

未知环境下无人车自主导航探索与地图构建*

满恂钰¹ 刘元盛² 齐含³ 严超¹ 杨茹锦¹

(1.北京联合大学,北京市信息服务工程重点实验室,北京 100101;2.北京联合大学,机器人学院,北京 100101;3.北京联合大学,智慧城市学院,北京 100101)

【摘要】针对自主导航探索算法易陷入局部区域的问题,提出了融合采样与深度强化学习的探索算法。首先,局部采用长短期记忆(LSTM)网络获得无人车历史位姿信息进而避免重复走向已探索区域;其次,利用深度强化学习输出策略最优的动作并设计奖励函数以激励无人车充分探索未知区域;最后,考虑无人车水平移动因素,通过解非对称旅行商问题(ATSP)生成一条符合其当前姿态的全局探索路径。2 000 s 矿道仿真环境中,所提出的算法相较于无人车自主探索(TARE)算法,探索面积增加 346.3 m²,总行驶距离减少 209.4 m;在真实场景试验中,该探索算法用时 1 014 s 完成面积为 3 444.3 m² 的地下车库探索返回起点,并完成环境地图构建。

关键词:自主导航探索 长短期记忆网络 深度强化学习 地图构建

中图分类号:U469.6;TP242.6 **文献标识码:**A **DOI:** 10.19620/j.cnki.1000-3703.20230277

Autonomous Navigation Exploration and Map Construction for Unmanned Vehicles in Unknown Environments

Man Xunyu¹, Liu Yuansheng², Qi Han³, Yan Chao¹, Yang Rujin¹

(1. Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101; 2. College of Robotics, Beijing Union University, Beijing 100101; 3. Smart City College, Beijing Union University, Beijing 100101)

【Abstract】For the problem that the autonomous navigation exploration algorithm is easy to fall into the local area, this paper proposed an exploration algorithm combining sampling and deep reinforcement learning. First, the Long-Short-Term Memory (LSTM) network was used locally to obtain the historical pose information of the unmanned vehicle to avoid repeated exploration of the explored area; secondly, the optimal action of the deep reinforcement learning strategy was used to output using deep reinforcement learning and the reward function was designed to encourage the unmanned vehicle to fully explore the unknown area; Finally, the horizontal movement factor of the unmanned vehicle was considered to generate a global exploration path conforming to its current attitude by solving the Asymmetric Travel Salesman Problem (ATSP). In the 2 000 s mine tunnel simulation environment, compared with the Technologies for Autonomous Robot Exploration (TARE) algorithm, the proposed algorithm increased the exploration area by 346.3 m² and reduced the total driving distance by 209.4 m; in the real scene test, the exploration algorithm completed the exploration of the underground garage with an area of 3 444.3 m² and returned to the starting point in 1 014 s and built the environment map.

Key words: Autonomous navigation exploration, Long-Short-Term Memory (LSTM) network, Deep reinforcement learning, Map building

【引用格式】满恂钰,刘元盛,齐含,等.未知环境下无人车自主导航探索与地图构建[J].汽车技术,2023(11):34-40.

MAN X Y, LIU Y S, QI H, et al. Autonomous Navigation Exploration and Map Construction for Unmanned Vehicles in Unknown Environments[J]. Automobile Technology, 2023(11): 34-40.

1 前言

具备自主移动功能的无人车^[1]可用于灾害救援^[2]、

地下勘探^[3-4]、行星探索^[5]及建筑作业^[6]等场景探索未知环境。自主导航探索与地图构建的任务涉及环境感知、同步定位与地图构建(Simultaneous Localization

*基金项目:国家重点研发计划项目(2021YFC3001300);国家自然科学基金重点项目合作项目(61931012);

北京联合大学高水平孵化项目和新进博士孵化项目(ZK10202208);

北京市属高等学校高水平科研创新团队建设支持计划项目(BPHR20220121)。

通讯作者:刘元盛(1973-),男,硕士,教授,研究方向为无人驾驶感知定位技术,yuansheng@buu.edu.cn。

And Mapping, SLAM)和路径规划等多个领域的结合应用^[7]。尽管自主导航探索技术已经取得长足发展,但面对长时段、大规模以及复杂环境等任务场景时依然面临导航定位丢失、探索区域不全、计算效率过低等挑战。常见自主导航探索问题的解决方法可分为基于边界的探索方法、基于采样的探索方法以及基于学习的探索方法。边界法的思想是寻找已知空间和未知空间之间的所有边界,随机选择其中一个边界作为下一个需要探索的目标区域。Yamauchi等^[8]开创了边界法,使无人车不断朝距离当前位置最近的边界移动,随着无人车的移动,边界不断更新。基于边界的方法采用贪婪策略引导无人车探索,在大规模复杂环境中会导致探索效率低下。

基于采样的方法通过随机采样一些可视点,设计增益函数选择能探索到更多未知空间的视点组并执行导航探索。常见算法包括快速扩展随机树(Rapidly-exploring Random Tree, RRT)算法^[9]、快速扩展随机图(Rapidly-exploring Random Graph, RRG)算法^[10]、下一步最佳视图规划(Next Best View Planning, NBVP)算法^[11]、分层和图的探索规划器(Graph-Based Path planning, GBP)^[12]、运动图元的探索规划器(Motion primitives-Based Path planning, MBP)^[13]和无人自主探索(Technologies for Autonomous Robot Exploration, TARE)算法^[14]等。然而,基于采样的算法本质属于贪心算法,易导致无人车陷入局部区域而无法完成完整环境探索。

基于学习的方法可直接从环境中学习而无需人工干预。Wang等^[15]提出改进TARE局部探索算法,引入注意力机制学习局部地图区域间多尺度依赖关系,但该方法

未考虑长期增益,易导致无人车重复进入已探索空间。

本文对自主导航探索任务进行重新思考,设计考虑历史信息的深度强化学习框架,选择最优目标点以解决无人车探索过程中陷入局部区域的问题,将无人车水平移动因素作为约束条件以获得一条最短全局探索路径。最终,将本文方法在仿真和真实环境中进行验证。

2 融合采样与深度强化学习的探索算法

本文基于激光雷达和惯性导航单元融合的优化正态分布变换(OPTimization Normal Distributions Transform, OP-NDT)^[16]算法提供实时点云和状态估计,提出基于采样与深度强化学习融合的探索(Exploration based on Sampling and Deep Reinforcement Learning, ESDRL)算法完成自主导航探索任务,设计含历史信息的视点选择软演员评论家网络(Viewpoint Selection network with Historical information Soft Actor-Critic, VSH-SAC)解决传统方法局部探索时无法存储和利用历史导航信息问题,引入长短期记忆(Long Short-Term Memory, LSTM)网络优化下一步视点选择,并采用软演员-评论家(Soft Actor-Critic, SAC)算法^[17]对选择的目标点进行评判以输出最优策略。将当前车辆所在局部空间之外的历史未探索区域记录为多个全局未探索空间,利用采样算法得到全局未探索空间质心,由A*算法生成全局子空间距离矩阵,并结合无人车当前位姿计算水平移动成本,通过解非对称旅行商问题(Asymmetric Traveling Salesman Problem, ATSP)得到一条串联所有未探索空间的最短路径。图1所示为ESDRL算法的整体框架。

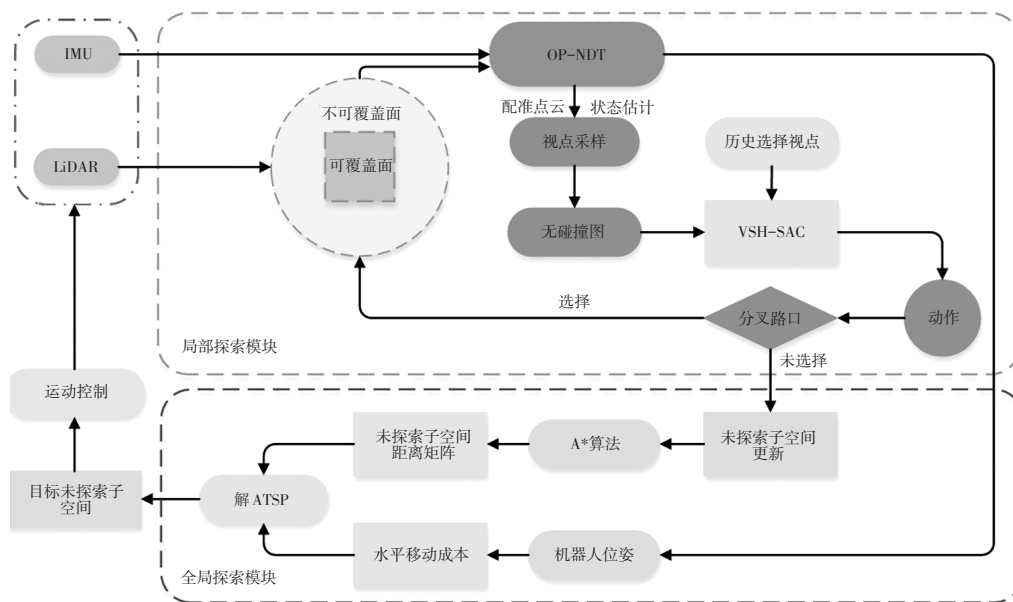


图1 ESDRL算法框架

定义 SCR^3 为全局需探索区域,随着探索的进行,全局区域可分为已探索子空间集合 S_k 、未探索子空间集合 S_u 和正在探索的子空间 S_v ,当不存在正在探索的子空间时探索结束,此时 $S=S_k \cup S_u$ 。设当前位置激光雷达扫描的范围为 F ,以固定大小 $20\text{ m} \times 20\text{ m}$ 作为可覆盖面 F_c ,其余范围为不可覆盖面 \bar{F}_c 。定义激光雷达传感器采集视点由集合 $V=[P_v, Q_v]$ 表示,其中 P_v 和 Q_v 分别为视点位置信息和姿态信息,可覆盖面内视点 V_c 设置为已访问视点。局部子空间内无人车的目标是根据采样视点集合实时找到最短的无碰撞轨迹 Γ^* 并覆盖 \bar{F}_c ,当前正在探索的子空间不存在 \bar{F}_c 时即完成局部区域探索:

$$\Gamma^* = \operatorname{argmin} C(\Gamma) \quad (1)$$

式中, Γ 为所有采样视点构成的路径矩阵; C 函数计算路径采样点间欧式距离并求和,从而得到每条路径长度。

以构成最短路径的视点集作为控制点,经过 B 样条曲线(B-Spline Curve)平滑后处理获得最终可行驶局部路径 Γ_{smooth}^* 。

无人车探索过程中遇到分叉路口,未被选择的分叉路段记为未探索空间,如图2所示。定义所有未探索子空间质心集合为 $\hat{g}=\{g_0, g_1, g_2, \dots, g_n\}$,其中 g_0 为初始无人车所在的未探索子空间质心, n 为未探索子空间的数量,根据探索任务的进行, \hat{g} 不断更新。全局探索问题定义为找到一条全局路径 Γ_{global}^* ,该路径穿过当前无人车所在视点位置 v_t^{cur} 和 \hat{g} 中的每个子空间的质心。

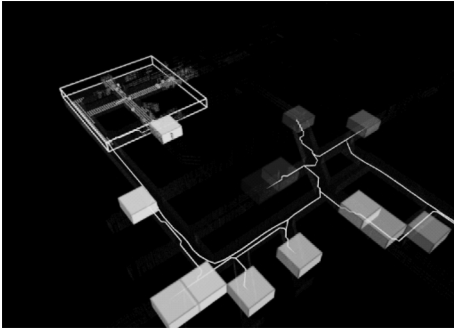


图2 全局探索子空间

2.1 VSH-SAC局部探索算法

本文根据无人车在 t 时刻激光雷达采集的视点集 $V_t=[v_t^1, v_t^2, v_t^3, v_t^4, \dots]$, $\forall v_t^k \in [P_{v_t^k}, Q_{v_t^k}]$ 两两相连,构建无碰撞图 $G_t=(V_t, E_t)$,其中 $P_{v_t^k}$ 、 $Q_{v_t^k}$ 分别为 t 时刻下第 k 个视点的位置和姿态, E_t 为一组可遍历的边。之后将 G_t 作为VSH-SAC算法模型的输入。VSH-SAC算法模型由1个编码器和1个解码器组成:编码器利用自注意力机制从当前局部子空间采样视点构成的图中提取特征;基于提取的特征、无人车当前位置以及通过LSTM存储的上一时刻无人车位姿信息,解码器输出关于相邻视点的策

略,该策略可用于决定下一步访问的视点。本文使用SAC算法来训练策略网络,其中评论家(Critic)网络用于预测状态-行动值。图3所示为VSH-SAC模型整体结构。

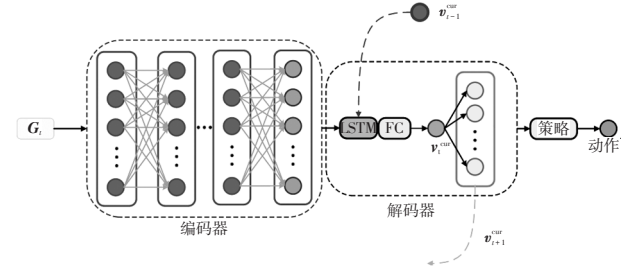


图3 VSH-SAC模型整体结构

编码器由多个堆叠注意层组成,首先将 G_t 中视点嵌入 d 维视点特征中,并计算所有相邻视点边缘掩码 $m_{ij} = \begin{cases} 1, (v_i, v_j) \in E_t \\ 0, (v_i, v_j) \notin E_t \end{cases}$,其中 v_i, v_j 为选择的2个视点, i, j 为选择的视点编号,所有边缘掩码组成边缘掩码矩阵 M 。将视点特征传递给注意层,每个注意层都以前一个注意层的输出作为输入。应用边缘掩码矩阵 M 限制每个视点只能访问其相邻视点的特征。解码器引入LSTM结构实现历史视点信息利用,以选择出符合无人车运动学约束的下一目标点。设当前为 t 时刻,上一时刻的视点 $v_{t-1}^{\text{cur}}=[P_{v_{t-1}}, Q_{v_{t-1}}]$ 与当前无人车所在位置 P_t 和姿态 Q_t 的变化共同影响对下一视点的选择,使SAC能选择出在一条相对平滑路径上且转角较小的下一目标视点。无人车在 t 时刻的观察状态表示为 $o_t=(G_t, P_t)$,VSH-SAC模型输出一个随机策略:

$$\pi_\theta(a_t | o_t) = \pi_\theta(P_{t+1} = v_{t+1}^{\text{cur}}, (P_t, v_t^{\text{cur}}) \in E_t | o_t) \quad (2)$$

式中, θ 为神经网络的权值集合; a_t 为输出的动作。

为了激励无人车尽可能以最短路径和最小转角进行探索,在完成每一个动作后,无人车会收到由4个部分组成的奖励:前3个部分分别为 $r_s=(\bar{F}_c | P_{t+1})$ 、 $r_p=-|P_t - P_{t+1}|$ 和 $r_a=-|Q_t - Q_{t+1}|$,分别是在新时刻下观测到的不可覆盖面以及对新、旧位置和新、旧偏航角之间差距的惩罚,第4个部分为无人车探索完当前局部子空间后得到的固定奖励,即 $r_f = \begin{cases} 0, V_c \neq \emptyset \\ 10, V_c = \emptyset \end{cases}$ 。因此,奖励函数可设计为:

$$r_t(o_t, a_t) = w_s r_s + w_p r_p + w_a r_a + r_f \quad (3)$$

式中, w_s, w_p, w_a 为子奖励函数对总奖励函数的影响系数。

综上,可得SAC最优策略公式为:

$$\pi^* = \operatorname{argmax}_{\pi} E_{(o_t, a_t)} \left[\sum_{t=0}^T \gamma^t (r_t + \alpha H(\pi(\cdot | o_t))) \right] \quad (4)$$

式中, T 为决策步骤数量,从 $t=0$ 开始进行训练; γ 为折扣

因子; α 为正则化系数,用于控制策略熵 H 的重要程度; π 为VSH-SAC模型输出的随机策略; $E_{(\omega, \pi)}$ 为求期望,旨在考虑动作分布的不确定性。

2.2 引入ATSP的全局探索算法

待探索子空间序列的生成对探索任务最终能覆盖的最大区域以及探索效率至关重要。多数方法采用传统的最近邻点策略生成探索序列,每次取当前位置与未探索空间质心的欧氏距离最小的空间作为下一个待探索空间。旅行商问题(Traveling Salesman Problem, TSP)的传统求解方法虽然算法简单,但本质上是一种贪心算法,在大规模环境探索时效率低下。本文采用解ATSP方法生成待探索子空间序列,同时考虑与无人车当前坐标位置距离最短、与无人车当前车身朝向最接近以及需要无人车横、纵向速度变化最小的未探索空间作为下一个待探索子空间。具体地说,首先利用A*算法计算未探索子空间质心间的最短路径并生成距离矩阵 D ,然后将无人车水平移动因素,包括欧氏距离、偏航角变化、待探索空间与无人车当前所在位置之间的速度方向变化作为成本,最后通过解ATSP生成更好的全局探索序列:

$$c_a(\mathbf{g}_i) = D_{\min}(\mathbf{g}_o, \mathbf{g}_i) \quad (5)$$

式中, $c_a(\mathbf{g}_i)$ 为当前无人车所在子空间质心位置 \mathbf{g}_o 与第 i 个未探索子空间质心位置 \mathbf{g}_i 之间位置的成本函数; $D_{\min}(\mathbf{g}_i)$ 为由A*算法生成的与第 i 个未探索子空间之间的最短路径。

同时考虑无人车当前速度 \mathbf{v}_o 和偏航方向对下一个需探索区域的影响,设计成本函数 $c_b(\mathbf{g}_i)$ 和 $c_c(\mathbf{g}_o, \mathbf{g}_i)$:

$$c_b(\mathbf{g}_i) = \arccos \frac{(\mathbf{g}_i - \mathbf{P}_o) \cdot \mathbf{v}_o}{\|\mathbf{g}_i - \mathbf{P}_o\| \|\mathbf{v}_o\|} \quad (6)$$

$$c_c(\mathbf{g}_o, \mathbf{g}_i) = \min \left\{ \frac{L(\mathbf{g}_i, \mathbf{P}_o)}{V_{\max}}, \frac{\min(|Q_o - Q_i|, 2\pi - |Q_o - Q_i|)}{Q_{\max}} \right\} \quad (7)$$

式中, \mathbf{P}_o 为当前无人车所在位置; $L(\mathbf{g}_i, \mathbf{P}_o)$ 为 $\mathbf{g}_i, \mathbf{P}_o$ 之间的欧氏距离; V_{\max} 为无人车可达到的最大速度; Q_o, Q_i 分别为第 i 个未探索子空间质心偏航方向(生成未探索子空间时的无人车朝向)和当前无人车偏航方向; Q_{\max} 为无人车最大偏航角度。

根据式(6)计算得出无人车当前所在位置与下一个待探索子空间位置受无人车速度约束的成本函数,使得规划的下一个待探索子空间横、纵向速度变化最小。最终,得到解ATSP的成本矩阵:

$$M_{\text{ATSP}}(\mathbf{g}_o, \mathbf{g}_i) = c_a(\mathbf{g}_o, \mathbf{g}_i) + w_b c_b(\mathbf{g}_i) + w_c c_c(\mathbf{g}_o, \mathbf{g}_i) \quad (8)$$

式中, w_b, w_c 为2个成本权重系数。

根据成本矩阵可得到一条串连所有未探索子空间的全局路径。

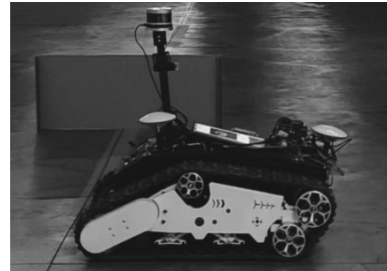
3 仿真与试验验证结果与分析

3.1 验证设施及环境

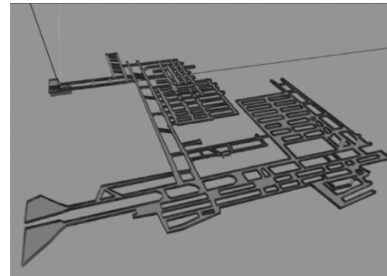
本文通过Gazebo仿真平台进行仿真测试并利用搭载激光雷达及惯性导航设备的履带式差速运动无人车进行实车测试,仿真和试验平台如图4所示。仿真环境采用文献[14]中搭建的矿道模型,长约为280 m,宽约200 m,可行驶道路总面积约为16 000 m²。



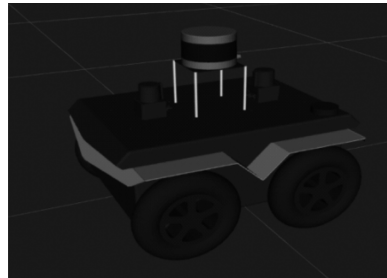
(a)真实地下矿洞场景



(b)真实履带车平台



(c)矿道仿真场景



(d)差速仿真无人车模型

图4 仿真和试验平台

VSH-SAC模型训练环境为Ubuntu 20.04系统,硬件平台采用AMD Ryzen 7 5800H处理器、16 GB内存和GeForce RTX 3060显卡。仿真训练环境为采用Gazebo搭建的矿道环境。由于探索环境规模较大,在网络训练时,设置最大决策步骤数 $T=1\ 000\ 000$ 、折扣因子 $\gamma=1$ 。

使用Adam优化器,设置学习率为0.000 2。真实场景试验测试设备为基于Ubuntu和机器人操作系统(Robot Operating System, ROS)的车载CPU工控机设备,其CPU型号为i5-7260U。

3.2 评价指标

自主导航探索与建图关注探索的效率和覆盖范围,本文采用文献[14]中提出的探索时间和探索区域面积曲线、探索时间和无人车行驶距离作为基本指标,同时,针对不同算法对比单次运行时间和探索效率。在探索时间越短且无人车行驶路程越短的情况下能够获得更多的探索区域是自主导航探索领域追求的更优结果。

3.3 对比仿真与结果分析

为验证本文算法的效果,人为控制探索时间分别为1 000 s和2 000 s,从而比较本文方法与传统的采样方法(NBVP、GBP、TARE)在Gazebo搭建的矿道仿真场景内探索面积和行驶距离等的差异,仿真结果如表1所示。在相同探索时间条件下,本文提出的ESDRL算法能在行驶较少路程的情况下探索更多的未知区域。尽管深度强化学习方法相较于传统方法少量增加了单步的运算时间,但加入视点选择奖惩机制使得无人车在局部探索过程中避免了大角度转向和多次向已探索区域的运动,从而提高了整体探索效率。

表1 Gazebo矿道仿真环境下不同算法对比

探索算法	探索时间/s	探索面积/m ²	行驶距离/m	单次时间/s	探索效率/m ² ·s ⁻¹
NBVP	1 000	2 287.9	3 026.1	7.49	2.29
GBP		1 692.7	892.6	8.86	1.70
TARE		3 774.0	1 689.4	1.01	3.78
ESDRL		3 806.9	1 532.8	1.25	3.81
NBVP	2 000	2 921.8	7 489.9	7.46	1.46
GBP		2 153.1	1 856.7	8.83	1.08
TARE		5 763.9	3 895.6	1.12	2.88
ESDRL		6 110.2	3 686.2	1.28	3.06

图5展示了不同算法在该矿道仿真环境下的探索效果,无人车重复经过的区域探索面积不被统计,即图5中曲线水平部分代表无人车此刻处于已探索过的区域。值得注意的是,由于矿道环境的复杂性,存在许多死胡同和道路交叉口,无人车进入这些死胡同或交叉口时需折返探索或选择探索方向。由图5可知:本文方法和TARE方法的探索面积远超GBP和NBVP方法;本文方法在第750~900 s和第1 250~1 600 s处进行掉头折返探索,无效探索时间约500 s,TARE方法在第500~750 s、第900~1 300 s、第1 750~2 000 s处进行掉头折返探索,

无效探索时间约900 s,因此本文方法相较于TARE方法重复经过的路段更少。此外,本文方法和TARE方法在不同时间点(第500 s、第750 s和第1 500 s)下探索时间与探索面积曲线的差异,体现了不同算法计算得到的下一目标点的不同,从而引导无人车在交叉路口处驶入了不同的矿道区域。

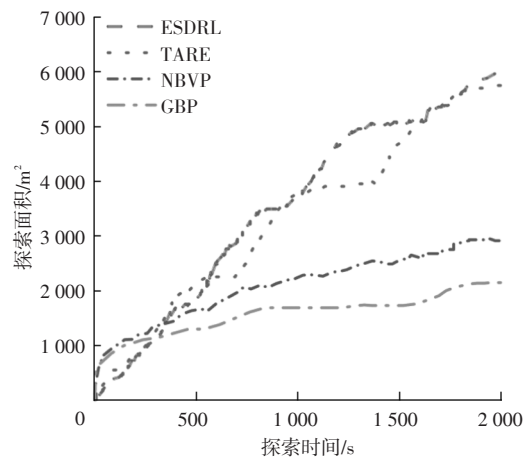


图5 不同算法在仿真矿道环境下探索时间与探索面积对比

综上所述,本文所提出的ESDRL方法相较于传统方法在相同的探索时间内能够在行驶较少路程的情况下探索更多的未知区域,具备更高的探索效率和准确性。

在矿道仿真环境探索中,ESDRL方法成功探索矿道未知区域并建立了三维点云鸟瞰地图,如图6所示。此外,将ESDRL方法建立的三维点云鸟瞰地图与TARE方法建立的三维点云鸟瞰地图进行部分区域对比,如图7所示:在对比区域1中,本文方法在探索路径上相较于TARE方法更加平滑且没有掉头探索,而TARE方法则折返从第2条矿道进行探索,增加了探索路程;对比区域2中,本文方法对历史探索信息进行选择,不进入窄道继续探索,引导无人车驶入其他未探索区域,而TARE方法重复2次进入窄道探索;对比区域3中,本文方法考虑历史和当前车身位姿,相对TARE方法较少出现大角度转弯的情况,使得探索路径更为平滑。

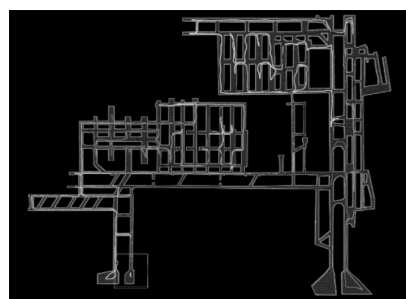


图6 ESDRL算法探索仿真矿道环境建立三维点云鸟瞰图及探索轨迹

通过仿真对比,验证了ESDRL算法在大规模未知汽车技术

矿道仿真环境探索的可行性,证明了ESDRL算法相较于目前最先进的方法在探索效率和探索覆盖率方面表现更为出色。

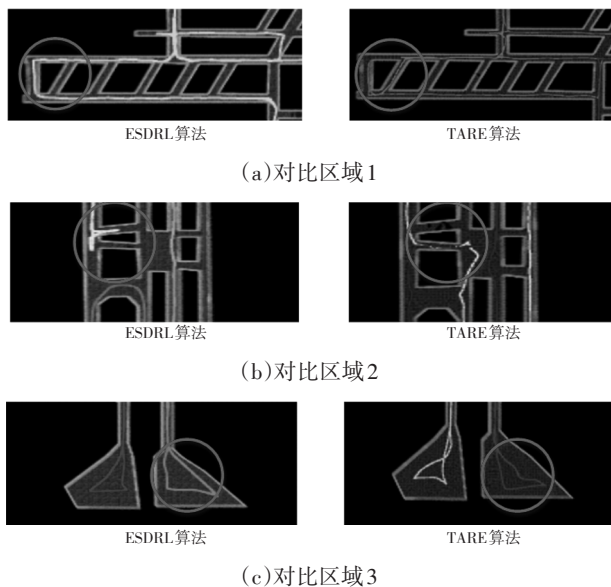


图7 ESDRL与TARE算法三维点云鸟瞰图部分区域对比

3.4 真实场景试验

真实试验环境是北京联合大学北四环校区南门地下车库场景,由于该环境之前未进行建图,因此可以归类为未知场景,如图8所示。地下车库面积约为3444.3 m²。设定无人车最大速度 $v=0.5$ m/s,局部探索区域大小设定为40 m×40 m×15 m,采用基于OP-NDT的激光SLAM算法为无人车提供实时定位信息并同步建图。将本文算法与目前最先进的TARE算法进行实车探索对比,最终获得的点云及轨迹路线如图9所示。



图8 北京联合大学北四环校区南门地下车库场景

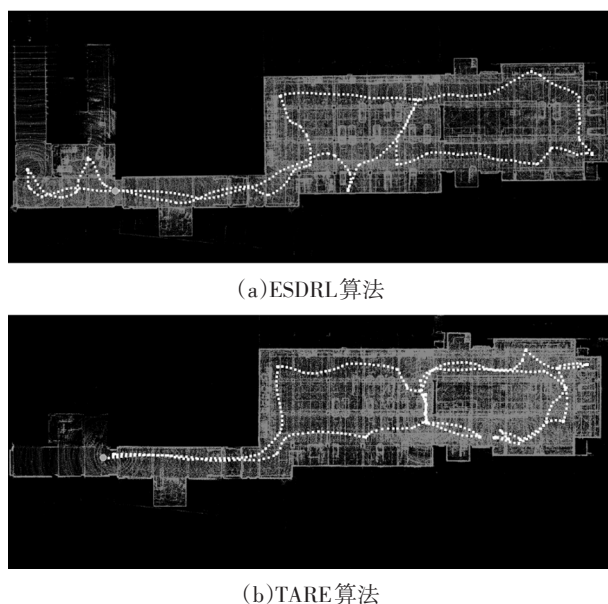


图9 2种算法探索获得的地下车库场景点云及探索轨迹

ESDRL算法在1014 s内完成3444.3 m²的地下车库探索并建立环境三维点云地图。由于初始无人车位姿朝向图9a中正右方,因此无人车首先进行地下车库右半部分探索,之后折返进行左半部分探索,最终回到起点位置。TARE算法总运行时长1500 s,有效探索面积约3365.1 m²。TARE算法出现陷入局部环境并重复绕圈的情况,最终折返至起点处,TARE算法认为不存在未探索区域而并未进入地下车库左半部分探索。对比试验结果表明,本文自主导航探索框架可部署于真实车辆进行未知场景无碰撞探索,并能结合SLAM算法实时精确定位,建立未知环境的场景地图,同时验证了本文所提出的算法在探索完成度和探索效率方面均优于TARE算法。

4 结束语

本文针对自主导航探索算法易陷入局部区域的问题,提出了融合采样与深度强化学习的ESDRL探索算法。局部引入长短期记忆网络获得无人车历史位姿信息进而避免重复走向已探索区域。全局导航探索基于改进传统旅行商问题算法,以无人车水平移动因素作为成本,通过解非对称旅行商问题生成更好的全局探索序列。仿真结果表明,本文方法有助于无人车在行驶途中尽可能减少大角度转弯和重复探索已知区域的情况,对比传统方法,在相同探索时间内行驶的路程更短且能较多地探索未知空间。实车试验中,本文方法在短时间内对地下车库进行了完整探索,证明了其可行性。

本文方法中采用了深度强化学习算法,该算法计算时间较长,导致性能过低的车载工控机设备在探索过程

中出现了设备死机的情况。未来的工作将侧重于将局部探索方法轻量化,在保证探索精度的同时提升探索效率。

参 考 文 献

- [1] MA N, LI D Y, HE W, et al. Future Vehicles: Interactive Wheeled Robots[J]. Science China (Information Sciences), 2021, 64(5): 212-214.
- [2] DELMERICO J, MINTCHEV S, GIUSTI A, et al. The Current State and Future Outlook of Rescue Robotics[J]. Journal of Field Robotics, 2019, 36(7): 1171-1191.
- [3] KULKARNI M, DHARMADHIKARI M, TRANZATTO M, et al. Autonomous Teamed Exploration of Subterranean Environments Using Legged and Aerial Robots[C]// 2022 International Conference on Robotics and Automation (ICRA). Philadelphia, PA, USA: IEEE, 2022: 3306-3313.
- [4] DE PETRIS P, NGUYEN H, DHARMADHIKARI M, et al. RMF- Owl: A Collision- Tolerant Flying Robot for Autonomous Subterranean Exploration[C]// 2022 International Conference on Unmanned Aircraft Systems (ICUAS). Dubrovnik, Croatia: IEEE, 2022: 536-543.
- [5] NG C S X, LUM G Z. Untethered Soft Robots for Future Planetary Explorations?[J]. Advanced Intelligent Systems, 2021, 5(3).
- [6] 刘子毅, 李铁军, 孙晨昭, 等. 基于BIM的建筑机器人自主导航策略优化研究[J]. 计算机工程与应用, 2022, 58(15): 302-308.
- LIU Z Y, LI T J, SUN C Z, et al. Research on Optimization of Autonomous Navigation Strategy of Construction Robot Based on BIM[J]. Computer Engineering and Applications, 2022, 58(15): 302-308.
- [7] AZPÚRUA H, SABOIA M, FREITAS G M, et al. A Survey on the Autonomous Exploration of Confined Subterranean Spaces: Perspectives from Real- Word and Industrial Robotic Deployments[J]. Robotics and Autonomous Systems, 2023, 160.
- [8] YAMAUCHI B. A Frontier-Based Approach for Autonomous Exploration[C]// 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation. Monterey, CA, USA: IEEE, 1997: 146-151.
- [9] SCHMID L, NI C, ZHONG Y, et al. Fast and Compute- Efficient Sampling- Based Local Exploration Planning via Distribution Learning[J]. IEEE Robotics and Automation Letters, 2022, 7(3): 7810-7817.
- [10] XU Z F, DENG D, SHIMADA K. Autonomous UAV Exploration of Dynamic Environments via Incremental Sampling and Probabilistic Roadmap[J]. IEEE Robotics and Automation Letters, 2021, 6(2): 2729-2736.
- [11] BIRCHER A, KAMEL M, ALEXIS K, et al. Receding Horizon “Next- Best- View” Planner for 3D Exploration [C]// 2016 IEEE International Conference on Robotics and Automation (ICRA). Stockholm, Sweden: IEEE, 2016: 1462-1468.
- [12] DANG T, TRANZATTO M, KHATTAK S, et al. Graph- Based Subterranean Exploration Path Planning Using Aerial and Legged Robots[J]. Journal of Field Robotics, 2020, 37(8): 1363-1388.
- [13] DHARMADHIKARI M, DANG T, SOLANKA L, et al. Motion Primitives- Based Path Planning for Fast and Agile Exploration Using Aerial Robots[C]// 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France: IEEE, 2020: 179-185.
- [14] CAO C, ZHU H B, CHOSET H, et al. TARE: A Hierarchical Framework for Efficiently Exploring Complex 3D Environments[C]// Robotics: Science and Systems, 2021.
- [15] CAO Y Y, HOU T X, WANG Y Z, et al. Ariadne: A Reinforcement Learning Approach Using Attention- Based Deep Networks for Exploration[EB/OL]. (2023-01-27)[2023-05-07]. <https://arxiv.org/abs/2301.11575>.
- [16] ZHONG K, LIU Y, YANG J, et al. A Factor Graph Optimization Mapping Based on Normaldistributions Transform[J]. Turkish Journal of Electrical Engineering and Computer Sciences, 2022, 30(3): 1127-1141.
- [17] 赵克刚, 石翠铎, 梁志豪, 等. 基于柔性演员-评论家算法的自适应巡航控制研究[J]. 汽车技术, 2023(3): 26-34.
- ZHAO K G, SHI C D, LIANG Z H, et al. Research on Adaptive Cruise Control Based on Soft Actor- Critic Algorithm[J]. Automobile Technology, 2023(3): 26-34.

(责任编辑 斛 畔)

修改稿收到日期为2023年5月7日。