

一种改进的皮层网络环境认知模型

武悦^{1,2} 阮晓钢^{1,2} 黄静^{1,2} 柴洁^{1,2}

摘要 前额皮层是哺乳动物环境认知能力的重要神经生理基础,许多研究基于皮层网络结构对前额皮层进行计算建模,使机器人能够完成环境认知与导航任务。但是,对皮层网络模型神经元噪声(一种干扰神经元规律放电的内部电信号)鲁棒性方面的研究不多,传统模型采用的奖励扩散方法存在着导航性能随噪声增大而下降过快的问题,同时其路径规划方法效果不好,无法规划出全局最短路径。针对上述问题,本文在皮层网络的基础上引入波前传播算法,结合全局抑制神经元来设计奖励传播回路,同时将时间细胞和位置偏好细胞引入模型的路径规划回路以改善路径规划效果。为了验证模型的有效性,本文复现了心理学上两个经典的环境认知实验。实验结果表明,本模型与其他皮层网络模型相比表现出更强的神经元噪声鲁棒性。同时,模型保持了较好的路径规划效果,与传统路径规划算法相比具有较高的效率。

关键词 皮层网络,波前传播,神经元噪声,环境认知,类脑计算

引用格式 武悦,阮晓钢,黄静,柴洁.一种改进的皮层网络环境认知模型.自动化学报,2021,47(6):1401-1411

DOI 10.16383/j.aas.c190715

An Improved Cortical Network Model for Environment Cognition

WU Yue^{1,2} RUAN Xiao-Gang^{1,2} HUANG Jing^{1,2} CHAI Jie^{1,2}

Abstract Prefrontal cortex is important physiological foundation of environment cognition ability in mammals. Many research seek to make computation model of prefrontal cortex based on cortical network structure, in order to enable robots realize tasks related to environment cognition and navigation. However, there are few works involving in cortical network model's robustness to neuron noise, which is an internal electric signal that generally impedes regular spiking of neurons. Tradition models using reward diffusion method have problem of rapid deterioration of navigation performance under increasing neuron noise. To solve this problem, on the basis of cortical network, this paper recruits wavefront propagation method combined with globally inhibitory neuron to design reward propagating circuit, and introduces time cell and position preference cell into path planning circuit. Two classic environment cognition experiments were reproduced to verify the model. Results show that comparing to other cortical network model, our model exhibits more robustness to neuron noise. Meanwhile, this model keeps good results of environment cognition, and has higher path planning efficiency comparing to traditional path planning algorithms.

Key words Cortical network, wavefront propagation, neuron noise, environment cognition, brain-inspired computing

Citation Wu Yue, Ruan Xiao-Gang, Huang Jing, Chai Jie. An improved cortical network model for environment cognition. *Acta Automatica Sinica*, 2021, 47(6): 1401-1411

环境认知是哺乳动物觅食和生存的基本技能,研究表明^[1-3],哺乳动物普遍具有强大的环境认知能力。研究并复现哺乳动物的环境认知机制对于提高

智能体导航能力、增进人们对生物环境认知机制的理解有着重要意义。有研究发现^[4],前额皮层与高层次的环境认知有关,对前额皮层实施组织切除术后的大鼠不能完成 Morris 水魔方环境认知任务。皮层网络实质上是不同神经元在大脑中的组织方式,其组成单元皮质柱是一种按照神经元功能排列的分层结构。皮层网络对于哺乳动物环境认知的重要作用在于建立感官输入与动作决策之间的联系,形成完整的环境认知能力。

为了使机器人表现出像哺乳动物那样强大的环境认知能力,很多学者对前额皮层网络进行了计算建模研究。皮层网络建模通常涉及奖励传播及路径规划两个问题。对于前者,目前广泛采用奖励扩散^[5-9](Reward diffusion)的方法。该方法以固定的传播

收稿日期 2019-10-16 录用日期 2020-02-23
Manuscript received October 16, 2019; accepted February 23, 2020

国家自然科学基金(61773027),北京市教育委员会科技计划(KM201810005028),北京市自然科学基金(4174083)资助

Supported by National Natural Science Foundation of China (61773027), Project of S&T Plan of Beijing Municipal Commission of Education (KM201810005028), Beijing Natural Science Foundation (4174083)

本文责任编辑 刘艳军

Recommended by Associate Editor LIU Yan-Jun

1. 北京工业大学信息学部 北京 100124 2. 计算智能与智能系统北京市重点实验室 北京 100124

1. Information Department, Beijing University of Technology, Beijing 100124 2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124

系数将奖励值传播给周围神经元, 具有计算简单的优点. 但是, 由于离子通道的随机开闭或神经元的形态变化, 神经元内部存在着噪声信号, 干扰了神经元的规律放电^[10]. 在考虑神经元噪声的情况下, 基于奖励扩散方法的模型由于噪声在神经元传递过程中的累积效应, 导航表现有所下降, 且奖励信息随距离增加而发生衰减. 对于路径规划问题, 皮层网络模型的路径规划主要采用向量法^[11]和位置细胞法^[12-13]. 向量法通过对头朝向细胞的活动加权平均作为动作方向, 位置细胞法以位置细胞中心为规划路径点. 虽然这两种方法都能完成绕近路任务(例如托尔曼三通道迷宫^[14]), 但是规划出的路径均依赖于已探索区域, 无法规划出全局上最短的路径. 一些学者试图将两种方法结合起来以规避缺点^[15-19], 但其模型普遍存在易受神经元噪声影响的问题.

为了提高模型对于神经元噪声的鲁棒性, 本文将波前传播 (Wavefront propagation) 方法引入模型. 波前传播方法^[20-22]利用整合放电 (Integrate-and-fire, IF) 神经元避免奖励信息的衰减, 由此增强模型抗神经元噪声的能力. 其现象已在电生理学实验中被观测证实, 是一种具有生物合理性 (Bio-plausibility) 的计算方法^[23]. 目前, 已有学者将其应用于基于图的导航中^[24-26], 但关于该方法在皮层网络模型中的应用还鲜有报道. 受此启发, 本文将波前传播应用于模型的奖励传播回路, 结合全局抑制神经元, 提出一种改进的皮层网络模型以提高对神经元噪声的抗干扰性. 另一方面, 为了保证模型的

路径规划效率, 模型引入了位置细胞和时间细胞, 利用时间细胞活动随时间增长的特性及其时间常数衰减机制选出全局最佳路径点, 进而提高路径规划性能.

1 皮层网络环境认知与导航模型

生理学研究^[27]表明, 前额皮层中存在对奖励信息编码的神经元, 提供与环境交互所需的动机信号. Preston 等证明了海马体与前额皮层间存在信息交互^[28], 前额皮层接收海马体感官信息的投射, 其中包括对空间位置具有选择性的位置细胞^[29]和对时间编码的时间细胞^[30]. 此外, 前额皮层中存在着同时传输奖励信息和运动信息的神经元^[31].

基于前额皮层生理学结构及其环境认知功能的研究结论, 本文对前额皮层网络进行计算建模. 如图 1 所示, 皮层网络包含奖励传播回路和路径规划回路. 奖励传播回路由奖励细胞和中间神经元构成, 负责在皮层网络中传播奖励信息, 中间神经元通过 STDP 机制对奖励传播路径进行记忆; 路径规划回路由中间神经元和位置偏好细胞构成, 并接收来自海马体位置细胞和时间细胞的输入, 其作用是利用奖励传播回路保存的环境信息进行路径规划, 通过时间细胞对位置偏好细胞的生长性输入提高路径规划效率.

1.1 奖励传播回路

奖励传播回路包含奖励细胞 r 、中间神经元 q_1 .

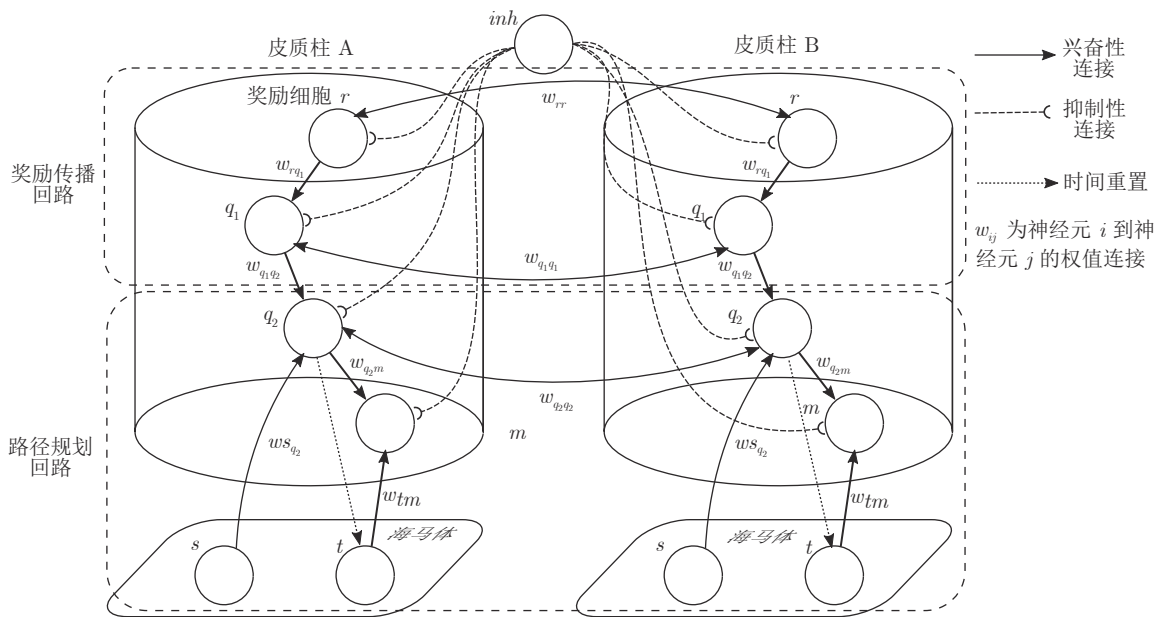


图 1 皮层网络模型结构示意图

Fig.1 Scheme of the cortical column network model

奖励细胞对奖励信息进行编码, 一旦机器人感知到奖励的存在 (如到达逃生平台或放置有食物、水的位置), 该细胞即产生脉冲放电. 为了实现奖励信息在皮层网络中的传播, 奖励细胞之间存在双向连接 w_{rr} . 当奖励细胞被激活后, 脉冲放电向其他奖励细胞传递, 同时也向中间神经元 q_1 传播. 奖励回路采用基于 IF 神经元的波前传播方法, 波前 (某时刻所有放电的神经元构成的膜电位场) 从奖励位置出发传播到自身位置. 在波前传播方法中, 奖励细胞 r 和中间神经元 q_1 的膜电位服从式 (1):

$$\tau \cdot \frac{d(V(t) + N(t))}{dt} = -V(t) + I(t) - V_{inh}(t) \quad (1)$$

其中 $V(t) \in [0, 1]$ 是神经元膜电位, τ 是时间常数, $N(t)$ 是神经元噪声, $V_{inh}(t)$ 是来自全局抑制神经元的抑制性输入. $I(t)$ 是接收其他神经元的整合输入, 服从式 (2):

$$I_i(t) = \sum_{j \in J} w_{ij} \cdot V_{r,j}(t) \quad (2)$$

其中 $V_{*,j}$ 为第 j 个皮质柱中类型为 $*$ 的神经元. 若 $V(t)$ 超过放电阈值 V_{thr} , 则神经元产生放电, 膜电位提升至 $V(t) = 1$. 放电后神经元进入一段时间的抑制状态, 不接受其他神经元的输入, 膜电位服从式 (3):

$$\tau \cdot \frac{d(V(t) + N(t))}{dt} = -V(t), \quad t_f < t < t_f + t_d \quad (3)$$

其中 t_f 是神经元放电的时刻, t_d 是神经元进入抑制状态的持续时间.

中间神经元 q_1 之间的权值 w_{ji} 根据 STDP 学习律变化, 权值改变服从式 (4):

$$\Delta w_{ji} = \begin{cases} +M \cdot \exp(n/\tau_{STDP}), & n \leq -0.02 \\ 0, & -0.02 < n < 0.02 \\ -M \cdot \exp(-n/\tau_{STDP}), & n \geq 0.02 \end{cases} \quad (4)$$

其中 M 是幅值, $n = t_i^f - t_j^f$ 是突触后神经元与突触前神经元间的放电时间差. τ_{STDP} 控制放电时间差对权值变化的影响.

文献 [32] 证明了首个到达目标点的波前总是通过最短的路径, 进而得出波前传播方法等价于 Dijkstra 算法. 因此, 记录首个到达目标点的奖励信息波前传播路径即找到了通往奖励的最短路径. 为了实现对波前传播路径的记忆, 并利用其逆路径进行路径规划, 权值 $w_{q_1 q_1}$ 共享给路径规划回路中的中间神经元 q_2 用于路径规划.

由于每个神经元接收来自相邻神经元的输入, 其中包含了噪声, 该噪声在单一神经元上的积累可能导致其误放电, 从而导致模型失效. 为了解决这

一问题, 本模型在皮质柱外引入了全局抑制性神经元 inh . 抑制性输入需要满足一定的取值范围, 输入过小无法达到消除神经元的误放电的目的, 而输入过大则有可能破坏神经元的正常放电. 考虑到神经元噪声的分布特性^[10], 本文将抑制性神经元 inh 的膜电位 V_{inh} 取值范围设置为服从式 (5):

$$(k+1)\sigma < V_{inh}(t) < \frac{1}{\tau} - (k+1)\sigma - V_{thr} \quad (5)$$

其中 k 为与当前皮质柱存在连接关系的皮质柱数量, σ 为神经元噪声标准差.

1.2 路径规划回路

路径规划回路负责根据奖励传播回路保存的奖励信息进行路径规划. 该回路具体设计思路如下: 如前所述, 为了追求最短导航路径, 须使神经元信号沿所记忆的波前传播路线进行传播, 令中间神经元 q_2 接受中间神经元 q_1 和位置细胞 s 的投射, 利用奖励传播回路权值 $w_{q_1 q_1}$ 在路径规划回路中的副本 $w_{q_2 q_2}$ 将波前传播至目标点. 但是, 由于可能没有位置细胞对全局最短路径涉及的区域进行编码, 该传播路径不一定是全局最短路径. 为了解决这一问题, 在模型中引入时间细胞和位置偏好细胞, 实现分段路径规划. 时间细胞充当计时器, 并为位置偏好细胞提供增长性输入, 使模型在子路径点的规划中经过多个位置细胞, 其构成的直线路径能够穿越潜在的未探索区域, 从而提高效率.

子目标点由位置偏好细胞确定, 由于选取子目标点涉及不同位置偏好细胞的比较, 而神经元放电的瞬时性决定其无法提供这一比较, 因此在该回路中除中间神经元 q_2 外均采用非放电神经元模型. 中间神经元 q_2 需为路径规划回路提供信息传播时序, 仍然采用整合放电神经元模型.

中间神经元 q_2 采用式 (1) 描述的整合放电神经元模型以实现波前传播, 其作用是充当奖励传播回路和路径规划回路的连接, 并为路径规划提供时序. 当奖励传播回路完成奖励信息从目标点到自身位置的传播后, 中间神经元 q_2 被激活, 进入路径规划阶段. 为了使路径规划回路利用奖励传播回路计算的奖励信息, 中间神经元 q_2 共享中间神经元 q_1 的权值, 即 $w_{q_2 q_2} = w_{q_1 q_1}$.

海马体位置细胞 s 编码空间位置, 对特定空间位置产生强烈反应. O'Keefe 通过电生理实验得出海马体位置细胞放电频率的最佳拟合为一高斯函数^[29], 因此本模型位置细胞膜电位 V_s 采用高斯函数描述, 服从式 (6):

$$V_s = \exp\left(-\frac{(x - x_c^2)}{\sigma_s^2}\right) \quad (6)$$

其中 x 是机器人位置, x_c 是位置细胞位置野中心, σ_s 决定位置野的大小. 位置细胞活动是判断是否新增皮质柱的依据, 当所有位置细胞膜电位均小于阈值 $V_{s,thr}$ 时, 说明在当前位置没有相应的位置细胞对其编码, 则新增一个皮质柱并与前一皮质柱建立连接.

位置偏好细胞 m 编码对皮质柱所在位置的偏好程度, 决定路径规划中的子路径点. 在路径规划的每一个子段中, 机器人向值最大的位置偏好细胞直线移动. 位置偏好细胞采用非放电神经元模型, 接收来自中间神经元 q_2 和海马体时间细胞 t 的输入, 其膜电位服从式 (7):

$$V_m(t) = V_{q_2}(t) \cdot V_t(t) \quad (7)$$

其中 $V_{q_2}(t)$ 为中间神经元 q_2 的膜电位, $V_t(t)$ 为时间细胞的膜电位.

为了提高规划路径的效率及对动态环境的自适应性, 在路径规划回路中引入时间细胞. 时间细胞可以使模型在子路径点的规划中经过多个位置细胞, 构成的直线路径可以穿越潜在的未探索区域, 从而提高路径效率. 在路径规划的每个子段中, 时间细胞的活动随时间增长而增加, 从而增加接受其投射的位置偏好细胞活动, 其膜电位服从式 (8):

$$V_t(t) = 1 - e^{-\tau_t(t-t_0)} \quad (8)$$

其中 t_0 为路径规划中每个子段的初始时刻, 即每个子段中神经元 q_2 首次放电的时刻, τ_t 为时间常数.

在时间细胞未饱和阶段 $V_t(t) < V_{t,thr}$, 路径规划子路径点即膜电位最大的位置偏好细胞通过多个皮质柱, 使子路径更加接近子段中的最短路径, 从而在全局上提高路径规划效率. 遇到障碍物时, 增大时间细胞时间常数 τ_t 重新进行子段的路径规划. 路径规划执行过程如算法 1 所示.

1.3 认知地图 (Cognitive Map) 的构建

根据模型的设定, 在机器人探索过的区域需要有相应的位置细胞对其进行编码. 在对新环境的探索过程中, 若所有位置细胞活动均不足, 不能对该位置做出有效的空间感知, 则新增一个编码当前位置的皮质柱, 并建立与前一皮质柱的连接. 因此, 在经过一定时间或距离的运动后, 将产生具有一定规模的皮层网络拓扑图. 皮层网络拓扑图中的奖励细胞能够对奖励信息进行认知, 形成由奖励到空间位置的映射, 由此生成的认知地图可以用于执行寻找食物奖励或逃生等环境认知任务.

认知地图概念最早是由 Tolman 等所提出^[14]. Tolman 等在三通道迷宫实验中发现, 无论环境如何改变, 大鼠均能够找到通往目标点的最短路径,

由此提出大鼠记住的不是某些路径, 而是在脑中生成了一幅认知地图. 本模型所生成的皮层网络拓扑图与认知地图相对应, 是认知地图在计算模型上的体现.

算法 1. 路径规划回路算法

```

1) While  $x \neq x_g$  do
2)   初始化:  $t_0 \leftarrow t, V_t \leftarrow 0$ ;
3)   for  $V_t < V_{t,thr}$  do
4)     更新时间细胞:  $V_t \leftarrow 1 - e^{-\tau_t(t-t_0)}$ ;
5)     更新中间神经元  $V_{q_2}$ ;
6)     更新位置偏好细胞  $V_m$ ;
7)     创建用于下一时刻计算的神经元活动副本
        $V_{old} \leftarrow V$ ;
8)      $t \leftarrow t + \Delta t$ 
9)   end
10)  选取子目标点:  $x_{sub} \leftarrow x_i, i = \arg \max_i V_m$ ;
11)  转向子目标点:  $h \leftarrow \arctan(\rightarrow xx_{sub})$ ;
12)  While  $x \neq x_{sub}$  do
13)    if 未检测到障碍物 then
14)       $x \leftarrow x + \text{step} \cdot [\cos(h), \sin(h)]$ ;
15)      更新位置细胞  $V_s$ ;
16)    else
17)       $\tau_t \leftarrow \eta \tau_t$ ;
18)       $V_{q_2,i} \leftarrow 1, i = \arg \max_i V_{s,i}$ ;
19)      break
20)    end
21)  end
22) end

```

2 实验设计与实验结果

本节通过两个心理学上的经典实验 (托尔曼 14 单元 T 型迷宫、Morris 水迷宫) 对模型的环境认知能力、神经元噪声鲁棒性、动态环境的适应能力及路径规划效率进行测试.

2.1 仿真实验参数

本实验采用仿真机器人作为主体, 具备距离传感器, 能够探测正前方 0.2 m 内是否存在障碍物, 忽略体积将其视为一个质点. 仿真机器人采用本文提出的皮层网络环境认知模型. 对于整合放电神经元 r 、 q_1 、 q_2 , 膜时间常数 $\tau = 0.8$, 放电阈值 $V_{thr} = 0.3$, 放电后抑制阶段持续时间 $t_d = 0.3$ s. 神经元的其他参数在表 1 中列出. 为了奖励信息在皮层网络中传播而不发生衰减, 奖励细胞之间存在饱和双向连接 $w_{rr} = 1$. 中间神经元 q_2 的激活条件是奖励信息

表 1 模型参数值设定
Table 1 Parameter setting of the model

神经元	类型	参数
奖励细胞 r	整合放电型	$w_{rr} = 1, w_{rq_1} = 1$
中间神经元 q_1	整合放电型	$w_{q_1q_2} = 0.1, \tau_{STDP} = 0.02, M = 1$
中间神经元 q_2	整合放电型	$w_{q_2q_2} = w_{q_1q_1}, w_{sq_2} = 0.1$
位置偏好细胞 m	非放电型	$w_{q_2m} = 1, w_{tm} = 1$
位置细胞 s	非放电型	$\sigma_s = 0.35, V_{s,thr} = 0.5$
时间细胞 t	非放电型	$\tau_t = 10, \eta = 2, V_{t,thr} = 0.95$
全局抑制神经元	非放电型	$V_{inh} = 0.1$

传播至自身所在位置, 仅在中间神经元 q_1 和位置细胞 s 同时激活时放电, 因此设置 $w_{q_1q_2} = 0.1, w_{sq_2} = 0.1$. 将神经元参数值代入式 (5), 同时考虑到实际的神经元中抑制性输入通常为较低水平, 本实验设置 $V_{inh} = 0.1$ 为基准值. 仿真时间间隔设置为 $\Delta t = 0.02 \text{ s}$.

2.2 托尔曼 14 单元 T 型迷宫实验

2.2.1 实验背景

托尔曼^[33]曾设计了 14 单元 T 型迷宫实验来观察大鼠在迷宫中的空间认知表现, 并由此提出了认知地图的假设. 如图 2 所示, 该迷宫有唯一一条通往奖励的路径和 14 个无效通道. 将大鼠放置在迷宫的起点, 在另一端放置食物, 测试大鼠直到获取食物时经过的无效通道的次数, 该指标能够反映大鼠对环境的认知程度. 本文采用该实验测试模型的噪声鲁棒性以及动态环境的自适应性.

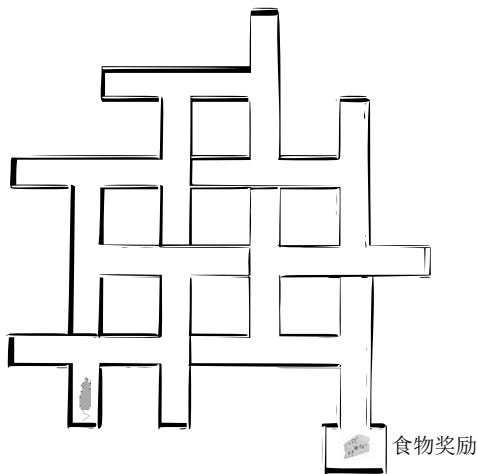


图 2 托尔曼 14 单元 T 型迷宫实验示意图

Fig.2 Sketch of Tolman 14-unit T-maze experiment

2.2.2 实验设计与结果

托尔曼 14 单元 T 型迷宫实验验证了认知地图

理论, 本文在相同的环境下对模型的认知地图构建、神经元噪声鲁棒性和对动态环境的自适应性进行测试, 实验分为准备阶段和测试阶段.

1) 认知地图构建

在准备阶段中机器人从起点出发对迷宫进行探索, 直至找到放置在目标点处的奖励. 机器人在每次直线前进 0.3 m 后进行 60° 以内的随机偏转, 若遇到障碍物则继续改变朝向. 经过探索, 机器人建立起皮层网络 (图 3). 图中, 黑色圆点代表皮质柱, 它们在探索过程中动态生成、连缀成网, 形成皮层拓扑网络, 以这种形式反映智能体对环境的认知, 即构成认知地图.

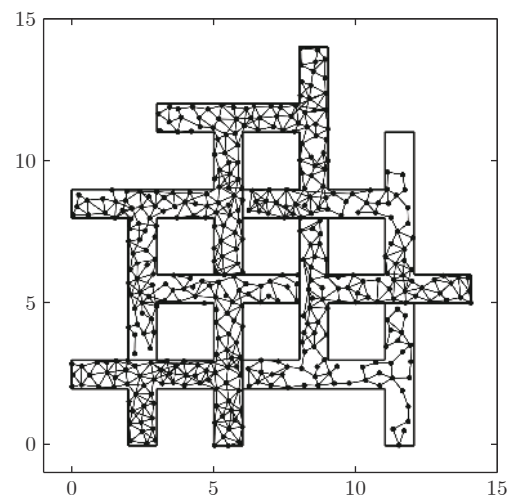


图 3 机器人探索迷宫后建立的皮层网络拓扑图 (m)

Fig.3 Cortical column topological map after exploring the maze (m)

2) 神经元噪声影响

神经元噪声是神经元内部产生的一种噪声电信号, 干扰神经元的正常放电. 文献 [10] 通过电生理学实验得出神经元噪声服从高斯分布, 基于此结论在本实验中向皮层网络中所有神经元添加高斯分布噪声 $N \sim N(0, \sigma)$. 在测试阶段中, 将机器人放置于

起始位置,模拟的目标点奖励使认知地图中对应的奖励细胞放电,激活皮层网络模型使机器人向目标点导航.受神经元噪声的影响,皮层网络内的神经元放电顺序可能被打乱,造成部分奖励信息的丢失.若大鼠停留在同一位置超过 5 s,则认为模型陷入局部最优,此时仿真大鼠随机移动 4 m 后重新进行路径规划,记录到达目标点过程中经过无效通道的次数.为了使实验结果具有统计学意义,总共做 18 组实验,在每组实验中设置不同的噪声标准差进行导航测试,并与采用奖励扩散方法的皮层网络模型进行对比,奖励值传播系数 $\beta = 0.98$,该系数为多次试验得出的能够产生最佳的导航结果的系数.

实验结果如图 4 所示,从箱线图的上下边缘、上下四分位数以及中位数上看,当标准差在 0.2 以内时,本模型经过无效通道次数显著少于奖励扩散方法.箱线图的箱体长度代表数据的波动程度,从图中可以看出,两种模型经过无效通道次数的波动程度均随神经元噪声增大而增加,但是在神经元噪声标准差不超过 0.2 的范围内,本模型对应的箱体更短,表明失败次数的波动更少,进而说明了本模型的路径规划性能相对稳定(图中双星号代表在 5% 的显著性水平下存在显著性差异).

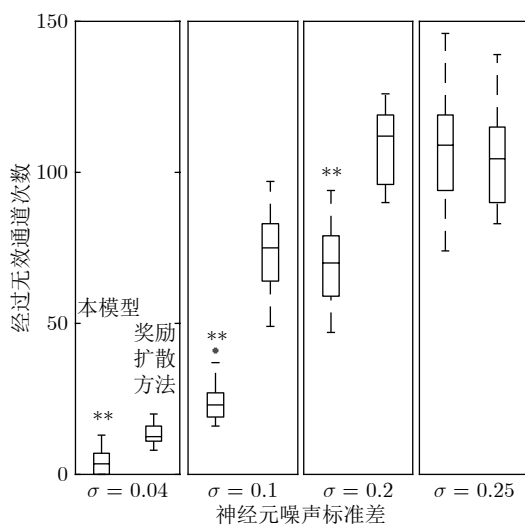


图 4 噪声标准差对导航结果的影响

Fig. 4 Influence of neuron noise standard variation on navigation results

当噪声标准差继续增大至 0.25 时,两种方法的导航表现均接近盲目探索(盲目探索时经过无效通道的平均次数为 122 次).然而在实际中,噪声标准差超过 0.2 是一个小概率事件^[10].实验结果表明,在合理的神经元噪声范围内,本模型的神经元噪声鲁

棒性优于采用奖励扩散方法的皮层网络模型.

3) 对动态环境的自适应性

当环境发生变化时,动物能很快适应这一变化,迅速做出调整,反映出动物在环境认知上的自适应性.本实验模拟这一过程,测试模型对动态环境的自适应能力.如图 5 所示,测试包含 3 次实验:第 1 次,在托尔曼 14 单元 T 型迷宫中移除其中一处墙壁,标记为“1”(图 5 (b));第 2 次,恢复“1”处墙壁,标记为“2”,再移除迷宫中另外两处墙壁,分别标记为“3”、“4”(图 5 (c));第 3 次,在图 5 (c) 的基础上增加一处墙壁,标记为“5”(图 5 (d)).3 次实验对环境的改变逐渐增多,既包含对障碍的移除,也包含对障碍的增加.

令机器人利用准备阶段建立的认知地图在 3 次实验中完成导航,结果如图 5 所示.可以看出,机器人在 3 次环境改变中都能自适应地调整巡航行为,重新找到通往目的地的较短路线,说明模型具有对动态环境的自适应性.

2.3 Morris 水迷宫实验

2.3.1 实验背景

1982 年 Morris 等^[34]在测试大鼠的空间记忆能力时设计了 Morris 水迷宫实验.如图 6 所示, Morris 水迷宫为一圆形水池,其中存在一个隐藏的逃生平台,被放置于水池中的大鼠需要找到并保持在逃生平台上以继续存活.实验结果表明,大鼠首次找到平台后,在后续的测试阶段从水池不同位置放入的大鼠均能够采取一条近似直线的路径逃往平台.然而,破坏海马体组织后大鼠无法成功地找到逃生平台,证明海马体具有空间记忆功能.此外,逃生时采取的路径也体现了海马体相关回路的路径寻优能力.本文采用该环境对模型的噪声鲁棒性和路径规划效率进行测试.

2.3.2 实验设计与结果

本文实验仿照 Morris 水迷宫实验设计如下:水迷宫为一直径 10 m 的圆形水池,水池中有一边长为 1 m 的正方形隐藏逃生平台;水池中的一组中心对称的障碍物用来测试模型的避障能力,中心对称结构能够防止逃生平台与障碍物线索的关联.实验分为准备阶段和测试阶段,首次逃生任务属于准备阶段,后续逃生任务均属于测试阶段.

1) 认知地图构建

在准备阶段中机器人首次执行逃生任务,在逃生过程中对环境进行初次探索.机器人在每次直线前进 0.3 m 后进行 30° 以内的随机偏转.大鼠在

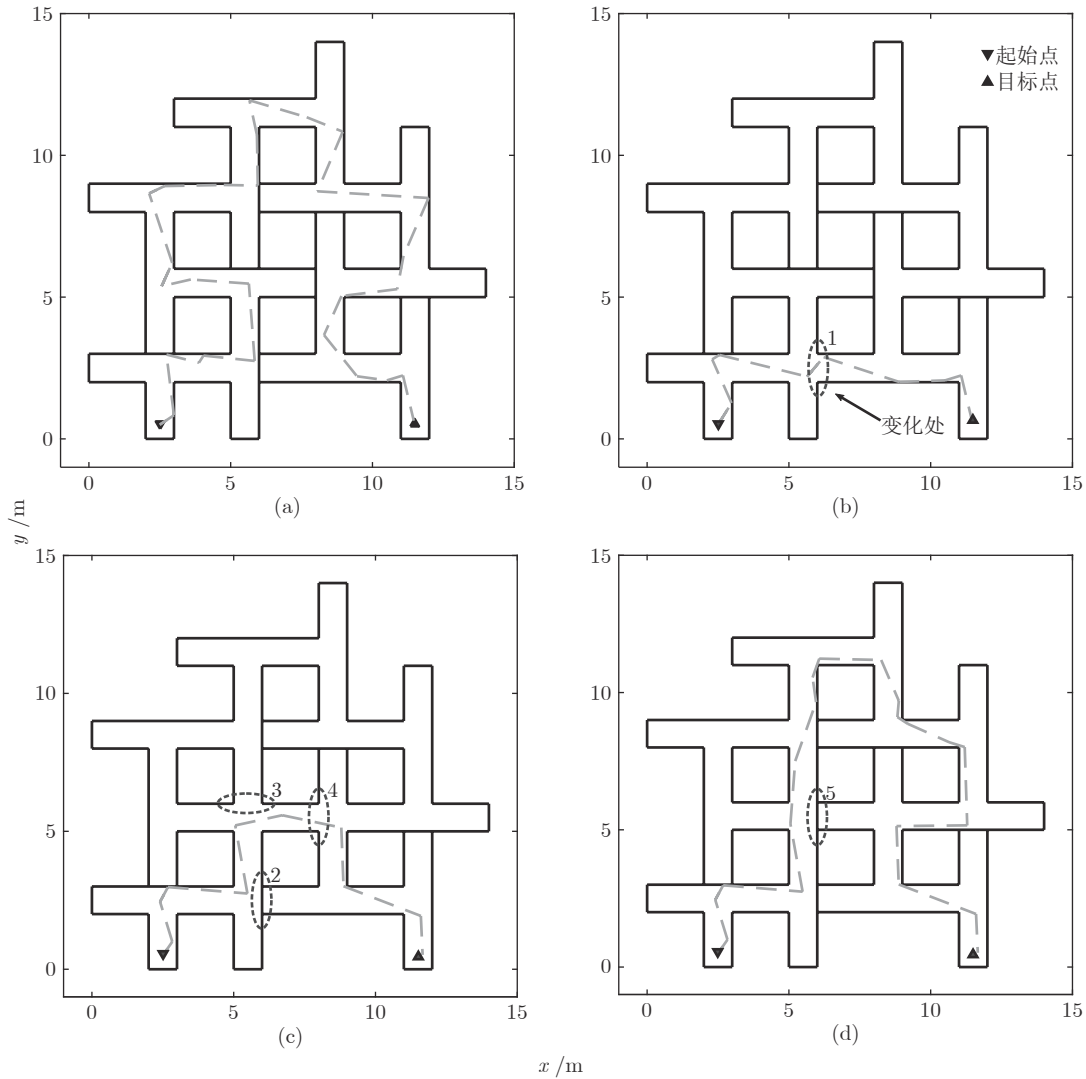


图 5 机器人在环境发生变化前后采取的路线

Fig.5 Path planned before and after environmental change

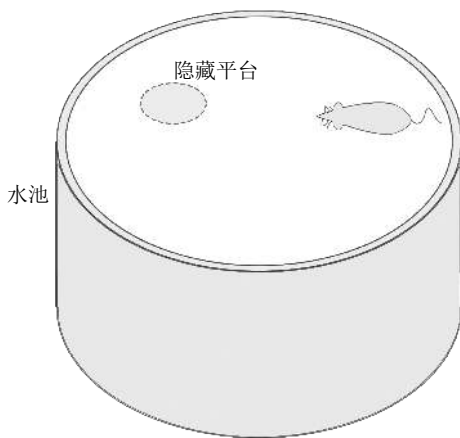


图 6 Morris 水迷宫示意图

Fig.6 Sketch of Morris water maze

Morris 水迷宫实验中表现出沿边缘行进而避开中央区域的行为, 本实验采用相似的设置, 使机器人避开距中心 3 m 内的区域. 模型以皮层网络的拓扑结构来描述机器人逃生时经过的区域, 是认知地图在计算模型层面的具体体现 (图 7).

2) 神经元噪声影响

为测试模型对神经元噪声的鲁棒性, 根据文献 [9] 指出的神经元噪声服从高斯分布的结论, 在实验中向皮层网络中所有神经元添加高斯分布噪声 $N \sim N(0, \sigma)$. 在测试阶段中, 将机器人放置于水池内不同起始位置, 模拟的逃生本能将对对应逃生平台位置的奖励细胞激活, 使机器人向逃生平台导航. 若大鼠停留在同一位置超过 10 秒, 则认为模型陷入了局部最优, 该次导航失败. 为排除偶然因素的干扰, 共进行 18 组相同的测试, 在每组中设置不同

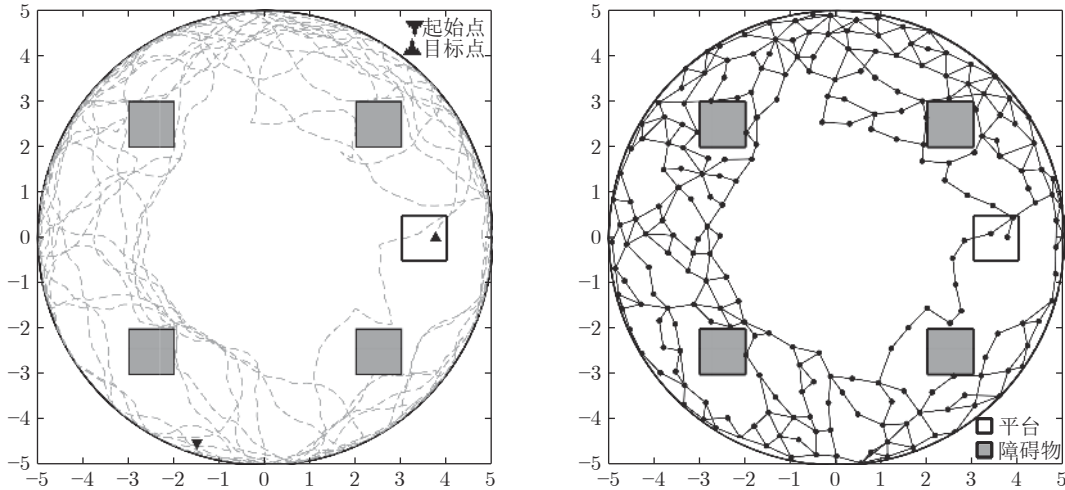


图 7 Morris 水迷宫路逃生实验中的移动轨迹 (左) 和建立的皮层网络拓扑图 (右) (m)

Fig.7 Moving trace on the preparing stage (left) and the established cortical column network (right) (m)

的噪声标准差进行导航测试, 并与采用奖励扩散方法的皮层网络模型进行对比.

实验结果如图 8 所示, 由于波前传播的整合放电特性以及全局抑制神经元的抑制作用, 本模型在神经元噪声标准差较小时 ($\sigma = 0.02$) 导航未受影响, 其导航失败率为 0, 而奖励扩散方法的失败率中位数达到了 20%. 当噪声标准差增大至 0.05 和 0.1 时, 本模型的导航失败率的最大最小值、上下四分位数以及中位数始终低于奖励扩散方法, 当噪声标准差较大时 ($\sigma = 0.2$), 本模型导航失败率中位数达到 0.77, 而奖励扩散方法完全失败. 该结果表明, 本论文模型的神经元噪声鲁棒性显著优于采用奖励扩散方法的皮层网络模型 ($p = 0.05$).

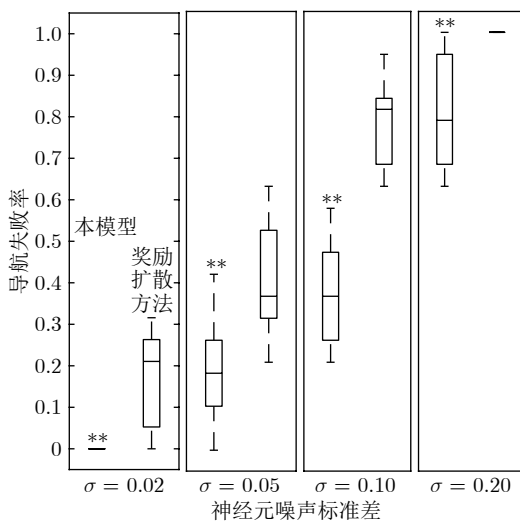


图 8 不同噪声对导航的影响

Fig.8 Influence of neuron noise on navigation results

3) 路径规划效率

在测试阶段, 将机器人放置于水池内不同起始位置执行逃生任务, 神经元噪声设置为 $N \sim N(0, 0.01)$. 在其中几次逃生中 (图 9), 机器人能够穿越未探索过的区域 (与图 7 所建立的认知地图相比), 说明本模型的路径规划是基于全局而完成的.

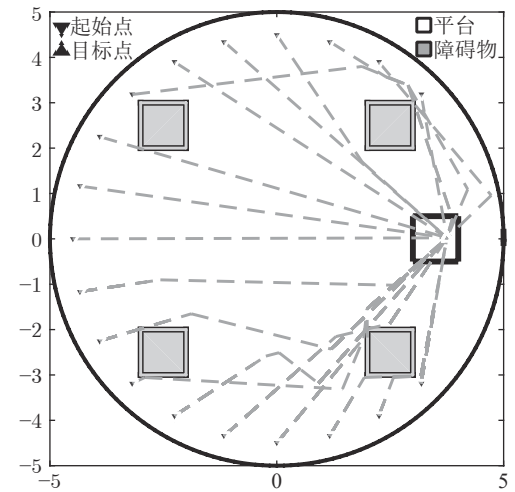


图 9 从不同位置出发的逃生路线 (m)

Fig.9 Escape trace from different starting points (m)

为了说明本模型的路径规划效果, 本实验选择了 A* 算法^[35] 和滚动窗口快速扩展随机树算法 (Rapid-exploring random tree, RRT)^[36] 进行了对比. 两者均为路径规划中经常使用的流行算法. 为使 A* 算法与 RRT 算法在本实验中表现出最佳效果, 本文通过多次实验设计算法细节及参数如下: A* 算法中节点评估函数为 $f(x) = g(x) + h(x)$, 其中 $g(x)$ 为已选取节点的累计路程, $h(x)$ 为当前位置

到目标点的欧氏距离; 滚动窗口 RRT 算法中窗口半径 $r = 1.5$ m, 随机树生长步长 $\rho = 0.3$ m, 窗口范围内随机树节点数量 $K = 25$. 步长与本论文中皮质柱平均间隔相近, 以保证在相近的条件下进行路径规划的比较.

在路径规划效果指标的选取上, 本文参考了文献 [37], 采用转弯次数、角度以及路径长度作为对比指标. 转弯次数及角度的增加通常对机器人的运动速度带来限制, 造成频繁的减速、停止和加速, 降低导航效率. 因此, 较少的转弯次数、转弯角度以及较短的导航路径通常意味着更好的路径规划效果.

图 10 对不同方法所规划的路径长度进行了对比. 从该图可以看出, 除出发点角度 (出发点与水池中心的连线与 x 轴正半轴夹角的绝对值) 为 60° 时本模型规划的路径长度无明显优势外, 其余出发点角度下本模型规划的路径长度都明显短于另两种算法.

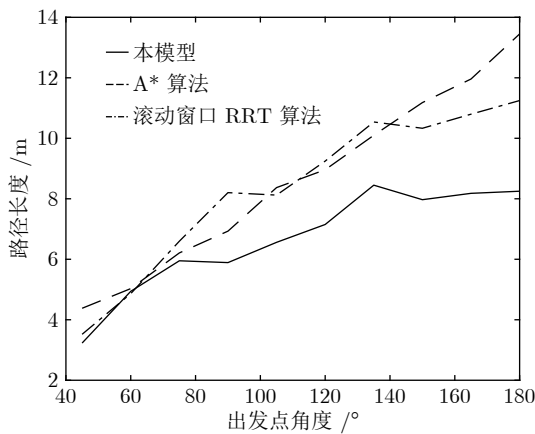


图 10 不同规划方法所规划路径长度

Fig. 10 Length of planned path by different planning method

表 2 列出了本实验中从所有角度出发的平均转弯次数和平均累计转弯角度. 由实验结果可以看出, 本模型的平均转弯次数及平均累计转弯角度远低于 A* 算法和滚动窗口 RRT 算法, 说明规划的路径更加平滑, 有效地提高了路径规划及导航的效率.

表 2 不同方法规划路径的转弯次数及转弯角度对比
Table 2 Comparison of turning counts and angle of path planned by different path planning methods

神经元	平均转弯次数	平均累计转弯角度
本模型	1.9	28.36°
A* 算法	17.55	331.9°
滚动窗口 RRT 算法	12.46	177.25°

3 结论

哺乳动物的皮层网络整合感官信息并输出动作决策, 该信息处理中枢对环境认知的重要性引起了许多对皮层网络的计算建模研究. 针对目前的皮层网络模型受神经元噪声影响较大的问题, 本文利用整合放电模型的非线性特性对神经元进行建模, 将波前传播方法引入奖励传播回路, 同时结合全局抑制神经元, 使模型具有更强的抗噪声干扰性. 同时, 为了防止路径规划效率的下降, 在路径规划回路中引入位置偏好细胞和来自海马体的时间细胞, 通过时间细胞及其时间常数衰减机制选出全局最佳路径点. 模型结构、各神经元功能设计及波前传播方法都基于目前的神经生理学研究结论, 具有一定的生物合理性.

为了验证模型的有效性, 本文复现了心理学上两个经典的环境认知实验: 托尔曼 14 单元 T 型迷宫实验和 Morris 水迷宫实验, 并与其他皮层网络模型和路径规划方法进行了对比. 实验数据表明, 本文所提出的模型与其他皮层网络模型相比具有更强的神经元噪声鲁棒性, 较之 A* 算法和滚动窗口 RRT 算法具有更好的路径规划效果. 同时, 模型也表现出较强的对动态环境的自适应能力.

另一方面, 目前模型中时间细胞的时间常数衰减机制较为简单, 需要进一步优化以提高最佳路径点的选取能力及搜寻速度. 为了使模型具有更佳的环境适应能力, 时间常数随环境的变化机制需要进一步探讨. 如何构建时间常数衰减机制涉及的回路, 使其符合生理学研究发现, 也是下一步研究的重点.

References

- Contreras M, Pelc T, Llofriu M, Weitzenfeld A. The ventral hippocampus is involved in multi-goal obstacle-rich spatial navigation. *Hippocampus*, 2018, **28**: 853–866
- Vorhees C, Williams M. Assessing spatial learning and memory in rodents. *ILAR Journal*, 2014, **55**(2): 310–332
- Bucci D, Chiba A, Gallagher M. Spatial learning in male and female long-evans rats. *Behavioral Neuroscience*, 1995, **109**(1): 180–183
- Granon S, Poucet B. Medial prefrontal lesions in the rat and spatial navigation: Evidence for impaired planning. *Behavioral Neuroscience*, 1995, **109**(3): 474–484
- Martinet L E, Sheynikhovich D, Benchenane K, Arleo A. Spatial learning and action planning in a prefrontal cortical network model. *Public Library of Science Computational Biology*, 2011, **7**(5): 1–21
- Martinet L E, Passot J B, Fouque B, Meyer J A. Map-based

- spatial navigation: A cortical column model for action planning. In: Proceedings of International Conference Spatial Cognition, Freiburg, Germany: Springer, 2008. 39–55.
- 7 Erdem U M, Hasselmo M E. A biologically inspired hierarchical goal directed navigation model. *Journal of Physiology-Paris*, 2014, **108**(1): 28–37
- 8 Chersi F, Pezzulo G. Using hippocampal-striatal loops for spatial navigation and goal-directed decision-making. *Cognitive Processing*, 2012, **13**(1): 125–129
- 9 Kaplan R, Friston K J. Planning and navigation as active inference. *Biological Cybernetics*, 2018, **112**(4): 323–343
- 10 Destexhe A, Rudolph-Lilith M. *Neuronal Noise*. New York: Springer, 2012, 1–2
- 11 Arleo A, Smeraldi F, Gerstner W. Cognitive navigation based on nonuniform gabor space sampling, unsupervised growing networks, and Reinforcement learning. *IEEE Transactions on Neural Networks*, 2004, **15**(3): 639–652
- 12 Strosslin T, Sheynikhovich D, Chavarriaga, Gerstner W. Robust self-localisation and navigation based on hippocampal place cells. *Neural Networks*, 2012, **18**(9): 1125–1140
- 13 Forster D J, Morris R G, Dayan P. A model of hippocampally dependent navigation, using the temporal difference learning rule. *Hippocampus*, 2000, **10**(1): 1–16
- 14 Tolman E C, Honzik C H. “Insight” in rats. *University of California Publications in Psychology*, 1931, **4**: 215–232
- 15 Edvardsen V, Bicanski A, Burgess N. Navigating with grid and place cells in cluttered environments. *Hippocampus*, 2019: 1–13
- 16 Erdem U M, Hasselmo M. A goal-directed spatial navigation model using forward trajectory planning based on grid cells. *European Journal of Neuroscience*, 2012, **35**: 916–931
- 17 Gonner L, Vitay J, Hamker F H. Predictive place-cell sequences for goal-finding emerge from goal memory and the cognitive map: a computational model. *Frontiers in Computational Neuroscience*, 2017, **11**(84): 1–19
- 18 Tejera G, Llofriu M, Barrera A, Weitzenfeld A. Bio-inspired robotics: A spatial cognition model integrating place cells, grid cells and head direction cells. *Journal of Intelligent and Robotic Systems*, 2018, **91**(3): 85–99
- 19 Bicanski A, Burgess N. A neural-level model of spatial memory and imagery. *eLife*, 2018, **7**: e33752
- 20 Ponulak F, Hopfield J J. Rapid, parallel path planning by propagating wavefronts of spiking neural activity. *Frontiers in Computational Neuroscience*, 2013, **7**(98): 1–14
- 21 Bi G Q, Poo M M. Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience*, 1998, **18**: 10464–10472
- 22 Kang L, DeWeese M R. Replay as wavefronts and theta sequences as bump oscillations in a grid cell attractor network. *eLife*, 2019, **8**: e46351
- 23 Ellender T J, Nissen W, Colgin L L, Mann E O, Paulsen O. Priming of hippocampal population bursts by individual perisomatic-targeting interneurons. *Journal of Neuroscience*, 2010, **30**(17): 5979–5991
- 24 Zennir M, Benmohammed M, Martinez D. Robust path planning by propagating rhythmic spiking activity in a hippocampal network model. *Biologically Inspired Cognitive Architectures*, 2017, **20**: 47–58
- 25 Khajeh-Alijani A, Urbanczik R, Senn W. Scale-free navigational planning by neuronal traveling waves. *Public Library of Science One*, 2015, **10**(7): 1–15
- 26 Palmer J, Keane A, Gong P. Learning and executing goal directed choices by internally generated sequences in spiking neural circuits. *Public Library of Science Computational Biology*, 2017, **13**(7): e1005669
- 27 Hok V, Save E, Lenck-Santini P P, Poucet B. Coding for spatial goals in the prelimbic/infralimbic area of the rat frontal cortex. *Proceedings of the National Academy of Sciences*, 2005, **102**(12): 4602–4607
- 28 Preston A R, Eichenbaum H. Interplay of hippocampus and prefrontal cortex in memory. *Current Biology*, 2013, **23**(17): 764–773
- 29 O’Keefe J. Place units in the hippocampus of the freely moving rat. *Experimental Neurology*, 1976, **51**(1): 78–109
- 30 Eichenbaum H. Memory on Time. *Trends in Cognitive Sciences*, 2013, **17**: 81–88
- 31 Ramakrishnan A, Byun Y W, Rand K, Pedersen C E, Levedev M A, Nicolelis M A. Interplay of hippocampus and prefrontal cortex in memory. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, **114**(24): 4841–4850
- 32 Dorst L, Trovato K. Optimal path planning by cost wave propagation in metric configuration space. In: SPIE Advances in Intelligent Robotics Systems, Cambridge, USA: SPIE, 1989. 186–197.
- 33 Tolman E C. Cognitive maps in rats and men. *The Psychological Review*, 1948, **55**(4): 189–208
- 34 Morris R, Garrud P, Rawlins J, O’Keefe J. Place navigation impaired in rats with hippocampal lesions. *Nature*, 1982, **297**: 681–683
- 35 Dechter R, Pearl J. Generalized best-first search strategies and the optimality of A*. *Journal of the ACM*, 1985, **32**(3): 505–536
- 36 Kang Liang, Zhao Chun-Xia, Guo Jian-Hui. Improved path planning based on rapidly exploring random tree for mobile robot in unknown environment. *Pattern Recognition and Artificial Intelligence*, 2009, **22**(3): 337–343
(康亮, 赵春霞, 郭剑辉. 未知环境下改进的基于 RRT 算法的移动机器人路径规划. 模式识别与人工智能, 2009, **22**(3): 337–343)
- 37 Bu Xin-Ping, Su Hu, Zou Wei, Wang Peng, Zhou Hai. Smooth path planning based on non-uniformly modeling and cubic bezier curves. *Acta Automatica Sinica*, 2017, **43**(5): 710–724
(卜新苹, 苏虎, 邹伟, 王鹏, 周海. 基于非均匀环境建模与三阶 Bezier 曲线的平滑路径规划. 自动化学报, 2017, **43**(5): 710–724)



武悦 北京工业大学信息学部硕士研究生. 2017 年获得西南交通大学学士学位. 主要研究方向为环境认知建模与类脑计算.

E-mail: wuy50271@gmail.com

(WU Yue Master student at the Faculty of Information Technology,

Beijing University of Technology. He received his bachelor degree from Southwest Jiaotong University in 2017. His research interest covers environment cognition modelling and brain-inspired computing.)



阮晓钢 北京工业大学信息学部教授. 1992 年获得浙江大学博士学位. 主要研究方向为自动控制, 人工智能与智能机器人.

E-mail: adrxg@bjut.edu.cn

(RUAN Xiao-Gang Professor at the Faculty of Information Techno-

logy, Beijing University of Technology. He received his Ph. D. degree from Zhejiang University in 1992. His research interest covers automatic control, artificial intelligence, and intelligent robot.)



黄静 北京工业大学信息学部副教授. 2016 年获得北京工业大学控制科学与工程博士学位. 主要研究方向为认知机器人与机器学习. 本文通信作者.

E-mail: huangjing@bjut.edu.cn

(HUANG Jing Associate professor at the Faculty of Information Techno-

logy, Beijing University of Technology. She received her Ph. D. degree in control science and engineering from Beijing University of Technology in 2016. Her research interest covers cognitive robotics, machine learning and industrial big data. Corresponding author of this paper.)



柴洁 北京工业大学信息学部博士研究生. 主要研究方向为认知学习和认知导航.

E-mail: chajie@emails.bjut.edu.cn

(CHAI Jie Ph. D. candidate at the Faculty of Information Technology, Beijing University of Technology.

Her research interest covers cognitive learning and cognitive navigation.)