

基于模糊 K 均值聚类 and Sarsa (λ) 算法的自适应爬壁机器人路径规划

李静静

(河北工程大学 科信学院, 河北 邯郸 056038)

摘要: 针对现有爬壁机器人规划算法难以实现在线自适应高效规划的问题, 设计了一种基于模糊 K-Means 算法和经典 Sarsa (λ) 算法自适应爬壁机器人规划算法; 首先, 对爬壁机器人的动力学模型进行了建模和分析, 然后, 对爬壁机器人规划中的状态进行自适应聚集从而实现值函数的近似, 设计了 K 值可变的改进模糊 K 均值聚类算法对状态进行自适应地在线聚类, 将聚类中心对应的值函数作为整个聚类所有数据对象的值函数的近似值, 最后, 对基于模糊 K 均值聚类算法和 Sarsa (λ) 算法的爬壁机器人在线规划算法进行了定义和描述, 在 MATLAB 环境下对简单障碍物场景和复杂障碍物场景分别仿真实验, 实验结果表明文中方法能有效地进行路径规划, 随着情节数的增加, 规划结果逐渐收敛到最优值, 同时在环境变化时, 收敛效果不受影响, 具有较好的稳定性, 是一种高效地实现爬壁机器人在线规划的方法。

关键词: 路径规划; 模糊 K 均值; 状态聚类; 值函数

Adaptive Path Planning of Wall-Climbing Robot Based on MIP and Improved Fuzzy K-Means Algorithm and Sarsa (λ)

Li Jingjing

(College of Kexin, Hebei University of Engineering, Handan 056038, China)

Abstract: Aiming at the given wall-climbing robot path planning algorithm difficult to plan effectively online, an adaptive path planning algorithm based on K-Means algorithm and classic Sarsa (λ) algorithm are designed. Firstly, the dynamical model for wall-climbing robot is designed. Then the states of the planning space is clustered adaptively to realize the value function approximating, the improved fuzzy K-Means with the variable K is designed to cluster on-line, using value of the cluster center as the approximate value for all the sample of the whole cluster. Finally, the algorithm based on fuzzy K-Means and classic Sarsa (λ) algorithm for wall-climbing robot path planning is defined and described. The simulation experiment with simple barrier and complicate barrier is operated in the MATLAB, the result shows the method in this paper can realize the path planning, and with the increase of the episode, the planning result is converged to the optimal value, and also the convergence effect is not subject to the change of the environment, so it has strong feasibility. It is a on-line planning method for wall-climbing robot path planning.

Keywords: path planning; fuzzy K-Means; state clustering; value function

0 引言

爬壁机器人^[1]是一种模拟生物界壁虎运动的轻巧结构, 能代替人工完成很多复杂危险动作^[2-3], 可以用于大量民用和军事领域, 如用于对高层楼房的外墙和玻璃窗等进行清洗, 对天然气存储罐检测等, 对航天飞机的外舱进行维修^[4]。

目前已经有的爬壁机器人行动规划方法主要有基于智能化算法的规划方法、基于人工势场和神经网络的规划方法、基于整数规划的规划方法和基于强化学习的规划方法。

文献 [5] 研究了一种基于遗传算法的双足爬壁机器人越障路径规划算法。文献 [6] 研究了一种基于人工神经网络结构能量函数的爬壁机器人三维空间全局路径规划方法, 并根据爬壁机器人吸附的物体形状确定不同的模拟退火初始温度以规划出最短路径。文献 [7] 提出了一种基于 City-Climber 的路径方法, 采用 MILP 混合整数线性规划方法对控制输入进行描述。文献 [8] 设计了一种基于 KNN 算法最近邻分类法和时

间差分方法结合的爬壁机器人路径规划方法。

上述工作中前 3 种方法不能处理在线不确定性问题, 而后一种方法中需要不断更新和保存 k 个最近状态, 不具备自适应性。为此, 文中设计了一种基于 K-Means 无监督聚类和 Sarsa (λ) 算法的爬壁机器人路径规划算法, 并通过实验证明了文中方法的优越性。

1 爬壁机器人运动学建模

文中的爬壁机器人采用双足爬壁机器人, 即由 2 个吸盘和 1 个多自由度移动机构组成, 移动机构驱动吸盘在移动壁面进行交替移动和吸附, 实现爬壁机器人跨越障碍物的运动, 以及在墙壁、天花板上爬行和在相邻表面交界处进行过渡。

图 1 所示为爬壁机器人的结构图, 主要包含: 吸盘足 r 和 f , 其坐标分别为 $\{x_r, y_r, z_r\}$ 和 $\{x_f, y_f, z_f\}$, 对于关节 1、2、3、4, 其坐标分别为 $\{x_1, y_1, z_1\}$ 、 $\{x_2, y_2, z_2\}$ 、 $\{x_3, y_3, z_3\}$ 和 $\{x_4, y_4, z_4\}$, 假设杆件 $i-1$ 到杆件 i 之间的齐次变换矩阵为 ${}^{i-1}T_i$, 采用 θ_i 表示杆件 i 的角度, 采用 D_1 和 F_1 分别表示 $\cos(\theta_1)$ 和 $\sin(\theta_1)$, 采用 D_{23} 表示 $\cos(\theta_2 + \theta_3)$ 、 F_{23} 表示 $\sin(\theta_2 + \theta_3)$ 、采用 D_{234} 表示 $\cos(\theta_2 + \theta_3 + \theta_4)$ 、 F_{234} 表示 $\sin(\theta_2 + \theta_3 + \theta_4)$, 则吸盘足 r 到吸盘足 f 之间变换的齐次变换矩阵⁰ T_4 可以表示为:

收稿日期:2014-05-11; 修回日期:2013-06-10。

作者简介:李静静(1981-),女,河北邯郸人,硕士,讲师,主要从事电子技术和数据挖掘方向的研究。

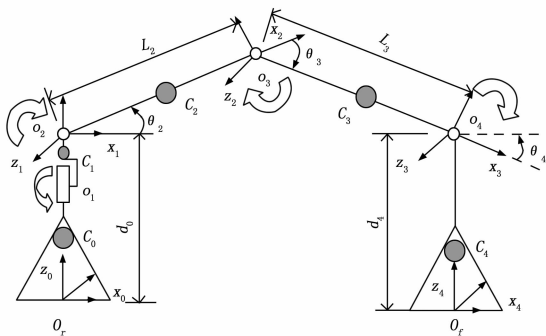


图 1 爬壁机器人结构图

$${}^0T_4 = {}^0T_1 T_2^2 T_3^3 T_4 = \begin{bmatrix} -D_1 D_{234} & -F_1 & -D_1 F_{234} & e_2 D_1 D_2 + e_3 D_1 D_{23} + c_4 D_1 F_{234} \\ F_1 D_{234} & -D_1 & -F_1 F_{234} & e_2 F_1 D_2 + e_3 F_1 D_{23} + c_4 F_1 F_{234} \\ F_{234} & 0 & D_{234} & c_0 + e_2 F_2 + e_3 F_{23} - c_4 F_{234} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

式 (1) 中, c_4 表示杆件 4 距离吸盘底盘面的距离, e_2 和 e_3 分别表示杆件 2 和杆件 3 的长度。

机器人吸盘 f 作相对吸盘 r 的运动, 当机器人在垂直壁上运动时, 其受力情况如图 2 所示。

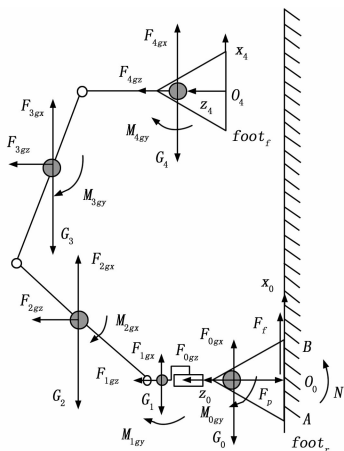


图 2 机器人壁面双足受力结构图

其中, 机器人受力满足下列方程:

$$\begin{cases} -M_g + F_f + \sum_{i=0}^4 F_{igx} = 0 \\ -F_p + F_N + \sum_{i=0}^4 F_{igz} = 0 \\ \sum_{i=0}^4 z_{ic} m_i g - F_p \frac{L_s}{2} + M_N = \\ \sum_{i=0}^4 (M_{igy} + x_{ic} F_{igz} - z_{ic} F_{igx}) = 0 \end{cases} \quad (2)$$

式 (2) 中, F_{igx} 和 F_{igy} 分别为杆件 i 在 x 和 y 方向的惯性力, M_{igy} 为杆件 i 的惯性力矩在 y 轴上分量, M_g 和 m_i 分别为机器人的质量和杆件 i 的质量, F_p 和 F_N 分别为吸盘吸附压力和支反力, z_{ic} 和 x_{ic} 分别为杆件 i 质心位置的 z 轴分量和 x 轴分量。

2 基于模糊 K-MEANS 算法的状态自适应聚集

2.1 自适应状态聚集

模糊 K-MEANS 算法是一种经典的动态聚类算法, 通过

不断最小化误差平方和, 并据此更新每个数据样本点到聚类中心的隶属度, 从而将隶属度最高的聚类作为数据样本点所属分类, 从而实现样本的分类, 但其存在着: 聚类数 K 需要在初始时刻确定, 且聚类数在确定后不能自适应的改变, 由于在爬壁机器人路径规划中, 当对状态进行连续表示时, 机器人在规划的过程中不断获取新的状态, 因此, 基于模糊 K-MEANS 算法的自适应状态聚集算法原理为:

首先设定一个初始的聚类数 K , 最大聚类数 K_{max} , 距离阈值为 θ_d , 当机器人在线获取一个新的状态时, 计算该新状态与已有的所有聚类中心的距离, 如果其与所有聚类中心的距离的最小值距离阈值为 θ_d 大于阈值 θ_d , 且聚类中心数 $K \leq K_{max}$ 时, 增加一个新的聚类 S_{K+1} 和聚类中心 c_{K+1} , 当状态聚集为类后, 一个聚类中的所有状态均采样聚类中心来表示, 即在后面的算法中, 采用聚类中心的 Q 值函数来表示聚类中所有状态的 Q 值。

2.2 自适应模糊 K-MEANS 聚集算法描述

算法 1: 基于改进模糊 K-MEANS 算法的在线状态自适应聚集算法描述如下。

1) 初始化距离阈值为 θ_d 、总距离阈值 θ 聚类数 K , 最大聚类数 K_{max} , 聚类集合 S_1, S_2, \dots, S_K 为空集, 聚类数据中心 c_1, c_2, \dots, c_K 为空, 当前迭代次数 $t = 1$;

2) 在线获取的数据样本 x_i , 计算每个数据样本点 x_i 到各个数据聚类 S_j 的隶属度 w_{ij} :

$$w_{ij} = \begin{cases} \left(\frac{\sum_{q=1}^K \left(\frac{\|x_i - c_j\|}{\|x_i - c_q\|} \right)^{2/(m-1)}}{\sum_{q=1}^K \left(\frac{\|x_i - c_j\|}{\|x_i - c_q\|} \right)^{2/(m-1)}} \right)^{-1} \\ 1 \|x_i - c_q\| = 0 (i = q) \\ 0 \|x_i - c_q\| = 0 (i \neq q) \\ 1 \leq i \leq n, 1 \leq j \leq K \end{cases} \quad (3)$$

其中: $\|x_{ik} - c_{jk}\|$ 表示数据对象 x_i 到数据中心 c_j 之间的欧几里得距离, w_{ij} 满足:

$$\sum_{j=1}^K w_{ij} = 1 (1 \leq i \leq n) \quad (4)$$

3) 更新各个数据聚类 S_j 的数据中心 c_1, c_2, \dots, c_K :

$$c_j = \frac{\sum_{i=1}^n w_{ij}^m x_i}{\sum_{i=1}^n w_{ij}^m} \quad 1 \leq j \leq K \quad (5)$$

4) 计算数据样本点 x_i 到所有聚类中心的距离的最小值 $J(W, C)$:

$$J(W, C) = \min w_{ij}^m \sum_{k=1}^D \sum_{j=1}^K \|x_{ik} - c_{jk}\| \quad (6)$$

5) 判断 $J(W, C)$ 是否大于距离阈值 θ_d , 如果大于阈值, 则判断 $K \leq K_{max}$ 是否成立:

如果成立, 则转入步骤 6);

否则, 转入步骤 7);

6) 以 x_i 作为数据中心重新建立一个新的聚类, $c_{K+1} = x_i$, 聚类 $S_K = \{x_i\}$, 并将聚类数 $K = K + 1$;

7) 查找 $J(W, C)$ 对应的聚类, 即使:

$$j = \underset{j}{\operatorname{argmin}} J(W, C) \quad (7)$$

根据找到的聚类 S_j , 调整聚类中心 c_j , 使得新的聚类中心为:

$$c_j = c_j + \eta [x_i - c_j] \quad (8)$$

其中: η 为学习率。

8) 判断是否达到最大迭代次数, 如果达到, 则算法结束; 否则, 判断所有数据点与所有数据中心的总距离是否小于 θ :

$$J_{sum}(W, C) = \sum_{i=1}^n \omega_{ij}^m \sum_{k=1}^D \sum_{j=1}^K \|x_{ik} - c_{jk}\| \quad (9)$$

其中: n 为样本总数, 如果式 (9) 小于 θ 则算法结束, 否则转入 2) 继续运行。

3 基于状态聚集和 Sarsa (λ) 算法的爬壁机器人路径规划

3.1 Sarsa (λ) 算法

Sarsa 算法是 1994 年时由 Rummery 和 Niranjan 提出的一种在策略时间差分算法, Agent 与环境的交互可以看成是一个采样四元组的过程, 即:

$$\langle \text{state1}, \text{action1}, \text{reward}, \text{state2}, \text{action2} \rangle \quad (10)$$

其中: state1 表示状态, action1 表示动作, reward 表示回报, state2 表示下一个状态, action2 表示在下一个状态时采取的动作。

3.2 基于状态聚集和 Sarsa (λ) 算法的爬壁机器人路径规划

经典的 Sarsa 算法中只计算一步更新的 Q 值, 因此在算法收敛前, 需要较多的迭代次数, 因此具有较低的收敛效率。为此, 采用带资格迹的 Sarsa 算法即 Sarsa (λ) 算法, 并结合模糊 K -MEANS 的状态聚类进行函数逼近。

算法 2: 文中基于模糊 KMEANS 算法和 Sarsa (λ) 的爬壁机器人路径规划算法描述如下。

1) 初始化状态空间 X 为所有爬壁机器人的位置, 状态动作空间 $X \times U$, 对于状态动作空间中的所有状态动作对 $\forall (x, u) \in X \times U, Q_0(x, u) = 0$, 探索因子 ϵ , 折扣因子 γ , 学习率 α , 最大迭代次数 T , 最大情节数 ET , 每个情节中的最大步数 ETS ;

2) 指定算法 1 中所需参数, 确定出初始的聚类、聚类中心、聚类个数和最大聚类个数;

3) 采用爬壁机器人的初始位置来初始化状态, 并把初始状态设置为当前状态 x_i , 令迭代次数 $t = 1$;

4) 当前情节数 $et = 1$;

5) 设置当前步数 $s = 0$;

6) 根据算法确定当前状态是否需要增加新的聚类, 并确定所属的聚类 S_j , 并对其数据中心作下列更新:

$$c_j = c_j + \eta[x_i - c_j] \quad (11)$$

7) 根据当前状态所属聚类的聚类中心的 Q 值, 确定其最优动作 a :

$$a = \arg \max_{a_i} Q(c_j, a_i) \quad (12)$$

根据 ϵ 贪心策略来确定当前状态下选取的最优动作, 即以 $1 - \epsilon$ 的概率选择最优动作 a , 以 ϵ 的概率选择其他动作。

8) 爬壁机器人根据由 ϵ 贪心策略确定的动作 u 来对当前位置进行更新, 得到后续状态 x' 和立即奖赏 r ;

9) 采用 ϵ 贪心策略来获得在后续状态 x' 时选择的动作 u' , 对爬壁机器人在当前状态下的 Q 值进行更新:

$$Q(x, u) = Q(x, u) + \alpha(r + \gamma Q(x', u') - Q(x, u)) \quad (12)$$

10) 将后续状态 x 赋给当前状态 $x \leftarrow x'$, 将下一个动作赋给当前动作 $u \leftarrow u'$;

11) $s = s + 1$, 判断是否达到情节所允许的最大步数:

如果达到最大步数, 则当前情节数 $et = et + 1$, 并转入 5) 继续执行;

否则, 算法迭代次数 $t = t + 1$, 判断 t 的值:

如果小于 T , 则转入 4) 继续执行; 算法结束;

否则算法结束, 根据最终学习到的 Q 值来求取最优路径。

4 仿真实验

采用 Matlab 仿真工具^[13]对文中方法进行验证, 设定 2 个障碍物较少的规划环境, 如图 3~4 所示。

采用文中方法对图 3 和图 4 所示的简单障碍物环境和复杂

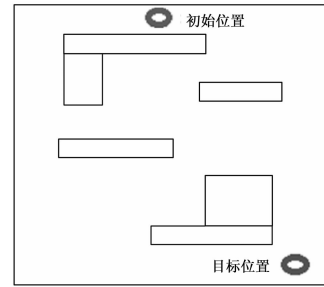


图 3 简单障碍物环境

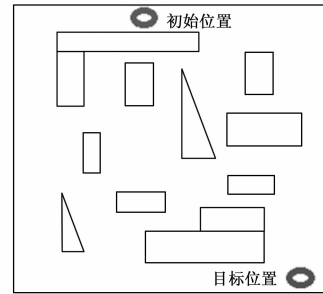


图 4 复杂障碍物环境

障碍物环境的爬壁机器人路径规划实例进行仿真, 参数设置如下: 算法 1 中距离阈值为 $\theta_d = 1.2$ 、总距离阈值 $\theta = 21$, 聚类数 $K = 1$, 最大聚类数 $K_{max} = 15$, 算法 2 中探索因子 $\epsilon = 0.2$, 折扣因子 $\gamma = 0.8$, 学习率 $\alpha = 0.5$, 最大迭代次数 $T = 30$, 最大情节数 $ET = 100$, 每个情节中的最大步数 $ETS = 1000$ 。

将文中方法得到的规划仿真结果收敛性与经典的 Sarsa (λ) 算法以及文献 [8] 所示的 KNN-Sarsa (λ) 方法进行比较, 得到的收敛情况分别如图 5 和图 6 所示。

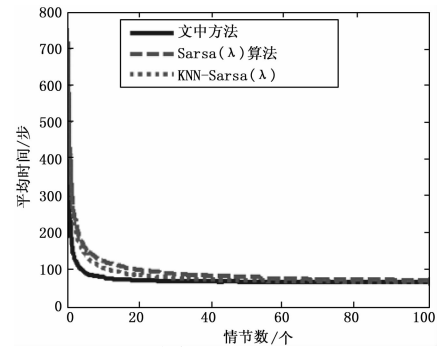


图 5 简单障碍物环境下收敛曲线

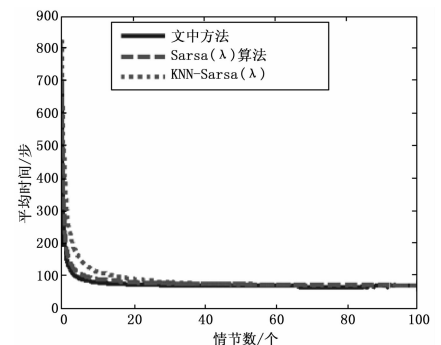


图 6 复杂障碍物环境下收敛曲线

图 5 和图 6 展示了简单障碍物环境和复杂障碍物环境下的爬壁机器人在不同方法下达到目标所需要的规划步数随情节变 (下转第 2885 页)

表 2 待识别样本模糊识别结果

样本序号	正常状态	轻微退化状态	严重退化状态	失效状态	识别状态
1	0.956 5	0.040 9	0.002 0	0.000 6	正常状态
2	0.081 2	0.879 0	0.027 6	0.012 2	轻微退化状态
3	0.059 6	0.935 1	0.003 2	0.002 1	轻微退化状态
4	0.043 2	0.067 5	0.735 1	0.154 2	严重退化状态
5	0.016 8	0.023 1	0.086 4	0.873 7	失效状态

有一定的促进意义。

参考文献:

[1] 钟秉林, 黄仁. 机械故障诊断学 [M]. 北京: 机械工业出版社, 2007.
 [2] Cheung M, Li W. Probabilistic Fatigue and Fracture Analysis of Steel Bridges [J]. Journal of Structural Safety, 2003, 23: 245-262.
 [3] 李兵, 张培林, 任国全, 等. 形态学广义分形维数在发动机故障诊断中的应用 [J]. 振动与冲击, 2011, 30 (10): 209-211.
 [4] Li B, Zhang P L, Wang Z J, etc. Morphological Covering Based Generalized Dimension for Gear Fault Diagnosis [J]. Nonlinear Dynamics, 2012, 67 (4): 2561-2571.

[5] Sharif M, Bilal M, Khan S, etc. Adaptive Filter and Morphological Operators Using Binary PSO [J]. Lecture Note in Computer Science, 2010, 6377 (1): 525-532.
 [6] 刘德馨, 李晓理, 周翔, 等. 模糊 C 均值聚类算法在高炉料面分类中的应用 [J]. 北京科技大学学报, 2012, 34 (6): 683-690.
 [7] Lee J, Qiu H, Yu G. Rexnord Technical Services. "Bearing Data Set" [EB/OL]. IMS, University of Cincinnati, NASA Ames Prognostics Data Repository. <http://ti.arc.nasa.gov/project/prognostic-data-repository>, NASA Ames, Moffett Field, CA, (accessed 09.05.09).
 [8] 潘玉娜, 陈进, 李兴林. 基于模糊 C-均值的设备性能退化评估方法 [J]. 上海交通大学学报, 2009, 43 (11): 1795-1797.

(上接第 2881 页)

化的情况, 从图 5 可以看出, 文中方法在第 15 个情节处就已经趋于收敛, 最终达到目标所需要的时间步为 99, 而经典的 Sarsa (λ) 算法和 KNN-Sarsa (λ) 方法则分别等到情节数为 40 和 24 时才收敛, 且所需要的时间步分别为 103 和 101。

在图 6 中, 虽然障碍物增加且布局更为复杂, 但文中方法仍然能在第 15 个情节处趋于收敛, 收敛对应的时间步为 98, 经典的 Sarsa (λ) 算法收敛于第 19 个情节, 对应的时间步为 98, KNN-Sarsa (λ) 方法则收敛于第 25 个情节, 收敛对应的时间步为 98。

综合上述, 文中方法不仅收敛效果好, 同时具有很强的稳定性, 在环境变化, 其收敛的效果不受影响, 因此, 较经典的 Sarsa (λ) 算法和 KNN-Sarsa (λ) 方法更优。

5 总结

为了实现爬壁机器人的自适应在线规划, 提出了一种基于模糊 K-Means 算法和 Sarsa (λ) 的机器人在线规划算法, 首先对爬壁机器人的动力学模型进行了建模和描述, 然后, 设计了改变的模糊 K 均值聚类算法的实现状态的自适应在线聚类, 将聚类中心对应的值函数作为整个聚类所有数据对象的值函数近似值, 最后, 对基于模糊 K 均值聚类算法和 Sarsa (λ) 的爬壁机器人在线规划算法进行了

详尽地定义和描述。实验表明文中方法不仅具有收敛效果好, 规划路径短的有点, 而且在环境变化的情况下, 具有很好的稳定性。

参考文献:

[1] 衣正尧, 弓永军, 王祖温, 等. 新型船舶壁面除锈爬壁机器人动力学建模与分析 [J]. 机械工程学报, 2010, 46 (15): 24-29.
 [2] Viet, Hoang Huu. Simulation-based evaluations of reinforcement learning algorithms for autonomous mobile robot path planning [J]. Springer Science business media B. V. 2012: 467-476.
 [3] 李志海. 轮足混合驱动爬壁机器人及其关键技术的研究 [D]. 哈尔滨: 哈尔滨工业大学, 2010.
 [4] Wang M, Chen X Q, Nayyerloo M, et al. A novel wall climbing robot based on Bernoulli effect [J]. Mechatronic and Embedded systems and applications, 2008, 210-215.
 [5] 李志海, 付宜利. 基于遗传算法的仿生双足爬壁机器人越障运动规划 [J]. 机器人, 2012, 6 (34): 751-757.
 [6] 禹建丽, 张晓梅, 程思雅, 等. 爬壁机器人路径规划研究 [J]. 系统仿真学报, 2009, 15 (21): 4748-4751.
 [7] 岳荣刚, 王少萍. 基于混合整数线性规划的爬壁机器人路径规划 [J]. 北京航空航天大学学报, 2013, 6 (39): 792-797.
 [8] 王金秋, 孙晓松, 秦华. 基于强化学习的爬壁机器人路径规划方法 [J]. 计算机测量与控制, 2013, 21 (11): 3093-3095.