列出了BH2月球车的动力学方程,并将避障行为解释为沿子目标点行走的过程。然后在月球车视觉局部范围内进行避障策略学习,选择车体、子目标点和障碍之间的相对位置矢量为学习过程的状态量,使策略学习对环境变化有鲁棒性;选择车轮的转向力矩为控制输入,降低了学习复杂度。实验证明此方法对环境变化有很好的适应能力。

关键词:月球车;避障;相对位姿;机器学习

DOI: 10.3778/j.issn.1002-8331.2009.07.009 文章编号: 1002-8331(2009)07-0029-03 文献标识码: A 中图分类号: TP242.3

1 引言

机器人运动规划是考虑运动学约束,在运动空间中得到一条连接起始点和目标点的可跟踪路径并得出相应控制序列的问题。月球车是一种特殊的轮式机器人,受其视觉系统和不确定环境的限制,其运动规划是全局与局部结合的结果。这里只研究月球车局部环境中的躲避障碍问题。基于精确环境图的避障学习方法^[1-2]根据特定的环境进行避障学习而生成连接固定起始点与目标点的路径,当环境发生改变或目标点发生变化时要重复学习过程。本文根据月球车与环境特征的相对位置作为学习状态进行避障学习,克服了以环境信息作为学习状态导致的学习结果适应性差的缺点。另外,现有的大多数针对轮式机器人的研究是基于运动学方程进行的,如文献[3-4],而实际的月球车系统是动力学系统,所以采用动力学方程描述其运动过程更合理。

首先对BH2月球车的动力学方程进行简化,得到机器人在驱动力矩下的运动规律;然后通过学习的方法求解小范围内

的动力学逆问题,得到期望运动需要的控制力矩;最后给出机器人在实验环境中的学习结果。

2 BH-2 轮式月球车动力学方程

实际的大多数机器人系统的质量不能忽略,在牵引系统的作用下机器人的速度不能瞬变,在 Δt 时间内以运动学状态方程规划出的控制量变化量 Δu 必须满足:

$$\Delta u \leq M^{-1} f \cdot \Delta t$$

其中 M 是惯性矩阵, f 是牵引系统的牵引力和转动力矩。所以用动力学方程进行运动规划更符合实际的机器人控制需要。

BH2月球车原理样车的平面结构如图1所示。其采用六轮独立驱动,每个驱动轮还配有方向电机。采用Maggi方程列写其动力学方程。根据6个驱动轮的无侧向滑移约束可列写出有6个非完整约束。为降低系统阶数并减少车轮与车体内应力,规定所有车轮滚动速度方向为其所在轨迹的切线方向,即 θ_1 、 θ_2 和 θ_3 之间有关系:

基金项目:国家高技术研究发展计划(863)(the National High-Tech Research and Development Plan of China under Grant No.2006AA12Z307)。

作者简介:潘海宁(1979-),男,博士生,主要研究领域:机器人控制、机器人运动规划;崔平远(1961-),男,博士,教授,主要研究领域:飞行器飞行动力学、深空探测器自主导航与控制;居鹤华(1969-),男,博士,副教授,主要研究领域:月球车导航与控制。

收稿日期:2008-11-17 修回日期:2008-12-26

$$\frac{\tan\theta_1}{\tan\theta_3} = \frac{r+w}{r}, \frac{\tan\theta_1}{\tan\theta_2} = \frac{r}{r-w} \cdot \frac{l_1}{l_2}$$

后轮有类似规定,便可将6个约束减少为2个:

$$\dot{\varphi}^x(t, q, \dot{q}) = 0, \chi = 2, 4$$

其中 $\dot{q} = (\dot{\phi}, \dot{\theta}_2, \dot{\theta}_4, \dot{x}, \dot{y})$ 为系统的状态变量。令 c 代表 \cos , s 代表 \sin , Q 为广义力, 选取一组准坐标 $(v_*^1, v_*^2, v_*^3, v_*^4, v_*^5)$, 并使下式成立:

$$\begin{bmatrix} v_*^1 \\ v_*^2 \\ v_*^3 \\ v_*^4 \\ v_*^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -l_{r1}c\theta_4 & 0 & 0 & -s(\varphi-\theta_4) & c(\varphi-\theta_4) \\ l_{r1}c\theta_2 & 0 & 0 & -s(\varphi-\theta_2) & c(\varphi+\theta_2) \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\theta}_2 \\ \dot{\theta}_4 \\ \dot{x} \\ \dot{y} \end{bmatrix}$$

由此可得 BH2 月球车的 Maggi 方程^[5]:

$$\sum_{i=1}^5 \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}^i} - \frac{\partial T}{\partial q^i} - Q_i \right) \frac{\partial q^i}{\partial v_*^j} = 0, j = \overline{1, 3}$$

其中 T 为系统的动能, 若忽略摇臂和车体其他部分的质量, 其由车体和6个车轮的质量构成, 根据 König 定理并假设1、2、3车轮有相同的转动动能, 4、5、6有相同的转动动能, 则月球车平面动能表示为:

$$2T = M_s (\dot{x}_c^2 + \dot{y}_c^2) + J_s \dot{\varphi}^2 + 3J_{r2} \dot{\theta}_{r2}^2 + 3J_{r1} \dot{\theta}_{r1}^2 + 6J_{r2} \dot{\varphi}_{r2}^2 + 6J_{r1} \dot{\varphi}_{r1}^2 + 6\dot{\varphi} (M_{r2} l_{r2} - M_{r1} l_{r1}) (\dot{x}_c \sin\varphi - \dot{y}_c \cos\varphi)$$

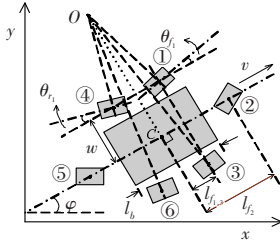


图1 BH2月球车平面结构图

3 避障学习设计

在当前视觉范围内, 设机器人初始位置为 s_0 , 则本文的避障控制描述为: 对期望的终止目标 s_{if} , 求一组控制量 u_Q , 使机器人从初始位置 s_0 运动到终端位置, 即机器人逆动力学求解问题。此类问题如用分析力学的方法求解, 不但计算量大而且不一定能得到唯一解, 不适合在线控制。文献[6]利用模糊控制方法实现了轮式机器人逆动力学求解问题, 但此方法求解逆动力学问题时规则设计比较复杂, 本文采用学习的方法求解。

3.1 离散状态定义

设车体与障碍物的距离为 d_a , 与子目标点的距离为 d_g , 车体航向角与障碍物方向的夹角为 γ_a , 航向角与子目标点方向的夹角为 γ_g 。如果根据状态 $s_i(d_g, d_a, \gamma_g, \gamma_a)$ 定义 Q 值泛化空间, 则 Q 值泛化空间为四维空间, 策略学习比较困难。考虑在沿障碍行走时并不关心与目标的距离 d_g , 而与障碍的距离 d_a 可作为回报函数的判断条件, 所以月球车在环境地图中的相对位置只需考虑 γ_g, γ_a 。此时 Q 学习过程的状态可用 $s_i(\gamma_g, \gamma_a)$ 表示, 将 γ_a 与 γ_g 划分为5个角度区间, 则每个区间代表 Q 值泛化空间的一个状态。令 L_f 与 L_r 分别是前轮与后轮的转向合力矩,

则学习过程中的动作可用 $a_i(L_f, L_r)$ 表示。将电机转向力矩在约束范围内分为8个区间作为策略学习的8个可选动作, 并规定 L_f 与 L_r 大小相等。

3.2 回报函数设计

遇到障碍的时候, 月球车根据局部视觉范围和与障碍的距离得到一系列子目标点, 其为障碍物边缘的法线方向与视觉范围的交点。为了学习控制月球车从起始点运动到子目标点的运动, 要对使月球车运动到子目标点的动作 $a_i(L_f, L_r)$ 给予相应的回报。

(1) 趋向子目标点回报。只要保证在局部视觉范围内的控制可将车体驱动至子目标点, 就能够将避障问题转换为无障碍条件下趋向目标的子过程, 保证车体能安全的驶向目标点。在图2中, 如果车体到达视觉边界时在子目标点附近则获得回报; 反之则没有回报。所以回报函数定义为:

$$r_q = 1, \text{ if } (|s_b - s_{sub}| \leq \varepsilon \text{ 且 } |s_b - s_0| = r)$$

其中 s_b 为车体位置; ε 为容许的月球车位置误差; r 为月球车视觉范围的半径。

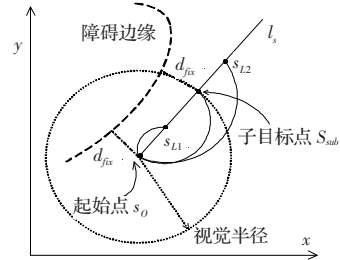


图2 回报函数定义

(2) 碰撞回报。如果月球车与障碍物的距离 d_a 小于避障容许的安全距离 d_{safe} , 则得到回报值-1, 即:

$$r_q = -1, \text{ if } (d_a < d_{safe})$$

3.3 学习过程

首先根据视觉信息和障碍信息生成当前状态的子目标点, 以恒定牵引力驱动月球车前进而只学习转向力矩的控制规律。学习过程按如下步骤进行:

- (1) 初始化 $Q(s, a)$ 值表;
- (2) 初始化当前状态 s_i ;
- (3) 随机选择动作 a_i 并执行;
- (4) 判断车体是否到达 t 时刻视觉边缘, 如已到达, 观察回报值 r_q 和新状态 s_{i+1} ;
- (5) 更新 Q 值, γ 为折扣因子并有 $0 \leq \gamma < 1$, $Q(s_{i+1}, a_{i+1}) \leftarrow Q(s_i, a_i) + [r_{i+1} + \gamma \cdot \max_{a_{max}} Q(s_{i+1}, a_{max}) - Q(s_i, a_i)]$;
- (6) 如果发生碰撞, 则在原地调整车体姿态并初始化 s_{i+1} , 然后根据 s_{i+1} 与视觉信息生成新的子目标点, 否则直接跳转到步骤(2)。

4 避障策略在线学习及结果分析

本章将进行 BH2 月球车原理样车在实验室环境中的避障学习实验, 图3即为其实验场景。月球车对环境的感知来自于车体的视觉系统, 包括: 1套用于近景精确测绘的高速激光雷达系统, 最高测速为5万点每秒; 18个红外避障测距探头, 用于紧急避障控制。计算机系统是一组 PC104 结构控制组件, 包

括控制主板和车载设备接插件。处理器为主频 700 MHz 的 Pentium III 处理器, 操作系统为 VxWorks 实时操作系统。



图3 BH2 避障策略实验

由于实验只是学习避障时的转向力矩控制规律, 所以无需模拟月表环境, 本实验中的障碍为室内会议桌椅。为了加快学习速度而在有限距离内获得更多的学习次数, 限定 BH2 激光雷达扫描范围为 1 m。车体行进速度设定为大约 10 cm/s, 并设定 6 个车轮转向力矩大小相同, 前方 3 个与后方 3 个车轮力矩方向相反。转向电机力矩控制通过改变其恒定电流实现, 电流范围控制在 0~400 mA。

学习结果如图 4 所示, 其中 x 轴代表 γ_a , y 轴代表 γ_g , z 轴代表车轮转向力矩 L 的大小, 与状态 $s_i(\gamma_g, \gamma_a)$ 相对应的动作 $a_i(L_f, L_r)$ 的 Q 可用三维空间内点的颜色深度值表示, 颜色越浅代表数值越大。可以看出在利用子目标点引导车体避障的过程中, 与障碍物的夹角对学习结果并不敏感, 转向力矩大小主要取决于与子目标的夹角 γ_g , 这种现象主要是因为用于选取

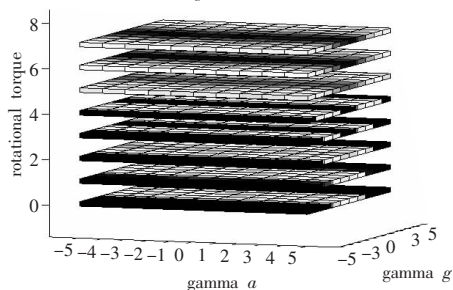


图4 状态-动作 Q 值表

(上接 6 页)

单信道误帧率的增大而可减小实际信噪比即减小传输能量。但能耗减小的幅度随 M 节点数增加而趋缓。直观地还可以看到, 在较大的误帧率范围内, 协作 SIMO 分集传输能耗都明显低于单信道非分集传输的最低能耗。

5 结束语

在无线传感器网络节点条件限制下, 采用 MIMO 分集传输数据, 不论分集各路误帧率 p_{fi} 是否相同以及采用何值, 其总能耗都将大于单路非分集采用最佳误帧率 $p_{f(1)-opt}$ 时的最低传输能耗。而如果采用协作 SIMO 分集传输, 将可能降低传输能耗。因此从能耗考虑, 在无线传感器网络中, 采用前述的 MIMO 分集传输是不合适的, 而采用协作 SIMO 分集传输显然是有益的。

应该强调, 本文采用的最低能耗指标间接来自于等效信噪比。当单信道传输时, 其反映的能耗关系是一致的。但是对于多信道分集传输, 当噪声分布特性不同时, 各信道信噪比的大小却不一定表示相互间的能耗大小。在这种情况下, 前述分析中单信道非分集传输的最低能耗, 应该是指信噪比满足最佳误帧率条件且实际能耗也是最低的信道能耗。如此说明后, 所得结论更具有普遍意义。

子目标点的 d_{fix} 数值比较大, 车体的转向动作不会与障碍物发生碰撞。较小的转向力矩在 γ_g 比较小的时候拥有较浅的颜色, 说明此时应该采用较小的转向力矩; 较大的转向力矩在 γ_g 较大的时候颜色较浅, 说明此时应该采用较大的转向力矩。

5 结论

基于局部相对位姿进行的策略学习有良好的环境适应能力, 其控制量为广义力而且规划过程只受其有界约束, 可以提高规划出的轨迹的可行性并改善学习策略的实际控制效果。但本文学习过程中转向力矩是恒定值, 当环境比较复杂时控制的精度不高, 这个问题可以依靠缩短控制周期来解决, 使离散的转向力矩序列实现更为平滑的控制效果。

参考文献:

- [1] Frazzoli E, Dahleh M A, Feron E. Robust hybrid control for autonomous vehicle motion planning[C]//Proceedings of the 39th IEEE Conference on Decision and Control, Sydney, 2000, 1: 821-826.
- [2] Fujisawa S, Kurozumi R, Yamamoto T, et al. Path planning for mobile robots using an improved reinforcement learning scheme[C]//Proceedings of the 2002 IEEE International Symposium on Intelligent Control, Vancouver, 2002, 1: 67-74.
- [3] Li Chuan-feng, Wang Chao-li. Robust stabilization of kinematic systems of uncertain nonholonomic mobile robots[C]//Proceedings of the 6th World Congress on Intelligent Control and Automation, Dalian, 2006, 1: 1142-1146.
- [4] Zhao Yi-lin, Ment B S L. Kinematics dynamics and control of wheeled mobile robots[C]//International Conference on Robotics and Automation, Nice, 1992, 1: 91-96.
- [5] Papastavridis J G. Maggi's equations of motion and the determination of constraint reactions[J]. Journal of Guidance, Dynamics and Control, 1990, 13(2): 213-220.
- [6] 居鹤华, 曹亮, 崔平远. 基于模糊逻辑的月球车逆运动学求解方法[J]. 宇航学报, 2006, 27(4): 643-648.

另外, 本文所探讨的仅是信道传输的发送能耗, 而未考虑节点电路能耗。这对于 MIMO 分集传输不影响所得结论, 但是对于协作 SIMO 分集传输, 将对 M 节点数有所限制。

参考文献:

- [1] Cui Shu-guang, Goldsmith A J, Bahai A. Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks[J]. IEEE Journal on Select Areas in Communications, 2004, 22(6): 1089-1098.
- [2] 郭金淮, 于宏毅. 一种用于无线传感器网的协作接收方案[J]. 电路与系统学报, 2007, 12(2): 43-46.
- [3] Min R, Chandrakasan A. A framework for energy-scalable communication in high-density wireless networks[C]//Proceedings of the Int Symp, Low Power Electronics Design, 2002: 36-41.
- [4] Xu Jian-cheng, Shi Hao-shan, Wang Jie-tai. Analysis of frame length and frame error rate for the lowest energy dissipation in wireless sensor networks[C]//Proceedings of the IEEE 4th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2008), 2008.
- [5] 徐建城, 史浩山. 无线传感器网络中最低能耗误帧率及判断准则[J]. 西北工业大学学报, 2008, 26(2): 137-141.
- [6] 徐建城, 史浩山, 王结太. 无线传感器网络中 ARQ 协议能量有效性分析[J]. 计算机工程与应用, 2008, 44(9): 13-23.