

DOI: 10.13382/j.jemi.B2206051

基于 Q 学习的生物启发式目标导向导航 路径规划模型*

王典^{1,2} 周阳³ 宋毅³ 代传金⁴

(1. 四川职业技术学院 遂宁 629000; 2. 西华师范大学 南充 637000; 3. 95486 部队 成都 610000;

4. 空军工程大学信息与导航学院 西安 710077)

摘要:为解决未知环境中移动机器人面向目标运行时最优路径获取问题,本文提出一种基于 Q 学习的生物启发式目标导向导航路径规划模型。该模型包括基于 Q 学习的空间探索、基于认知图运行控制和最优路径选择 3 部分。首先,在空间探索中,通过位置细胞的放电情况表征位置状态,采用 ε 动态取值方式进行状态-动作学习,生成认知图,并给出空间探索阶段最优路径。其次,在基于认知图的运行控制中,分别依据最大动作细胞放电率原则和群体动作细胞原则进行运行方向选择,采用多尺度位置更新间距进行位置更新,得到不同认知图下最优路径。最后,对比分析空间探索阶段和运行控制阶段路径规划结果,选取最优路径。仿真结果表明,所提模型可行,采用 ε 动态取值方式进行空间探索可得到较好的路径规划结果;运行体在充分的空间探索后,可提供可行、有效的面向目标运行的路径。

关键词: Q 学习;位置细胞;认知图;路径规划;目标导向导航;仿生导航

中图分类号: Q811; TP273 **文献标识码:** A **国家标准学科分类代码:** 510.80

Model of path planning in biological inspired goal-oriented navigation based on Q -learning

Wang Dian^{1,2} Zhou Yang³ Song Yi³ Dai Chuanjin⁴

(1. Sichuan Vocational and Technical College, Suining 629000, China; 2. China West Normal University, Nanchong 637000, China;

3. Army Unit 95486, Chengdu 610000, China; 4. Information and Navigation College,

Air Force Engineering University, Xi'an 710077, China)

Abstract: To solve the problem of obtaining the optimal path for mobile robots running during goal-oriented running in an unknown environment, a path planning model in biological inspired goal-oriented navigation based on Q -learning is proposed in this paper. The model includes three parts: Spatial exploration based on Q -learning, running control based on cognitive map and optimal path selection. Firstly, in space exploration, the location state is represented by place cells' firing statues, and the state-action is learned by using dynamic ε value, which can generate cognitive map and provide the optimal path in space exploration stage. Secondly, in the running control based on cognitive map, the running direction is selected respectively according to the principle of maximum action cells' firing and the principle of group action cells, and the multi-scale position update intervals are used to update the position. As a result, the optimal path based on different cognitive maps can be obtained. Finally, path planning's result from space exploration stage and running control stage is compared, and the optimal path is selected. Simulation results show that the proposed model is feasible. A better path planning result can be obtained by using the dynamic ε value in space exploration. Besides, a feasible and effective path can be provided for goal-oriented running after sufficient space exploration.

Keywords: Q -learning; place cells; cognitive map; path planning; goal-oriented navigation; bionic navigation

0 引言

目前,移动机器人智能路径规划技术是自主导航领域研究热点之一。生物启发式目标导向导航作为一种智能路径规划方式,可融合生物大脑细胞放电机理,如位置细胞(place cells)^[1]、网格细胞(grid cells)^[2]、头朝向细胞(head direction cells)^[3]、边界细胞(border cells)^[4]等,感知自身所处位置状态,并与动作行为进行关联、学习,以此实现面向目标运行的导航过程。该类导航可有效解决移动机器人未知环境中目标的找寻以及面向目标运行的路径规划问题,提供一种智能、自主的导航方式,是当前移动机器人自主导航研究领域重要的研究方向^[5-7]。

针对生物启发式目标导向导航实现问题,国内外已开展相关研究,如Kunz等^[8]研究了目标导向导航过程中大脑海马区细胞的放电机理,进一步明确了生物启发式目标导向导航生物学基础;Lofriu等^[9]在空间表征中采用多尺度思路进行位置细胞模拟,采用即时差分(temporal difference)学习算法实现状态-动作更新;赵辰豪等^[10]提出一种基于改进Q学习算法的导航认知图构建方法,解决了目标导向导航认知图构建效率低、方向信息不准确问题;方略等^[11]在目标导向导航研究中,基于位置细胞、运动神经元构建了一种前馈神经网络模型,采用Q学习算法实现大鼠面向目标导航任务;Li等^[12]提出一种基于多尺度空间表征的生物启发目标导向导航模型,采用高斯模型构建多尺度位置细胞,并引入模拟退火方法进行行为选择,等等。

从研究现状看,生物启发式目标导向导航主要涉及位置状态的构建、状态-动作的学习以及路径规划关键环节。其中位置状态构建中涉及大脑细胞放电机理的模拟过程,如网格细胞模拟^[13]、网格细胞到位置细胞转换^[14]、视觉位置细胞模拟^[15]等;状态-动作的学习主要通过增强学习算法实现^[16];路径规划部分则是在位置状态构建、状态-动作学习的基础上,提供面向目标运行的可行路径。就路径规划结果而言,不同空间探索方式以及基于不同状态-动作知识得到的路径规划结果不同。为解决最优路径获取问题,本文提出一种基于Q学习的生物启发式目标导向导航路径规划模型。在空间探索中,提出基于 ϵ 动态取值方式进行状态-动作的学习,在基于认知图的运行控制中,提出采用多尺度位置更新间距方式进行运行控制,最后综合分析空间探索阶段和基于认知图的运行控制阶段路径规划结果,以最短路径作为最优路径。

1 位置细胞和动作细胞

1.1 位置细胞

1971年O'Keefe等^[1]在自由活动大鼠海马内发现一些神经元的放电活动具有位置选择性,并将其定义为位置细胞。图1给出了生物实验记录到的大鼠海马位置细胞的放电活动情况,红点表示位置细胞在对应位置处发生了放电。研究表明,位置细胞是动物进行空间环境表征关键细胞,其特异性放电行为可以使动物感知自身所处的位置,并进一步影响动物的导航行为。

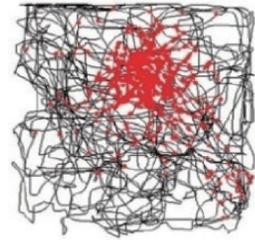


图1 位置细胞放电样式^[17]

Fig. 1 Place cells' firing pattern^[17]

位置细胞的放电活动主要受两类感知信息激发,一是内部自运动感知信息,二是外部视觉感知信息。其中内部自运动感知信息下位置细胞放电活动主要通过网格细胞到位置细胞的转换实现^[18]。外部视觉感知信息下位置细胞放电活动主要通过视觉感知信息与空间位置编码实现。本团队围绕位置细胞建模问题已开展了相关研究,提出了基于径向基神经网络的网格细胞到位置细胞模型^[19-20],基于环境感知和相似性度量的视觉位置细胞生成模型^[21]。鉴于本文研究重点是目标导向导航路径规划问题,因此后续位置细胞构建中,对位置细胞生成过程进行了简化,采用O'Keefe等^[22]提供的位置细胞数学描述方式进行模拟,计算式如下:

$$P_{PCi}(\vec{r}) = e^{-\frac{\|\vec{r} - \vec{r}_{i0}\|_2}{\sigma_{PCi}}}, i = 1, 2, \dots, N_{PC} \quad (1)$$

式中: $\vec{r}_{i0} = [x_{i0}, y_{i0}]$ 为位置细胞*i*放电野中心; $\vec{r} = [x, y]$ 为空间位置坐标; $P_{PCi}(\vec{r})$ 为位置细胞*i*在 \vec{r} 处放电率; σ_{PCi} 为位置细胞*i*的放电野调整参数; N_{PC} 为表征空间环境的位置细胞总数。通过设置 \vec{r}_{i0} 和 σ_{PCi} 可模拟不同样式的位置细胞。

图2给出了两组不同参数下模拟的位置细胞放电情况,颜色深浅代表放电率大小。可见,模拟的位置细胞具有生物位置细胞的放电特性,单个位置细胞放电活动对应于特定的空间区域,所处位置越靠近位置细胞的放电

野中心,位置细胞放电率越大。

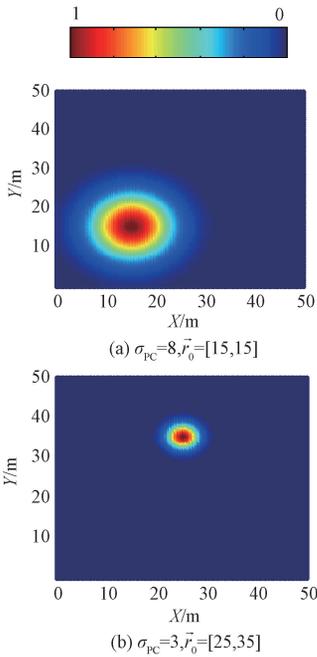


图 2 不同参数下位置细胞的放电情况

Fig. 2 Place cells' firing status in different parameters

1.2 动作细胞

目标导向导航中引入动作细胞是为了创建动作空间,每个动作细胞有一个对应的动作方向,多个动作细胞就可以将动作空间进行离散化。

以 x 轴正方向为基准,各动作细胞表示的方向定义如下:

$$\varphi_{AC_i} = (i - 1) \frac{360^\circ}{N_{AC}}, i = 1, 2, \dots, N_{AC} \quad (2)$$

式中: φ_{AC_i} 为动作细胞 i 表示的动作方向; N_{AC} 为动作细胞总数。通过设置不同数量的动作细胞可以得到不同元素大小的动作空间。

目标导向导航中,运行方向是通过分析实时动作细胞的放电率决定的。空间探索学习过程中,选取的运行方向为最大动作细胞放电率表征的方向。空间探索结束后,运行方向既可以基于最大动作细胞放电率决定,也可以基于群体动作细胞放电情况决定。其中,基于群体动作细胞放电情况进行运行方向选择时,本文采用重心估计原理对群体动作细胞放电率进行处理得到运行方向,计算式如下:

$$\theta(\vec{r}) = \frac{\sum_{i=1}^{N_{AC}} P_{AC_i}(\vec{r}) \varphi_{AC_i}}{\sum_{i=1}^{N_{AC}} P_{AC_i}(\vec{r})} \quad (3)$$

式中: $\theta(\vec{r})$ 表示位置 \vec{r} 处基于群体动作细胞得到的运

行方向; $P_{AC_i}(\vec{r})$ 为动作细胞 i 在 \vec{r} 处放电率。

2 生物启发式目标导向导航路径规划模型

2.1 基本原理

生物启发式目标导向导航路径规划示意图如图 3 所示,包括基于 Q 学习的空间探索、基于认知图的运行控制和最优路径选择 3 部分。整个模型基本原理为:首先,运行体对未知空间环境进行探索,依据最大动作细胞放电率原则选取运行方向,经历目标的找寻、位置细胞增量式生成和基于 Q 学习的状态-动作学习过程,得到基于位置细胞的空间表征知识、面向目标运行的导航知识(认知图)、空间探索阶段最优路径。其次,空间探索结束后,根据不同位置处动作细胞的放电情况,采用重心估计原理得到群体动作细胞下的认知图,并与最大动作细胞放电率原则生成的认知图一起,作为运行控制的导航知识,指引运行体进行面向目标运行的导航行为,得到不同认知图下路径规划结果。最后,分析空间探索阶段和基于认知图运行控制阶段路径规划结果,选择最优的路径。接下来将对以上过程具体实施方法进行详细介绍。

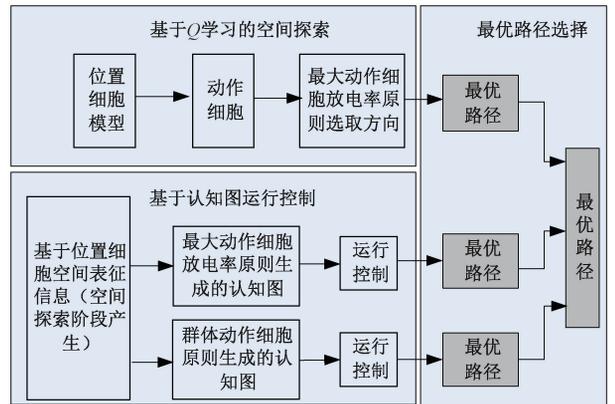


图 3 生物启发式目标导向导航路径规划示意图

Fig. 3 Sketch map of path planning in biological inspired goal-oriented navigation

2.2 实施方法

生物启发式目标导向导航路径规划流程如图 4 所示,主要过程与 2.1 节中示意图的 3 个部分对应。

基于 Q 学习的空间探索流程为:单次空间探索时,运行体从起点出发后,基于位置细胞模型构建位置细胞和判断当前所处状态,然后基于 Q 学习进行运行方向的选择和状态-动作知识的学习,当运行体发现目标或者到达指定位置时,完成该次空间探索,同时判定是否超过空间探索次数要求,若没有,继续从起点出发进行新的空间探索,直到完成规定次数的空间探索过程,得到一条空间探

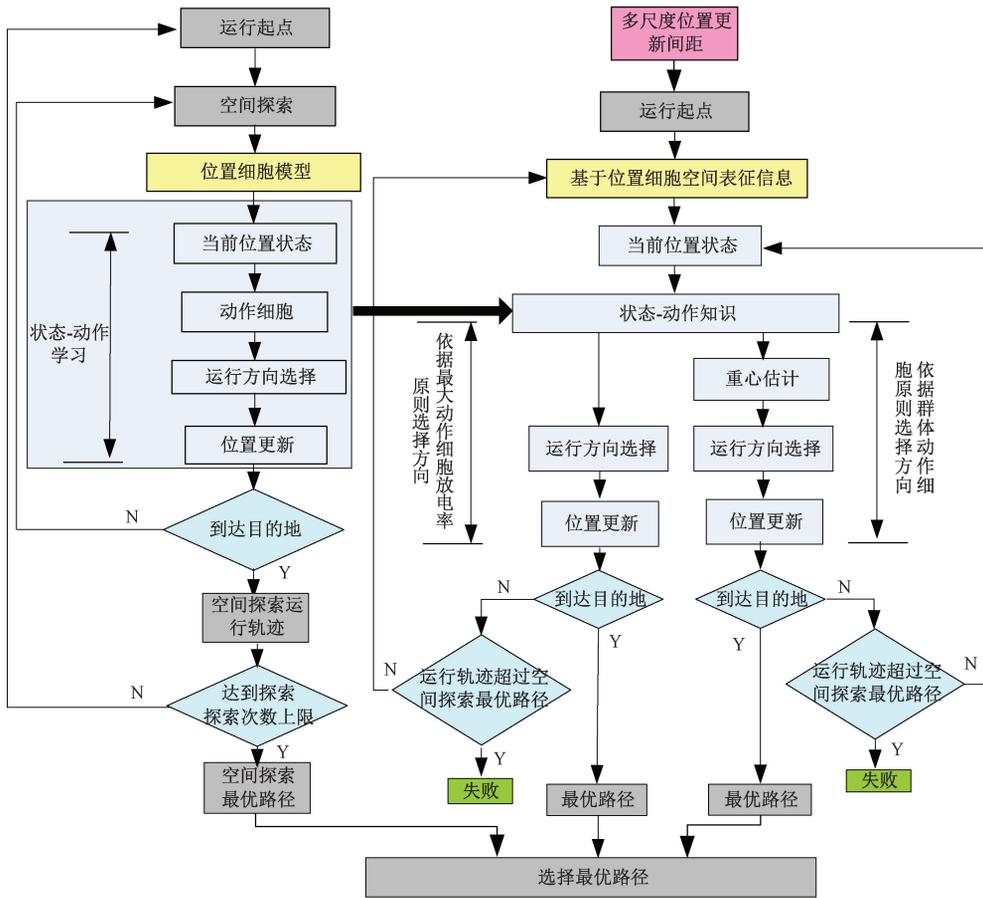


图4 生物启发式目标导向导航路径规划流程

Fig. 4 Flow chart of path planning in biological inspired goal-oriented navigation

索阶段最优路径。

基于认知图的运行控制流程为：运行体在出发前先设定位置更新间距长度，然后从起点出发，依据空间环境表征信息判断当前所处状态，并采取不同的运行方向选择原则，完成各位置更新距离下基于不同认知图的目标导向导航，得到各认知图下最优路径规划结果。

最优路径选取来源于3类路径规划结果，即空间探索阶段最优路径、多尺度位置更新间距下依据最大动作细胞放电率原则和群体动作细胞原则得到最优路径。最优路径选取原则为最短路径原则。

1) 基于 Q 学习的空间探索

空间探索解决的是未知环境中目标的找寻、位置细胞增量式生成和基于 Q 学习的状态-动作学习。

步骤(1) 所处位置状态判断。

运行体从起点出发后，根据式(1)计算当前位置处位置细胞放电情况，并依据各位置细胞放电情况感知当前所处位置状态，若当前时刻，位置细胞放电率低于设定的放电率阈值，还需执行位置细胞新增过程，并将新增位置细胞与已有位置细胞合并分析，执行当前位置状态的

判定过程，将最终位置状态表示为 s_i 。

步骤(2) 运行方向选择。

得到所处位置状态后，需要进行当前时刻运行方向选择问题。本文依据 ε -贪婪策略选择运行方向，即：以概率 ε 在定义的动作空间中随机选择一个运行方向，以概率 $1 - \varepsilon$ 选择当前状态下最大动作细胞放电率表征方向作为运行方向，即：

$$\varphi = \operatorname{argmax}_{AC_i} P_{AC_i}(s_i, \varphi_{AC_i}), i = 1, 2, \dots, N_{AC} \quad (4)$$

式中： $P_{AC_i}(s_i, \varphi_{AC_i})$ 表示状态为 s_i 时动作细胞 i 的放电率； φ 为所选的运行方向。

可见，运行方向选择需解决两个问题，一是 ε 的取值，二是动作细胞放电率的计算。

针对 ε 取值问题。考虑到 ε 取值越大，随机选择运行方向的概率越大，但空间探索后期运行体自身已生成较多可用的空间探索知识。因此，本文对 ε 固定取值方式进行了改进，采用动态取值方式，具体如下：

$$\varepsilon_i = \begin{cases} \varepsilon, & i < \lambda N_{\text{explore}} \\ -\frac{\varepsilon(i - N_{\text{explore}})}{N_{\text{explore}} - \lambda N_{\text{explore}}}, & i \geq \lambda N_{\text{explore}} \end{cases} \quad (5)$$

式中: ε_i 为第 i 次空间探索时贪婪策略取值, N_{explore} 为设定的空间探索次数, λ 命名为 ε 动态取值因子, $\lambda \in (0, 1)$ 。

ε 动态取值示意图如图 5 所示。可见, 在空间探索初期, ε_i 取值较大, 当探索次数达到 $\lambda N_{\text{explore}}$ 时, ε_i 呈线性减小, 直到为 0。

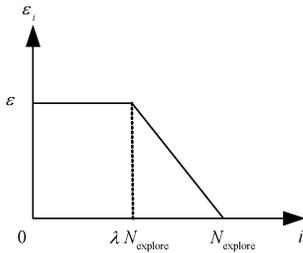


图 5 ε 动态取值示意图

Fig. 5 Sketch map of the dynamic ε value

针对动作细胞放电率计算问题。在基于 Q 学习的空间探索中, 动作细胞放电率可依据 Q 值得到。考虑到 Q 学习过程中, Q 值存在负数情况, 因此本文动作细胞放电率通过对 Q 值进行非负性处理得到, 计算式如下:

$$P_{AC_i}(s_i, \varphi_{AC_i}) = \frac{Q(s_i, \varphi_{AC_i}) - \min Q_{AC}(s_i, \varphi_{AC})}{\max Q_{AC}(s_i, \varphi_{AC}) - \min Q_{AC}(s_i, \varphi_{AC})},$$

$$i = 1, 2, \dots, N_{AC} \quad (6)$$

式中: $P_{AC_i}(s_i, \varphi_{AC_i})$ 、 $Q(s_i, \varphi_{AC_i})$ 分别为动作细胞 i 在状态 s_i 下放电率和对应的 Q 值, $Q_{AC}(s_i, \varphi_{AC})$ 为状态 s_i 下全体动作细胞的 Q 值。

步骤(3)状态-动作知识的学习。

空间探索时状态-动作的学习是通过 Q 学习算法中 Q 值的学习实现, 直到发现目标或者到达指定位置。 Q 值的更新式如下:

$$Q(s_t, \varphi) = Q(s_t, \varphi) + \alpha(R(s_t, \varphi) + \gamma \max_{\varphi_{AC_i}} Q(s_{t+1}, \varphi_{AC_i}) - Q(s_t, \varphi)) \quad (7)$$

式中: $Q(s_t, \varphi)$ 表示状态为 s_t 执行的动作方向为 φ 时对应的 Q 值; α 为学习率, $\alpha \in [0, 1]$; γ 为折扣因子, $\gamma \in [0, 1]$; $Q(s_{t+1}, \varphi_{AC_i})$ 表示状态为 s_{t+1} 动作方向为 φ_{AC_i} 时对应的 Q 值; $R(s_t, \varphi)$ 为状态为 s_t 执行动作方向为 φ 时的报酬值。

经过以上步骤, 运行体可完成单次空间探索过程, 当空间探索次数达到设定的总次数时, 完成整个空间的探索, 最终得到基于位置细胞的空间环境表征信息和状态-动作知识, 同时通过对比分析不同探索过程中的轨迹长度, 得到空间探索阶段的最优路径。

2) 基于认知图的运行控制

基于认知图运行控制就是在空间探索阶段得到的状态-动作知识基础上, 依据不同的运行方向选择策略, 完

成面向目标运行的导航过程, 并得到路径规划结果。

在设定的位置更新间距下, 运行体从起点出发, 各位置处分析位置细胞放电情况, 选取最大放电率位置细胞作为当前位置状态, 然后依据空间探索阶段得到的状态-知识, 采用最大动作细胞放电率表征的方向或者群体动作细胞表征的方向进行运行方向的选择, 并根据设定的位置更新间距更新位置, 直到完成目标导向导航并得到各位置更新间距下路径规划结果。若整个运行控制过程中, 运行轨迹超过空间探索过程中最优路径或者未到达指定位置, 则认为基于认知图的目标导向导航任务失败。

3) 最优路径选择

通过研究发现, 采用不同运行方向选择策略得到的路径规划结果不同, 且在运行控制时, 采用不同位置更新间距进行位置更新时到达目标位置的轨迹长度也有差别。因此, 本文在最优路径选取时, 将两种认知图下实现的路径规划结果均作为最优路径备选结果, 且在基于认知图的运行控制中, 对位置更新间距也进行了多尺度处理, 以此得到不同尺度位置更新间距下两种认知图的路径规划结果, 并将其与空间探索阶段得到的最优路径进行比较, 选择最短路径作为最优路径。

3 仿真实验及分析

3.1 模型仿真实现

考虑到生物启发式目标导向导航路径规划涉及位置细胞模拟和运行控制问题, 本文仿真参数设置时沿用文献[23-24]的设计思路, 对位置细胞模型、运行体运行状态、学习算法等参数进行了设置, 具体如下:

1) 空间区域大小为 $50 \text{ m} \times 50 \text{ m}$, 空间探索阶段采用离散方式进行导航控制和位置更新, 运行速度为 4 m/s , 位置更新周期为 1 s ;

2) 运行起点坐标为 $(0, 0)$, 目标位置坐标为 $(45, 45)$, 设定运行体与目标间距离小于 3 m 时, 认为发现目标或到达指定位置, 否则认为空间探索失败或者目标导向导航任务失败;

3) 空间探索阶段, 位置细胞放电率低于 0.5 时执行位置细胞新增过程;

4) Q 值学习时, 运行体超出设定的空间区域对应动作的 Q 值减少 0.0001 , 运行体到达指定区域时反馈的报酬值设置为 20 , 其余情况, 报酬值为 0 ;

5) 基于认知图的运行控制中, 多尺度位置更新间距取值区间为 $[1, 4]$, 相邻位置更新间距长度为 0.2 m ;

6) 将采用最大动作细胞放电率原则得到的认知图简化为“认知图 I”, 采用群体动作细胞原则得到的认知图简化为“认知图 II”;

7) 其余参数设置如表 1 所示。

表 1 仿真部分参数

Table 1 Partial parameters in simulation

参数	数值
学习率 α	0.8
折扣因子 γ	0.4
贪婪因子 ϵ	0.8
动作细胞数量 N_{AC}	8
ϵ 动态取值因子 λ	0.8

图 6 给出了空间探索结束后生成的位置细胞情况,共生成了 89 个位置细胞。图 7 给出了空间探索次数为 5 次时空间探索阶段最优路径和基于不同认知图得到的路径规划结果,其中图中的菱形表示生成的位置细胞放电野中心位置,箭头表示该位置状态下运行方向。结果表明:1)所提路径规划模型可行,空间探索过程中,运行体能增量式生成位置细胞来表征空间环境,并通过位置细胞的放电情况感知所处的位置状态,并以此进行目标导向导航。2)空间探索不充分时,基于认知图的运行控制无法完成目标导向导航,但空间探索阶段可以提供用于完成导航任务的路径,即便路径不是很理想。可见,采用最优路径选择模式,可以避免空间探索不充分时出现无法完成目标导向导航任务的情况。

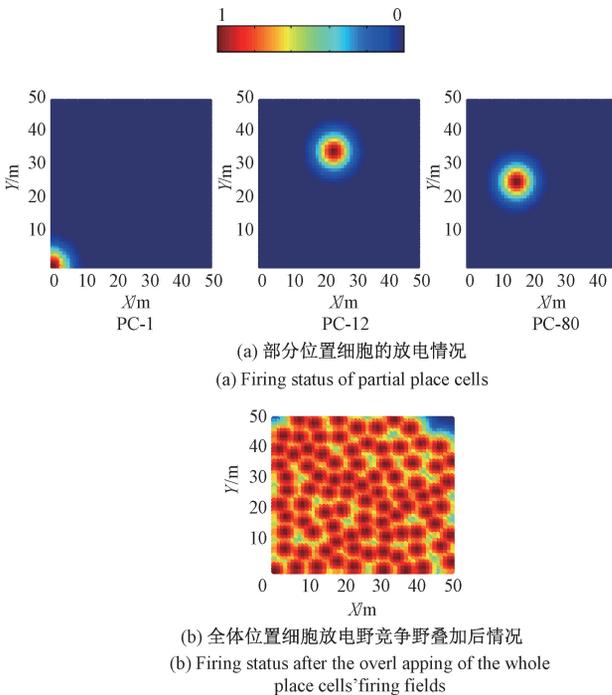
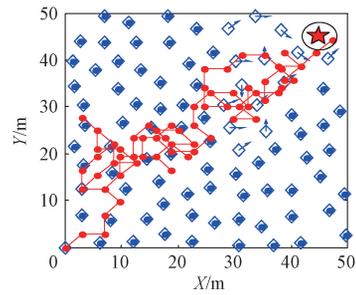


图 6 空间探索结束后生成的位置细胞情况
Fig. 6 Generated place cells after space exploration

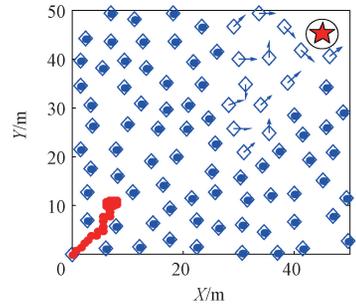
3.2 不同路径规划模式下性能

图 8 给出了空间探索次数为 30 次时不同路径规划模式下路径规划结果,其余参数见 3.1 节条件设置。结



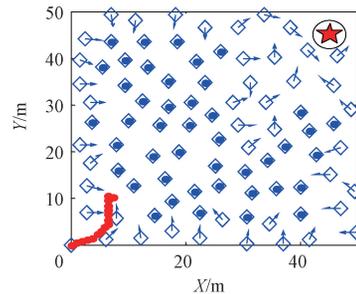
(a) 空间探索阶段最优路径

(a) Optimal path in space exploration stage



(b) 基于认知图 I 路径规划结果

(b) Path planning results based on cognitive map I



(c) 基于认知图 II 路径规划结果

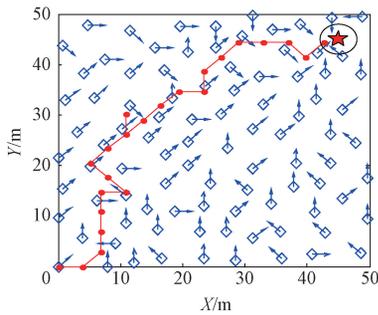
(c) Path planning results based on cognitive map II

图 7 空间探索次数为 5 次时路径规划结果

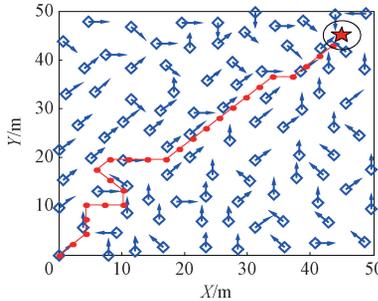
Fig. 7 Path planning results when the space explorations' number is set to 5

果表明,空间探索阶段最优路径轨迹总长 92 m,基于认知图 I 进行路径规划的轨迹总长为 75 m,对应的位置更新间距为 3 m,基于认知图 II 进行路径规划的轨迹总长为 66 m,对应的位置更新间距为 1 m。可见,不同路径规划模型得到的路径规划结果不同,依据最短路径原则,基于认知图 II 的路径规划性能明显优于空间探索阶段最优路径和认知图 I 的路径规划结果,因此可选择基于认知图 II 的路径规划结果作为最优路径,其轨迹总长相比空间探索阶段最优路径减少了 26 m,占比约 28.26%,相比基于认知图 I 的路径规划结果减少了 9 m,占比 12%。

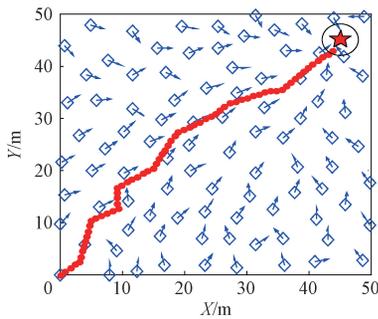
图 9 给出了空间探索次数为 40 次不同路径规划模式下路径规划结果,其余参数见 3.1 节条件设置。仿真结果进一步说明采用不同路径规划模式得到的最优路径结果不同,且明显存在路径长短之别,因此可以通过最短



(a) 空间探索阶段最优路径(总长92 m)
(a) Optimal path in space exploration stage



(b) 基于认知图I最优路径(总长75 m)
(b) Optimal path based on cognitive map



(c) 基于认知图II最优路径(总长66 m)
(c) Optimal path based on cognitive map II

图 8 空间探索次数为 30 次时路径规划结果

Fig. 8 Path planning results when the space explorations' number is set to 30

路径原则选择最终最优路径。可见,所提模型能够实现多模式下路径规划,并获取最优路径。同时,需要注意的是基于空间探索阶段得到的最优路径结果可能劣于其余两种模式,但是该模式是必不可少的,目的是为了出现基于认知图无法实现路径规划的情况,如前述图 7 的仿真情况。

3.3 不同 ε 取值方式下性能

图 10 给出了空间探索阶段采用 ε 动态取值和固定取值时得到的最优路径轨迹长度对比情况。仿真时,空间探索次数为 40, ε 固定取值大小为 0.7, ε 动态取值时对应的 λ 为 0.6,其余参数见 3.1 节条件设置。相同参数下各进行了 10 次实验。结果表明,空间探索阶段采用

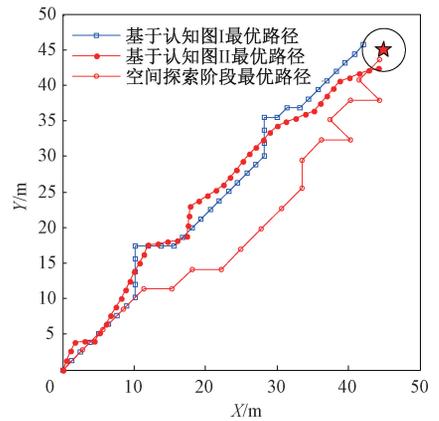


图 9 空间探索次数为 40 次时最优路径

Fig. 9 Optimal path when the space explorations' number is set to 40

ε 动态取值方式可行、有效,相比传统固定取值方式,可缩短路径规划结果的轨迹长度,即:采用 ε 动态取值方式得到的最优路径优于传统固定取值方式。

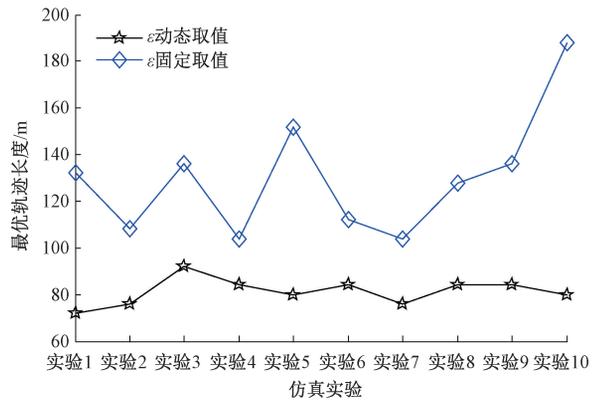


图 10 不同 ε 取值方式下最优路径运行轨迹长度

Fig. 10 Running trajectory length of optimal path when the ε value is set by different modes

3.4 不同位置更新间距下性能

图 11 给出了空间探索次数为 40 次时,不同位置更新间距下基于认知图路径规划运行轨迹长度,仿真中其余参数见 3.1 节条件设置。可见,位置更新间距不同,得到路径规划结果不同。基于认知图 I 的路径规划结果中,位置更新间距为 1.8 m 时,运行轨迹长度最短,为 68.4 m。基于认知图 II 的路径规划结果中,位置更新间距为 1.4 m 时,运行轨迹长度最短,为 65.8 m。对比分析后,选择认知图 II 下最优路径相比认知图 I 下最优路径可减少 2.6 m。因此,基于认知图的运行控制中,通过多尺度位置更新间距的设置可找寻到一条较短、可行的目标导向导航路径,相比传统固定位置更新间距而言,可缩短最优路径的轨迹长度。

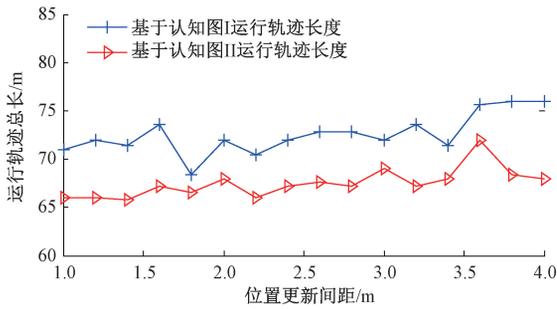


图 11 不同位置更新间距下基于认知图路径规划运行轨迹长度

Fig. 11 Running trajectory length of path planning based on cognitive map when location update intervals are set to different value

4 结 论

未知环境中移动机器人目标的找寻及面向目标的运行可通过目标导向导航过程实现。本文在生物位置细胞放电机理及传统目标导向导航启发下,提出了一种基于Q学习的生物启发式目标导向导航路径规划模型,该模型由基于Q学习的空间探索、基于认知图的运行控制和最优路径选择3部分组成。经仿真验证,所提模型可行,运行体在空间探索中能够增量式生成位置细胞表征环境,并通过位置细胞放电情况感知所处位置状态;当空间探索不充分时,可将探索阶段得到的最短路径作为最优路径;当空间探索充分时,可实现基于认知图的运行控制,并提供目标导向导航路径规划结果,且通过对比分析空间探索阶段和基于不同认知图运行控制阶段路径规划结果,选择最短路径作为最优路径。可见,本文模型能解决运行体空间探索不充分时无法实现目标导向导航过程的问题,且不同空间探索次数下均能通过位置更新间距和 ε 参数的动态取值得到一个最优的路径。

后续,本文将针对移动机器人,搭建实验平台,进一步验证所提模型多路径规划模式下提升路径规划性能的能力,并重点分析不同位置更新间距和不同 ε 取值方式对路径规划性能的影响。本文及后续内容的研究可对未知环境中移动机器人如何快速找寻目标,以及在线或离线生成面向目标运行的最优路径产生理论指导意义和实际应用价值。

参考文献

[1] O'KEEFE J, DOSLROVSKV J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat [J]. Brain Research, 1971, 34(1):171-175.
 [2] MOSER E I, MOSER M B. Grid cells and neural coding

in high-end cortices [J]. Neuron, 2013, 80 (3): 765-774.

- [3] TAUBE J S, MULLER R U, RANCK J B J. Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis [J]. Journal of Neuroscience, 1990, 10(2): 420-435.
 [4] MCNAUGHTON B L, BATTAGLIA F P, JENSEN O, et al. Path integration and the neural basis of the 'cognitive map' [J]. Neuroscience, 2006, 7:663-678.
 [5] SHUGUAN C, MENG M L, JIANTAO F, et al. Decoding route selection of pigeon during goal-directed behavior: A joint spike-LFP study [J]. Behavioural Brain Research, 2021, 409:113289.
 [6] 丛明,邹强,刘冬,等.定位细胞认知机理启发的机器人导航研究综述[J].机械工程学报,2019,55(23): 1-12.
 CONG M, ZOU Q, LIU D, et al. Review of robot navigation inspired by the localization cells' cognitive mechanism [J]. Journal of Mechanical Engineering, 2019, 55(23): 1-12.
 [7] 杨闯,刘建业,熊智,等.由感知到动作决策一体化的类脑导航技术研究现状与未来发展[J].航空学报, 2020, 44(1):023280.
 YANG CH, LIU J Y, XIONG ZH. et al. Brain-inspired navigation technology integrating perception and action decision: A review and outlook [J]. Acta Aeronautica et Astronautica Sinica. 2020. 41(1): 023280.
 [8] KUNZ L, WANG L, LACHNER-PIZ D, et al. Hippocampal theta phases organize the reactivation of large-scale electrophysiological representations during goal-directed navigation [J]. Science Advances, 2019, 5(7): eaav8192.
 [9] LOFRIU M, TEJERA G, CONTRERAS M. Goal-oriented robot navigation learning using a multi-scale space representation[J]. Neural Networks, 2015, 72:62-74.
 [10] 赵辰豪,吴德伟,何晶,等.基于改进Q学习算法的导航认知图构建[J].空军工程大学(自然科学版), 2020, 21(2): 53-59.
 ZHAO CH H, WU D W, HE J, et al. Navigation cognitive map construction based on improved Q-learning algorithm [J]. Journal of Air Force Engineering University (Natural Science Edition), 2020, 21 (2): 53-59.
 [11] 方略,何红军.基于鼠脑海马位置细胞与Q学习面向目标导航[J].生物信息学, 2019, 17(3):31-38.
 FANG L, HE H J. Goal oriented navigation based on place cells of rat's brain hippocampus and Q-learning [J]. Chinese Journal of Bioinformatics, 2019, 17(3): 31-38.

- [12] LI W L, WU D W, DU J, et al. A biologically inspired model based on a multi-scale spatial representation for goal-directed navigation [J]. *KSH Transactions on Internet and Information Systems*, 2017, (11) 3: 1477-1491.
- [13] EDVARSEN V. Long-range navigation by path integration and decoding of grid cells in neural network [C]. *Proceedings of the International Joint Conference on Neural Network*, Anchorage, AK, US, IEEE, 2017: 4348-4355.
- [14] XU X Y, WANG Y H, WANG R B. The place cell activity in three-dimensional space generated by multiple grid cell inputs [J]. *Nonlinear Dynamics*, 2022, 108(2): 1719-1731.
- [15] JAUFFRET A, CUPERLIER N, GAUSSIER P. From grid cells and visual place cells to multimodal place cell: A new robotic architecture [J]. *Frontiers in Neurobotics*, 2015, 9(1): 1-22.
- [16] 郭迟, 罗宾汉, 李飞, 等. 类脑导航算法: 综述与验证 [J]. *武汉大学学报·信息科学版*, 2021, 46(12): 1819-1828.
GUO CH, LUO B H, LI F, et al. Review and verification for brain-like navigation algorithm [J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(12): 1819-1828.
- [17] MOSER E I, KROPFF E, MOSER M B. Place cells, grid cells, and the brain's spatial representation system [J]. *Annual Review of Neuroscience*, 2008, 31: 69-89.
- [18] 陈孟元, 田德红. 基于多尺度网格细胞到位置细胞的仿生 SLAM 算法 [J]. *计算机辅助设计与图形学学报*, 2021, 33(5): 712-723.
CHEN M Y, TIAN D H. Bionic SLAM algorithm based on multi-scale grid cell to place cell [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2021, 33(5): 712-723.
- [19] ZHOU Y, WU D W. Grid-to-place cells model based on radial basis function network [J]. *Electronics Letters*, 2017, 53(3): 200-201.
- [20] 王典, 周阳, 代传金. 改进的基于 RBF 神经网络的网格细胞到位置细胞模型 [J/OL]. *成都理工大学学报 (自然科学版)*: 1-10.
WANG D, ZHOU Y, DAI CH J. Improved model of grid cells to place cells based on RBF neural network [J/OL]. *Journal of Chengdu University of Technology (Science & Technology Edition)*: 1-10.
- [21] ZHOU Y, WU D W. A model of generating visual place cells based on environment perception and similar measure [J]. *Computational Intelligence and Neuroscience*, 2016: 3253678.
- [22] O'KEEFE J, BURGESS N. Geometrical determinants of the place fields of hippocampal neurons [J]. *Nature*, 1996, 381: 425-428.
- [23] 赵辰豪, 吴德伟, 韩昆, 等. 基于多尺度网格细胞的路径整合模型 [J]. *系统工程与电子技术*, 2021, 43(10): 2961-2967.
ZHAO CH H, WU D W, HAN K, et al. Path integration model based on multi-scale grid cell [J]. *Systems Engineering and Electronics*, 2021, 43(10): 2961-2967.
- [24] 阮晓钢, 柴洁, 武悦, 等. 基于海马体位置细胞的认知地图构建与导航 [J]. *自动化学报*, 2021, 47(3): 666-677.
RUAN X G, CHAI J, WU Y, et al. Cognitive map construction and navigation based on hippocampal place cells [J]. *Acta Automatica Sinica*, 2021, 47(3): 666-677.

作者简介



王典, 2012 年于武汉理工大学获得学士学位, 现为四川职业技术学院讲师, 西华师范大学硕士研究生, 主要研究方向为信息通信、智能导航。

E-mail: 527037357@qq.com

Wang Dian received his B. Sc. degree from Wuhan University of Technology in 2012. Now he is a lecturer in Sichuan Vocational and Technical College, and he is a M. Sc. candidate in China West Normal University. His main research interests include information communication and intelligent navigation.



周阳 (通信作者), 2011 年于空军工程大学获得学士学位, 2013 年于空军工程大学获得硕士学位, 2017 年于空军工程大学获得博士学位, 现为 95486 部队工程师, 主要研究方向为仿生导航、智能导航。

E-mail: yydayl@sina.cn

Zhou Yang (Corresponding author) received his B. Sc. degree from Air Force Engineering University in 2011, M. Sc. degree from Air Force Engineering University in 2013, and Ph. D. degree from Air Force Engineering University in 2017, respectively. Now he is an engineer in Army Unit 95486. His main research interests include biomimetic navigation and intelligent navigation.