记忆神经网络在机器人导航领域的应用与研究进展

王作为， 徐征， 张汝波， 洪才森， 王殊

引用本文:
王作为, 徐征, 张汝波, 等. 记忆神经网络在机器人导航领域的应用与研究进展[J]. 智能系统学报, 2020, 15(5): 835–846.
WANG Zuowei, XU Zheng, ZHANG Rubo, et al. Research progress and application of memory neural network in robot navigation[J]. *CAAI Transactions on Intelligent Systems*, 2020, 15(5): 835–846.

在线阅读 View online: https://dx.doi.org/10.11992/tis.202002020

DOI: 10.11992/tis.202002020
网络出版地址: http://kns.cnki.net/kcms/detail/23.1538.tp.20200413.1849.002.html

记忆神经网络在机器人导航领域的应用与研究进展

王作为[1,2], 徐征[3,4], 张汝波[5], 洪才森[1], 王殊[1]
（1. 天津工业大学 计算机科学与技术学院, 天津 300387; 2. 天津工业大学 机械工程学院博士后工作站, 天津 300387; 3. 天津动核芯科技有限公司, 天津 300350; 4. 天津职业技术师范大学 汽车与交通学院, 天津 300222; 5. 大连民族大学 机电工程学院, 辽宁 大连 116600）

摘　要: 记忆神经网络非常适合解决时间序列决策问题, 将其用于机器人导航领域是非常有前景的新兴研究领域。本文主要讨论记忆神经网络在机器人导航领域的研究进展。给出几种基本记忆神经网络结合导航任务的工作机理, 总结了不同模型的优缺点; 对记忆神经网络在导航领域的研究进展进行简要综述; 进一步介绍导航验证环境的发展; 最后梳理了记忆神经网络在导航问题所面临的复杂性挑战, 并预测了记忆神经网络在导航领域未来的发展方向。

关键词: 记忆神经网络; 机器人导航; 深度强化学习; 可微神经计算机; 可微神经字典; 深度学习; 强化学习; 记忆网络
中图分类号: TP183　　**文献标志码:** A　　**文章编号:** 1673−4785(2020)05−0835−12

中文引用格式: 王作为, 徐征, 张汝波, 等. 记忆神经网络在机器人导航领域的应用与研究进展 [J]. 智能系统学报, 2020, 15(5): 835–846.
英文引用格式: WANG Zuowei, XU Zheng, ZHANG Rubo, et al. Research progress and application of memory neural network in robot navigation[J]. CAAI transactions on intelligent systems, 2020, 15(5): 835–846.

Research progress and application of memory neural network in robot navigation

WANG Zuowei[1,2], XU Zheng[3,4], ZHANG Rubo[5], HONG Caisen[1], WANG Shu[1]
(1. School of Computer Science and Technology, Tianjin Polytechnic University, Tianjin 300387, China; 2. College of Mechanical Engineering Post-doctoral Research Station, Tianjin Polytechnic University, Tianjin 300387, China; 3. DongHexin Technology Co., Ltd., Tianjin 300350, China; 4. College of Automobile and Transportation, Tianjin University of Technology and Education, Tianjin 300222, China; 5. College of Mechanical and Electrical Engineering, Dalian Minzu University, Dalian 116600, China)

Abstract: Memory networks are a relatively new class of models designed to alleviate the problem of learning long-term dependencies in sequential data, by providing an explicit memory representation for each token in the sequence, and they can be used for learning navigation policies in an unstructured terrain, which is a complex task. Memory neural networks are highly suitable for solving time series decision-making problems, and their application in robot navigation is a very promising and emerging research field. The research progress of memory neural networks in the field of robot navigation is primarily discussed in this paper. First, the working mechanism of several basic memory neural networks used for robot navigationis introduced, and the advantages and disadvantages of different models are summarized. Then, the research progress of memory neural network in navigation field is briefly reviewed, and the development of navigation verification environment is discussed. Finally, the complex challenges faced by memory neural networks in navigation are summarized, and the future development of memory neural networks in navigation field is predicted.

Keywords: memory neural network; robot navigation; deep reinforcement learning; differentiable neural computer; differentiable neural dictionary; deep learning; reinforcement learning; memory networks

收稿日期: 2020−02−27.　**网络出版日期:** 2020−04−14.
基金项目: 国家自然科学基金面上项目 (61972456); 天津市教委科研计划项目（2019KJ018）; 天津工业大学学位与研究生教育改革项目（Y20180104）.
通信作者: 王作为. E-mail: wangzuowei@126.com.

自主机器人导航所在环境一般是未知的、动态的、部分可观测的。自主机器人导航需要具备以下能力: 探索未知环境、构建地图、目标导航的

路径规划与执行、环境变化的适应性。

传统的导航方法基于全局定位与地图构建(simultaneous localization and mapping, SLAM), SLAM 由于定位飘移、传感器噪音、环境改变以及有限的计算规划能力使得该方法很难推广到实际应用[1]。近年来, 由于神经网络的强大的表征能力, 尤其是强化学习与深度神经网络的结合使得深度强化学习 (deep reinforcement learning, DRL) 广泛应用到机器人导航领域[2-5]。然而 DRL 基于当前感知做出决策, 很难具有泛化性和推理能力, 并且很难应用于部分观测环境中。递归神经网络 (recurrent neural network,RNN) 和长短时记忆神经网络 (long short-term memory, LSTM) 与 DRL 相结合在机器人导航领域虽然取得了一定进展[6-7], 然而隐藏节点和权重所能记住的数据十分有限, 且只能记住一些有一定内在规律和特征的信息, 对于长程记忆则显得无能为力。

为了解决神经网络长程记忆的问题, 近 3 年涌现出了各种的记忆神经网络 (memory neural networks, MNN) 模型, MNN 采用外部记忆矩阵实现, 将记忆与计算分离开来, 采用可微的读写机制访问外部记忆网络, 整个系统可微, 允许端对端的训练。MNN 与 DRL 结合非常适合解决时间序列决策问题, 将其用于导航领域是非常有前景的新兴研究领域[8-9]。

1　MNN 结合导航任务的工作机理

近年来将 MNN 用于导航领域主要有 3 种神经网络模型: MemNN(memory networks)、DNC(differentiable neural computer) 以及 DND(differentiable neural dictionary), 下面分别介绍将其用于导航领域的工作机理。

1.1　MemNN 在导航中的应用

Sukhbaatar 等[10] 首先提出 MemNN, 这是一种无写操作的记忆结构, 记忆存储是固定的。网络学到的内容是如何从固定记忆池中去访问和读取信息, 而不是如何去改写内容。该模型被广泛用于情感分析[11]、对话训练[12] 等领域。Oh 等[13] 首次将 MemNN 与 DRL 相结合并在三维 Minecraft 环境中实现导航任务。相继提出了: 记忆 Q 网络 (memory Q-network, MQN)、循环记忆 Q 网络 (recurrent memory Q-network, RMQN) 以及反馈循环记忆 Q 网络 (feedback recurrent memory Q-network, FRMQN), 如图 1 所示。

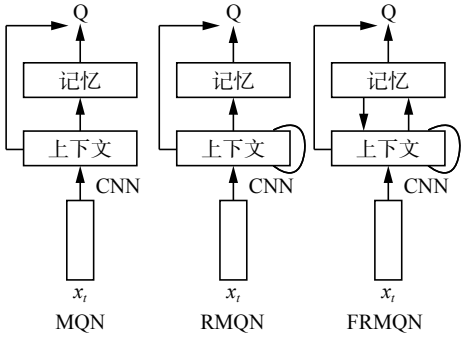

图 1　MemNN+DRL 结构

Fig. 1　MemNN+DRL structure

MemNN+DRL 结构采用一个递归控制器 DRL 与外部记忆 MemNN 进行交互, 基于时间上下文实现寻址机制, 有效处理了部分观测、长时依赖导航策略以及相似地图的知识迁移问题。MemNN+DRL 的导航工作机理如下: 将机器人最近遇到的 M 步观察经过编码写入到 MemNN 中, 相当于 M 步的情节记忆, 采用强化学习算法端对端训练参数, 最终获得导航能力。读写机制如图 2 所示。

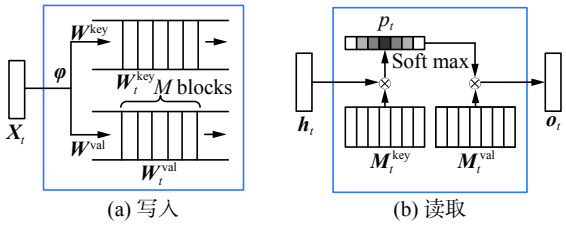

图 2　MemNN 读写机制

Fig. 2　Read and write mechanism of MemNN

具体模块功能如下。

1) 编码模块

将原始的图像信息提取出高层特征信息。将一个 c 通道的 $h\times w$ 维的图像 X_t 编码成一个 e 维特征向量 $\boldsymbol{e}_t$:

$$\boldsymbol{e}_t = \boldsymbol{\varphi}^{\text{enc}}(X_t) \tag{1}$$

2) 写记忆操作

将最近 M 步的观察实现矩阵转换, 以键记忆模块和值记忆模块形式分别存储到记忆中, 如式 (2)、(3):

$$\boldsymbol{M}_t^{\text{key}} = \boldsymbol{W}^{\text{key}}\boldsymbol{E}_t \tag{2}$$

$$\boldsymbol{M}_t^{\text{val}} = \boldsymbol{W}^{\text{val}}\boldsymbol{E}_t \tag{3}$$

式中: 矩阵 $\boldsymbol{M}_t^{\text{key}}$ 和 $\boldsymbol{M}_t^{\text{val}}$ 分别代表了键记忆模块和值记忆模块; $\boldsymbol{W}^{\text{key}}$ 和 $\boldsymbol{W}^{\text{val}}$ 是相应的线性转移矩阵; $\boldsymbol{E}_t$ 是最近 M 次观察的特征向量序列。

3) 读记忆操作

机器人导航过程中, 根据最近几步的观测值计算上下文向量 $\boldsymbol{h}_t$, 然后通过计算上下文向量 $\boldsymbol{h}_t$ 和键记忆模块 $\boldsymbol{M}_t^{\text{key}}$ 之间的内积, 再归一化后得

到一组注意力权重 $p_{t,i}$，即实现了注意力机制 (attention mechanism)。通过这种软注意力机制，机器人可以找到与当前观测向量 $\boldsymbol{h}_t$ 相关的那一部分记忆，即环境中的定位过程。读操作的输出 $\boldsymbol{o}_t$ 利用注意力权重 P_t 和值记忆模块 $\boldsymbol{M}_t^{\text{val}}$ 求出线性累加和。其中注意力权重公式和输出式分别为

$$p_{t,i}=\frac{\exp\left(\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{M}_t^{\text{key}}[i]\right)}{\sum_{j=1}^{M}\exp\left(\boldsymbol{h}_t^{\mathrm{T}}\boldsymbol{M}_t^{\text{key}}[j]\right)} \tag{4}$$

$$\boldsymbol{o}_t=\boldsymbol{M}_t^{\text{val}}P_t \tag{5}$$

4) 注意力机制

注意力机制在文本识别、图像识别、问答系统、机器翻译中被广泛深入研究[14-18]。注意力机制由一个注意力权重表示，越大的权重代表对应位置 i 越重要。在导航问题中，注意力机制不仅关注当前观测值与记忆模块的匹配度，而且考虑之前几步观测序列与记忆模块的匹配度，因此是一种基于时间序列的注意力机制。在 FRMQN 中，采用 LSTM 结构的注意力机制，如式 (6) 所示：

$$[h_t,c_t]=\text{LSTM}([e_t,o_{t-1}],h_{t-1},c_{t-1}) \tag{6}$$

其中，上一步召回的记忆 o_{t-1} 作为 LSTM 输入的一部分，这允许 FRMQN 不仅根据当前的观测序列还根据之前检索到的记忆来实现多级推理过程，这与 MemNN 中的多级跳结构非常类似[19]。

5) 预测行为值函数

记忆模块的输出是 $\boldsymbol{o}_t$，它表示了概率统计上的机器人的记忆模块和当前上下文输入最相关的特征向量，利用这个特征向量 $\boldsymbol{o}_t$ 产生相应的 Q 值输出，实现动作选择。这里的 q_t 是一个估计的状态行为值函数，用 MLP 多重前向网络实现。如式 (7) 所示：

$$\begin{aligned}\boldsymbol{g}_t&=f(\boldsymbol{W}^h\boldsymbol{h}_t+\boldsymbol{o}_t)\\ \boldsymbol{q}_t&=\boldsymbol{W}^q\boldsymbol{g}_t\end{aligned} \tag{7}$$

式中：$\boldsymbol{W}^h$、$\boldsymbol{W}^q$ 是其权值，最后一层利用 softmax 作为输出。利用目标值函数和当前值函数的误差来训练整个模型，整个过程数据流通非常平滑，全程可微，此模型可以利用误差反向传播进行训练，最终优化 $\boldsymbol{W}^{\text{key}}$、$\boldsymbol{W}^{\text{val}}$、$\boldsymbol{W}^h$、$\boldsymbol{W}^q$ 矩阵。

1.2　DNC 在导航中的应用

Google DeepMind 在 Nature 首次提出了 DNC 模型[20]。其强大的推理能力使其在自然语言理解、算法推理、视觉推理中被广泛深入研究[21-25]。DNC 具有递归神经控制器，可以通过执行可微的读操作和写操作去访问外部记忆资源。DNC 结构如图 3 所示。a 为递归控制器模块，其输入为外部输入向量和从记忆中读出 R 维向量，输出为外部输出向量和交互参量，这些交互参量用来确定读、写操作的参数。b 为多个读头和一个写头，用来实现对记忆的读写操作。c 为记忆模块，是一个 $N\times W$ 的记忆矩阵。d 是每个记忆位置的使用度向量，用来记录目前每个记忆位置的使用情况，其中时间链接矩阵记录了写入的顺序，时间顺序用箭头表示。Parisotto[26] 首次将 DNC 用于导航任务，将记忆模块看作神经地图 (neural map)，下面分别对每个模块进行说明。

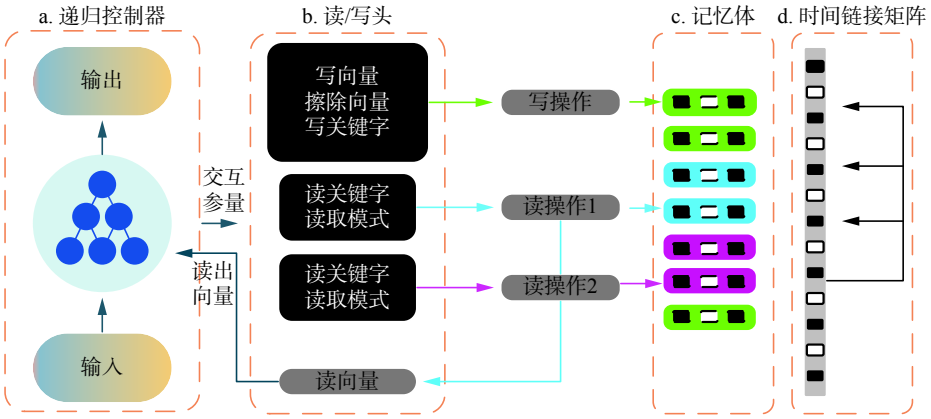

图 3　DNC 结构

Fig. 3　DNC structure

1) 递归控制器

每个时间步 t 控制器从环境接收当前感知向量 $\boldsymbol{s}_t$，机器人首先根据当前感知向量 $\boldsymbol{s}_t$ 和当前的全局读向量 $\boldsymbol{r}_t$ 产生一个上下文向量 $\boldsymbol{q}_t$，接着通过读头从上一时刻神经地图 $\boldsymbol{M}_{t-1}$ 中读取 R 维读入向量，机器人控制器根据当前输入向量 $\boldsymbol{s}_t$ 和 $\boldsymbol{r}_t$ 来输出向量 $\boldsymbol{c}_t$。$\boldsymbol{c}_t$ 用来得到策略输出 $\pi_t(a|s)$。另外，控制器也输出一个交互向量 $\boldsymbol{E}_t$，$\boldsymbol{E}_t$ 定义了当前时间步该如何与记忆交互。控制器可以采用任何神经网络结构实现，例如：CNN 结构、LSTM 结构或者

多级 LSTM 结构。

2) 读操作

上下文向量 $\boldsymbol{q}_t$ 基于当前输入 $\boldsymbol{s}_t$ 和 $\boldsymbol{r}_t$ 得到，利用上下文向量 $\boldsymbol{q}_t$ 和地图 $\boldsymbol{M}_t$ 中的每一个位置特征 $\boldsymbol{M}_t^{(x,y)}$ 做内积得到一个得分 $a_t^{(x,y)}$。得分正则化处理后得到在地图上所有位置的一个概率分布，即实现了软注意力机制。这个概率分布用来计算在所有位置特征 $\boldsymbol{M}_t^{(x,y)}$ 上的一个加权平均和 $\boldsymbol{c}_t$。这里读操作将神经地图看做联想记忆：机器人提供了一些不完全的信息 $\boldsymbol{q}_t$，读操作将返回一个与 $\boldsymbol{q}_t$ 最匹配的完整的记忆信息，类似于机器人可以回忆起当前的观察与记忆中的某些路标相似的东西。注意力权重公式和输出公式如式(8)所示，其中 $\boldsymbol{W}$ 是权重矩阵：

$$\begin{aligned} &\boldsymbol{q}_t = \boldsymbol{W}[\boldsymbol{s}_t, \boldsymbol{r}_t], a_t^{(x,y)} = \boldsymbol{q}_t \cdot \boldsymbol{M}_t^{(x,y)}, \\ &\boldsymbol{\alpha}_t^{(x,y)} = \frac{e^{a_t^{(x,y)}}}{\sum_{(w,z)} e^{a_t^{(w,z)}}} \\ &\boldsymbol{c}_t = \sum_{(x,y)} \boldsymbol{\alpha}_t^{(x,y)} \boldsymbol{M}_t^{(x,y)} \end{aligned} \tag{8}$$

3) 写操作

给定机器人当前时刻 t 的位置 (x_t,y_t)，写操作的输入为：当前感知向量 $\boldsymbol{s}_t$，全局读向量 $\boldsymbol{r}_t$，上下文读向量 $\boldsymbol{c}_t$，和当前的神经地图中 (x_t,y_t) 的特征向量 $\boldsymbol{M}_t^{(x_t,y_t)}$，通过一个深度神经网络 f_w 产生一个新的 c 维向量 $\boldsymbol{w}_{t+1}^{(x_t,y_t)} = f_w([\boldsymbol{s}_t, \boldsymbol{r}_t, \boldsymbol{c}_t, \boldsymbol{M}_t^{(x_t,y_t)}])$。这个向量作为新的局部 (x_t,y_t) 写候选向量。

写操作利用新的特征向量 $\boldsymbol{w}_{t+1}^{(x_t,y_t)}$ 替换机器人神经地图中 (x_t,y_t) 位置的特征向量，这是一种强写入机制。写操作修改了 t+1 时刻的神经地图 M_{t+1}。M_{t+1} 除了位置 (x_t,y_t) 上的特征信息有所改变，其余与旧神经地图一致，这是一个局部写入操作，如式(9)所示：

$$M_{t+1}^{(a,b)} = \begin{cases} w_{t+1}^{(x_t,y_t)}, & (a,b) = (x_t,y_t) \\ M_t^{(a,b)}, & (a,b) \neq (x_t,y_t) \end{cases} \tag{9}$$

4) 注意力机制

DNC 构建了 3 种注意力机制：基于内容的注意力机制、时间机制和动态记忆分配机制。其中基于内容寻址和动态记忆分配的方式决定写入记忆的位置；基于内容寻址和时间链接矩阵决定读出记忆位置。注意力机制由交互向量参数 $\boldsymbol{E}_t$ 决定。

实验在三维 ViZDoom 环境下验证，对于更加复杂的迷宫环境，其长时记忆能力、泛化性能力均优于 FRMQN。这是由于环境越来越大，越来越复杂，需要记忆的知识越来越多，MemNN 记忆结构只能记忆 M 步历史，而 DNC 可以记忆整个地图，并且可以根据环境改变动态修改地图，因此具有更好的不同环境间的知识迁移能力及适应动态环境的能力。然而 DNC 学到的参数较多，除了学习控制器网络参数外，还要学习读写操作的交互参数 $\boldsymbol{E}_t$。

1.3 DND 在导航中的应用

Pritzel 等[27]提出了神经情节控制模型 (neural episodic control, NEC)，用于实现机器人导航。作者指出当前的深度强化学习模型存在共同的弊端：所有深度强化学习模型 (包括 MemNN、DNC) 都是参数化模型，需要采用随机梯度下降法学习参数矩阵，如果参数矩阵较多，收敛速度缓慢，尤其是导航领域存在稀疏回报问题，整个过程很难收敛。而强化学习算法本身，尤其是 Q 学习是通过值迭代学习最优策略，而表格形式是最适合强化学习的知识表示形式。如果能将缓慢更新的状态表征用深度网络表示，将迅速更新的值函数用表格的形式表示，则更为有效。因此提出了一种无参数的记忆机构——可微神经字典 (DND)。类似于 Key-Value 记忆模型，将参数表示的键 (状态 S) 与表格表示的值 (行为值函数 V) 相结合，并在机器人选择行为期间使用基于上下文的软注意力机制来检索有用的值函数。允许自由读写，并且采用了追加写操作，使得写操作更加简单。每个行为 a 都有对应的 DND 记忆模块，学习采用 N 步 Q 强化学习算法，同时采用了类似于 DQN 中的回放机制。

1)NEC 结构

NEC 结构如图 4 所示。该结构分成 3 个部分：卷积神经网络；一系列 DND 记忆模块 (即行为记忆模块 M_a)；以及一个最终的网络，该网络将动作记忆模块的读出转换成 Q 值。卷积神经网络将视觉感知 s_t 转换成关键字 h_t。每个行为对应一个行为记忆模块 M_a，每个行为记忆模块 M_a 由键记忆模块 h_i 和值记忆模块 Q_i 组成。记忆模块从关键字 h_i 映射到值 Q_i 是一个联想关系，与数据字典类似，根据当前关键字 h_t 在记忆模块 M_a 中读出相应的值 Q_t，记忆模块 M_a 的输出即为对应行为 a 的 $Q(s,a)$ 值，不同的记忆 M_a 共享相同的卷积网络。机器人根据最高的 Q 值估计来决定在每一步中执行哪个动作，然后根据 N 步 Q 学习更新值函数和相应的权值。

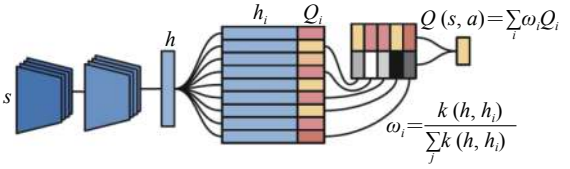

图 4 NEC 结构

Fig. 4 NEC structure

2) 读操作

读操作就是在 DND 上将当前关键字 h 映射为输出值 $Q(s,a)$, 如式(10)所示:

$$Q(s,a)=\sum_i w_i Q_i, w_i=\frac{k(h,h_i)}{\sum_j k(h,h_j)} \tag{10}$$

这里 h_i 是键记忆模块的第 i 个元素, Q_i 是值记忆模块的第 i 个元素。$K(x,y)$ 是一个相似度函数。因此 DND 的读操作相当于在记忆中搜索与 h 最匹配的那些记忆, 输出是记忆中对应 Q_i 值的加权和, 这是一种基于内容的注意力机制, 没有考虑时间相关性。从大容量的记忆里读取采用最近邻方法 (k-d 树, 详情介绍见文献 [28])。

3) 写操作

查找结束后, 将一个新的键−值对写入记忆。写入的过程是一个追加 (append-only) 写操作, 即将键−值对分别写入到键记忆模块和值记忆模块的末尾, 无需计算写入位置, 简化写入操作。如果键已经存在记忆中, 则对应的值函数 Q_i 根据 N 步 Q 学习更新, 写入操作如式(11):

$$\begin{aligned} &Q_i \leftarrow Q_i+\alpha(Q^{(N)}(s,a)-Q_i)\\ &Q^{(N)}(s_t,a)=\sum_{j=0}^{N-1}\gamma^j r_{t+j}+\gamma^N \max_{a'} Q(s_{t+N},a') \end{aligned} \tag{11}$$

这里的写操作类似于 Q 表更新, 只不过这里的 Q 表示随着时间动态增长的。学习率 α 设置较大, 类似于快门式学习, 学习过程不涉及参数调整, 是一种无参数记忆结构。

4) 参数更新

类似于 DQN 的回放机制, 将每次的转移实例 (s_t, a_t, r_t) 存储在回放缓冲区中, 其中 $Q^{(N)}(s,a)$ 作为目标函数。从回放缓冲区中随机取出的小批量样本用于反向误差更新, 这里的神经网络参数的更新率较小。因此是一种缓慢更新的卷积网络和迅速更新的值函数相结合的结构, 该模型大大提高了数据有效性、提高收敛速度。

该方法类似于基于实例的学习, 在 Atari 游戏中验证, 在数据有效性和收敛速度方面, 优于 DQN、A3C、Prioritised DQN 算法。

1.4　不同记忆神经网络的优缺点

3 种记忆结构都采用了软关注度机制, 利用 DRL 实现误差反向传播, 整个过程均是可微的、端对端的结构。用于部分可观测导航任务均取得了优于 LSTM+DRL 的效果。笔者分析了不同记忆神经网络的写操作、读操作、注意力机制、存储知识、训练参数、记忆结构, 以及将其应用于导航领域的各自优缺点, 如表 1 所示。从表 1 可以看出, MemNN 与 DND 存储知识是情节记忆, 即存储了大量的经验序列, 而 DNC 存储的是真正的空间地图。在训练时间上, DNC 训练参数最多, 训练时间长, 因此将其用于导航领域常常出现不收敛的问题; 而 DND 训练参数少, 训练时间快, 与基于实例的机器学习类似。

表 1　不同记忆结构的对比

Table 1　Comparisons of different memory structures

记忆神经网络	写操作	读操作	注意力机制	存储知识	训练参数	记忆结构	解决难题	存在问题
MemNN	固定写入 M步观察	根据内积运算求得相似度	基于时间上下文注意力机制	情节记忆	矩阵参数 $\boldsymbol{W}_{\text{key}}$、$\boldsymbol{W}_{\text{val}}$、$\boldsymbol{W}_{\text{h}}$、$\boldsymbol{W}_{\text{q}}$以及神经网络参数	Key-Value 结构	部分观测、长时记忆、相似地图的迁移学习	难以适应动态环境
DNC	适应性写操作、局部写操作、软写入机制	根据内积运算求得相似度	基于内容的注意力机制、时间链接注意力机制、动态记忆分配机制	空间地图	控制器网络参数、决定读写的交互参数$\boldsymbol{E}_{\text{t}}$	神经网络控制器+矩阵记忆结构	部分观测、长时记忆、不同地图间的迁移学习、动态环境的适应性	参数收敛慢
DND	简单追加(append-only)写操作	相似度函数+基于k-d树最近邻方法	基于内容的注意力机制	逐渐增加的情节记忆	快门式学习无参数记忆、卷积神经网络参数	卷积神经网络+无参数记忆结构	数据有效性、提高收敛速度、部分观测、延迟回报	记忆空间大、如何压缩记忆

2　MNN 在导航领域的研究进展

MNN 的飞速发展也就是近三年的事情，这些记忆结构大部分应用在自然语言处理、问题回答系统、视觉推理等领域，机器人导航领域没有得到广泛关注。在有限的一些文献中，主要分成以下几个改进方向。

2.1　关注度机制的改进

Neural Slam[29] 将 SLAM 与 DNC 深入结合，将 SLAM 中的运动预测和定位嵌入到软注意力寻址机制中，实现有偏的读写操作，DNC 作为环境地图的表示，整个过程采用深度强化学习 A3C 实现，是一个端对端的训练模型，Neural Slam 模型如图 5 所示。

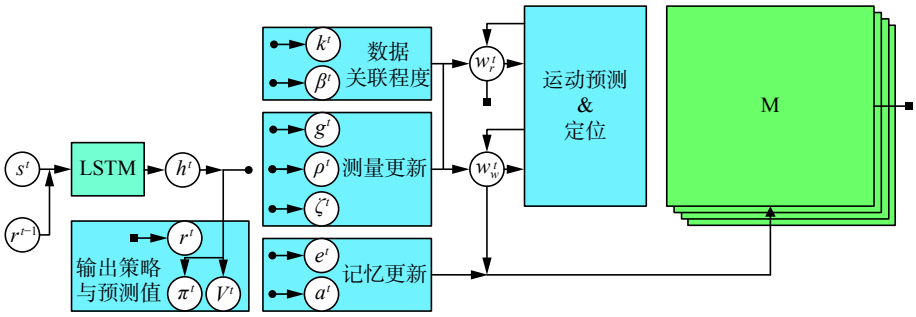

图 5　Neural Slam 结构

Fig. 5　Neural Slam structure

在每个时间步中，将输入直接提供给 LSTM 单元，它给出一个隐藏状态 h^t。使用这个隐藏状态 h^t 来发出一组交互参量，根据这些交互参量 k^t、β^t、g^t、ρ^t、ζ^t、e^t、a^t 由读头、写头计算其读权重 w_r^t 和写权重 w_w^t，这里与 Parisotto 等[26] 所提出的神经地图的区别是：神经地图中的位置信息 (x,y) 事先已知，而 Neural Slam 利用 SLAM 计算其位置的信念值。

该方法优点是将 SLAM 与 DNC 很好地融合，改进了 DNC 的软注意力机制，使得机器人不断更新其位置信念。缺陷是输入只是激光测距信息，没有高维视觉信息，构建的是一个度量地图。

2.2　写入机制的改进

如前所述，神经地图的主要缺点是机器人时刻知道自己的绝对位置，并且其写入机制是一种强写入机制 (只要重新写入，之前的信息就被替代)，难以实现长期信息的维护。因此 Emilio Parisotto 在进一步的研究工作中[26, 30]，将 DNC 看做一个 2 维空间地图，采用了基于 GRU 的写操作和自我为中心的神经地图 (ego neural map) 的模型，采用 A2C 算法学习。在更复杂的 3 维 ViZ-Doom 环境中验证，性能优于传统的 Neural Map 方法。

1) 软写入机制

写操作利用新的特征向量替换记忆中当前位置的向量，这是一种强写入机制，强写入机制不保留之前的记忆内容。文献 [31] 提出基于 GRU 的写入机制。GRU 写操作在递归神经网络中有着较长的研究历史，GRU 写操作比强写入机制具有更长久保持记忆的能力。

2) 主动神经定位

文献 [30] 进一步对绝对位置进行改进，提出了一种“主动神经定位器”，它是一种完全可微的神经网络，能够准确有效地进行定位。该模型融合了传统的基于滤波的定位方法的思想，利用具有乘法交互的状态结构化信念来传播信念，并将其与策略模型相结合，利用最少的步骤精确地进行定位。采用端到端强化学习的方法对主动神经定位器进行训练。

2.3　与 VIN 的融合

传统深度强化学习系统缺乏明确的规划计算。Tamar 等[32] 提出了值迭代网络 (value iteration networks, VIN)，这是一个嵌入了“规划模块”的完全可微的神经网络。方法的巧妙之处是观察到经典的值迭代 (VI) 规划算法可以由特定类型的 CNN 表示，通过在标准的前馈网络中嵌入 VI 网络模块，使得策略训练起来很简单，VIN 策略可以更好地泛化到新的、不可见的环境。但是该方法由于没有记忆模块，因此无法适应部分可观测环境。下面是将 VIN 与 MNN 相结合进行改进。

1) CMP

Gupta 等[33] 将地图构建和 VIN 模块结合，设计了一个 CMP(cognitive mapping and planning) 结构用来实现部分观测环境下的导航任务，采用模仿学习 DAGGER 算法实现真实室内场景下的导航，性能优于 LSTM+DRL 模型，CMP 结构如图 6 所示。

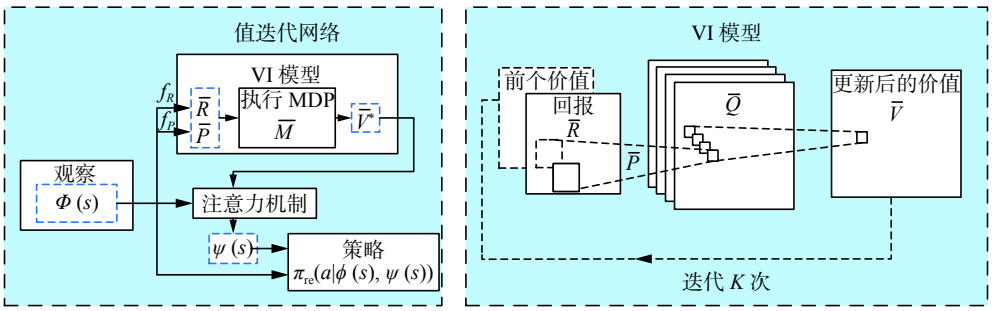

图 6　CMP 结构

Fig. 6　CMP structure

图 6 中模型的主要改进之处如下：

地图构建利用机器人的观察值得到，生成一个以自我为中心的多尺度信念地图。地图是一个二维的空间记忆结构，将一个三维环境投射到二维栅格环境中去。信念更新方式是训练一个卷积神经网络根据观察到的第一人称视图来预测更新。

规划器利用自我为中心的多尺度信念地图和目标位置来规划当前动作。规划器采用 VI 模型，使用一个可训练、可微的分层的值迭代网络。

2) MACN

Khan 等[34] 将 DNC 与 VIN 相结合应用到部分可观测环境下的导航问题，提出了一种记忆扩展控制网络 (memory augmented control network, MACN)。结构如图 7 所示，该方法并没有尝试将一个三维环境投射为二维栅格环境，而是直接计算环境的信念空间，并把这种信念值存入一个可微记忆 DNC 中，采用监督学习实现了连续控制的机器人在一个三维环境下的导航任务。

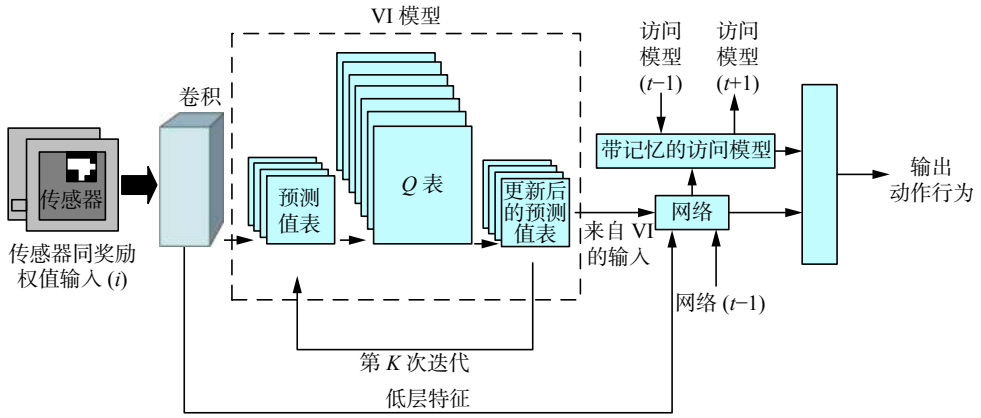

图 7　MACN 结构

Fig. 7　MACN structure

MACN 是利用 VI 模块来学习局部信念值，并将这种局部信念值存入一个可微记忆 DNC 中，DNC 描述了整个环境的信念空间。这恰恰是采用了分层强化学习中 option 的思想[35]，更适合高维度的状态空间和行为空间。

规划采用了分层的结构：低层采用 VIN 实现局部规划，高层利用 DNC 学习全局规划。低层规划模块利用丰富表征的特征信息计算局部环境的最优策略，高层规划将得到的局部策略和当前的稀疏表征作为输入，采用基于 DNC 的记忆模块，来产生一个全局环境的最优策略。

2.4　与基于模型的强化学习结合

基于模型的强化学习对于实现导航任务非常有效。然而生成时间模型 (generative temporal models, GTMs) 的构建在复杂的部分观测三维环境下是非常困难的。大多数 GTMs，例如隐马尔可夫模型[36] 和卡尔曼滤波器及其非线性扩展[37]，这些模型中使用的固定阶马尔可夫假设不足以描述实际系统的特性。递归神经网络比固定阶马尔科夫假设约束的模型具有显著的优势，最近的 GTMs，例如变分递归神经网络[38] 和深度卡尔曼滤波器[39] 都是建立在递归神经网络之上，原则上这些递归神经网络可以解决变阶马尔科夫问题。然而由于其参数太多使得实际应用起来效率极低。Gemici 等[40] 将记忆神经网络与生成时间模型相结合，提出了带记忆的时间生成模型 (GTMMs)，该模型实现了三维环境的感知建模，但没有实现导航任务。Fraccaro 等[41] 将生成时间模型与

DND 相结合，由于在部分观测的三维环境中学习生成时间模型非常困难，因此提出一个动作条件生成模型 (action-conditioned generative model) 来对环境建模，在二维和三维环境中实现上百步的一致预测。为了解决部分可观测问题，DeepMind 团队[42]引入了一种新的模型——外部记忆、RL 和状态推断网络相结合 (MERLIN)，MERLIN 结构如图 8 所示。

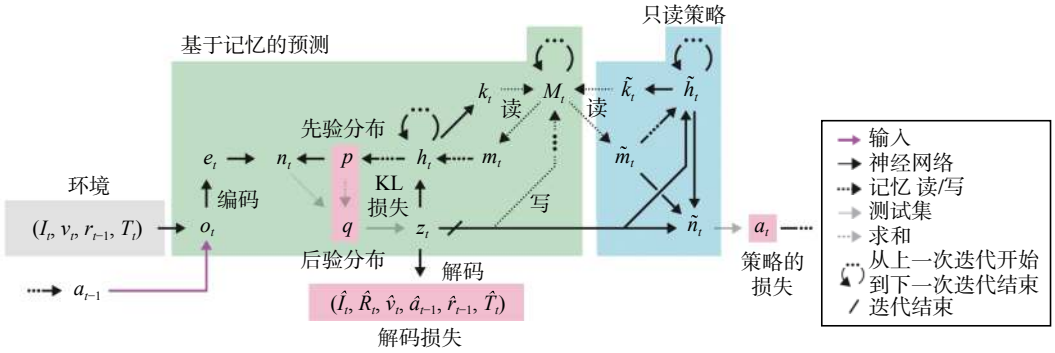

图 8　MERLIN 结构

Fig. 8　MERLIN structure

创新之处是提出基于记忆的预测器 (memory-based predictor, MBP)。MBP 是一个无监督模型。MBP 的输入来自于多模态信息 (例如图像信息 I_t，速度信息 v_t，回报值 r_{t-1}，行为 a_{t-1} 以及文本命令 T_t)，下一个状态根据记忆中保存的之前的状态变量和行为来预测。另一种概率分布，即后验概率，根据新的观测值修正了这一先验，从而形成对状态变量更好地估计。在 MERLIN 中，策略模块对记忆模块只能进行只读访问。MERLIN 在部分观测三维环境中验证，机器人快速地建立一个地图的近似模型，从这个模型中它可以快速导航回目标点。

综上所述，MNN 在导航领域的研究进展如图 9 所示。

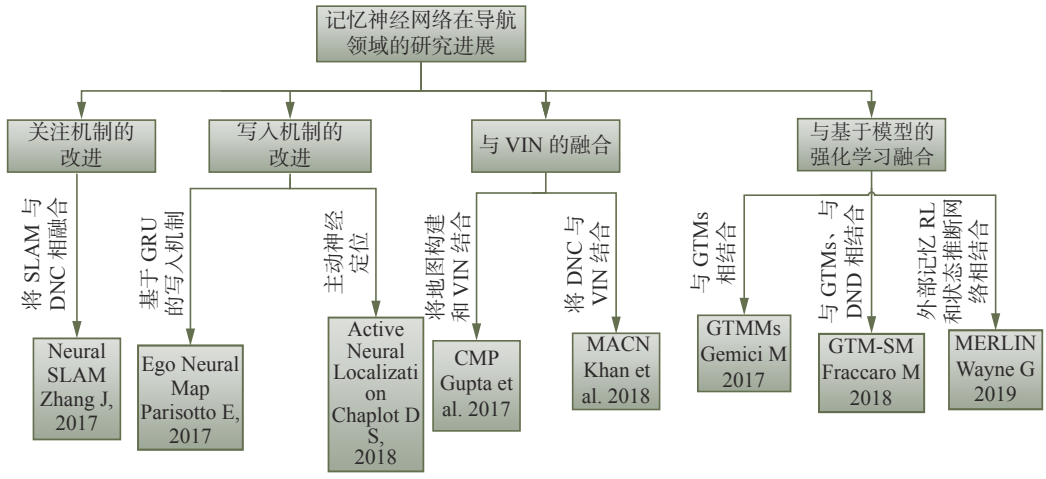

图 9　MNN 在导航领域研究进展

Fig. 9　Diagram of MNN's progress in the field of navigation

3　导航验证环境的发展

为了弥补仿真器和真实场景之间的鸿沟，许多研究者提出了大量的真实、高仿真度的仿真环境，可以进一步将算法移植到真实环境中去。常用的三维导航仿真环境有 DeepMind Lab (Beattie et al 2016)[43]、Malmo (Johnson et al 2016)[44]、ViZDoom(M. Kempka, 2016)[45]，然而这些仿真环境存在一个主要的问题：没有真实场景图片。

文献 [46] 提出了一个三维房屋模拟器 House3D，建立在 SUNCG[47] 的基础上，该数据集包含数千个不同的合成室内场景，配有各种对象和布局，它的视觉多样性和丰富的内容为研究强化学习智能体的语义泛化开辟了道路。另外，HoME[48] 和 MINOS[49] 也提供了合成的大范围室内环境，但是均没有提供与环境的交互。文献 [50-51] 构建了一个可交互的三维高仿真环境 AI2-THOR，它由逼真的三维室内场景组成，其中

智能体可以在场景中导航并与对象交互来执行任务。该环境可以实现深度强化学习、模仿学习、交互学习、规划学习、视觉问题回答、无监督表示学习、对象检测和分割以及认知学习模型。Mirowski 等[52]提出了一种新的交互式环境“StreetLearn”, 从真实世界图片和谷歌街景中获得信息。笔者从以下几个方面总结了不同仿真环境的境的特点, 如表 2 所示。

表 2　不同仿真环境的特点

Table 2　Characteristics of different simulation environments

不同仿真环境	三维	大范围环境	用户可定制	真实场景图片	物理引擎	与对象交互
DeepMind Lab (Beattie et al., 2016)	√		√			
Malmo (Johnson et al., 2016)	√	√	√			
ViZDoom(M. Kempka, 2016)	√		√			
House3D	√	√	√	√		
HoME	√	√	√			
MINOS	√	√	√			
AI2-THOR(Eric Kolve, 2019)	√		√	√	√	√

4　结束语

综上所述, 不同记忆神经网络的发展为解决视觉导航任务提供了很多有效的模型, 但是基于记忆神经网络的导航领域仍然存在多个方面的问题, 例如: 部分观测、延迟回报、泛化性差、数据有效性有待提高、环境模型构建等问题。以后的发展主要集中在如何构建更有效的记忆结构, 实现更有效的学习; 如何与新发展起来的 DRL 方法相结合, 例如元强化学习、多目标强化学习等; 如何与概率统计模型相结合, 实现更有效的基于环境模型的视觉导航; 以及与图模型相结合, 发展出更有效的图记忆模型。下面主要给出 3 个最有前景的发展方向。

1) 基于神经动力学联想记忆

以上提到的 DNC、MemNN、DND 的研究都是基于外部记忆的。这种外部记忆的存储形式及读、写机制存在以下问题: ①控制器与外部记忆完全分离, 是一个不严格端对端的结构, 影响学习效果; ②记忆的读、写过程类似于 CPU 访问存储器的过程, 缺乏生物学的解释。

基于神经动力学的联想记忆网络是日益兴起的一个热点领域。这种基于神经动力学的联想记忆具有更好的生物学解释性, 联想记忆网络一般不受特定结构限制、可以实现增量的序列学习, 并且以一种自组织、无监督的形式。Danihelka 等将一个联想记忆模型作为部件引入 LSTM 网络中, 从而在不引入额外参数的情况下增加网络容量。Paris 提出自组织联想记忆网络模型, 并且将其用于人机交互、时空特征的学习等领域, 但是笔者尚未发现将其用于导航领域, 因此如何将联想记忆模型和导航领域结合是最新的研究热点。

2) 基于图网络的记忆结构

图网络 (graph network, GN) 是一种最新兴起的研究方向, 还没有比较成熟的网络模型。图网络是将消息传递的思想扩展到图结构上的神经网络。图中的每个节点都用一组神经元来表示其状态, 每个节点都可以收到相邻节点的消息, 并更新自己的状态。

应用到不同任务, 有不同的图网络结构。例如图卷积网络 (graph convolutional network, GCN)、消息传递网络 (message passing neural network, MPNN) 等。实际上, 导航任务中的记忆地图适合用这种图网络来表示, 因此将记忆网络扩展到图网络结构中, 并应用于导航领域也是一个非常有前景的研究方向。

3) 与概率图模型相结合

概率图模型和神经网络有着类似的网络结构, 但两者也有很大区别。概率图模型中节点是随机变量, 概率图的结构主要描述随机变量间的依赖关系, 一般是稀疏连接, 优点是可以有效地进行统计推断。而神经网络中的神经元是计算节点, 每个神经元没有直观解释。近些年来概率图模型和神经网络结合越来越紧密, 例如利用神经网络强大的表示能力来建模图模型中的推断问题 (变分编码器), 生成问题 (生成对抗网络) 等, 包括 2.4 节中提到的模型都是将概率图模型与神经记忆网络相融合用于实现导航任务。概率图模型

与记忆神经网络深度融合将是导航领域最有前景的研究方向之一，最有希望实现空间地图的建模，实现基于模型的强化学习。

参考文献：

[1] 刘强, 段富海, 桑勇. 复杂环境下视觉 SLAM 闭环检测方法综述 [J]. 机器人, 2019, 41(1): 112–123, 136.
LIU Qiang, DUAN Fuhai, SANG Yong. A survey of loop-closure detection method of visual SLAM in complex environments[J]. Robot, 2019, 41(1): 112–123, 136.

[2] KULKARNI T D, SAEEDI A, GAUTAM S, et al. Deep successor reinforcement learning[J]. arXiv preprint arXiv: 1606.02396v1, 2016.

[3] MNIH V, BADIA A P, MIRZA M, et al. Asynchronous methods for deep reinforcement learning[C]//Proceedings of the 33rd International Conference on International Conference on Machine Learning. New York, USA, 2016: 1928−1937.

[4] ZHU Yuke, MOTTAGHI R, KOLVE E, et al. Target-driven visual navigation in indoor scenes using deep reinforcement learning[C]//Proceedings of 2017 IEEE International Conference on Robotics and Automation (ICRA). Singapore, 2016.

[5] MIROWSKI P, PASCANU R, VIOLA F, et al. Learning to navigate in complex environments[C]//Proceedings of the 5th International Conference on Learning Representations. Toulon, France, 2017.

[6] JADERBERG M, MNIH V, CZARNECKI W M, et al. Reinforcement learning with unsupervised auxiliary tasks[C]//Proceedings of the 5th International Conference on Learning Representations. Toulon, France, 2016.

[7] HEESS N, HUNT J J, LILLICRAP T P, et al. Memory-based control with recurrent neural networks[C]//Proceedings of the Workshops of Advances in Neural Information Processing Systems. Montreal, Canada, 2015: 301−312.

[8] RAMANI D. A short survey on memory based reinforcement learning[J]. arXiv preprint arXiv:1904.06736v1, 2019.

[9] SAVINOV N, DOSOVITSKIY A, KOLTUN V. Semi-parametric topological memory for navigation[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[10] SUKHBAATAR A, WESTON J, FERGUS R, et al. End-to-end memory networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada, 2015: 2440−2448.

[11] ZHANG Lei, WANG Shuai, LIU Bing. Deep learning for sentiment analysis: A survey[J]. Wiley interdisciplinary reviews: data mining and knowledge discovery, 2018, 8(4): e1253.

[12] YOUNG T, HAZARIKA D, PORIA S, et al. Recent trends in deep learning based natural language processing[J]. IEEE computational intelligence magazine, 2018, 13(3): 55–75.

[13] OH J, CHOCKALINGAM V, SINGH S, et al. Control of memory, active perception, and action in minecraft[C]//Proceedings of the 33nd International Conference on Machine Learning. New York, USA, 2016: 2790−2799.

[14] BOTHE C, MAGG S, WEBER C, et al. Conversational analysis using utterance-level attention-based bidirectional recurrent neural networks[C]//Proceedings of the 19th Annual Conference of the International Speech Communication Association. Hyderabad, India, 2018.

[15] 张新生, 高腾. 多头注意力记忆网络的对象级情感分类 [J]. 模式识别与人工智能, 2019, 32(11): 997–1005.
ZHANG Xinsheng, GAO Teng. Aspect level sentiment classification with multiple-head attention memory network[J]. Pattern recognition and artificial intelligence, 2019, 32(11): 997–1005.

[16] BAHDANAU D, CHOROWSKI J, SERDYUK D, et al. End-to-end attention-based large vocabulary speech recognition[C]//Proceedings of 2016 IEEE International Conference on Acoustics, Speech and Signal Processing. Shanghai, China, 2016: 4945−4949.

[17] JETLEY S, LORD N A, LEE N, et al. Learn to pay attention[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[18] 梁天新, 杨小平, 王良, 等. 记忆神经网络的研究与发展 [J]. 软件学报, 2017, 28(11): 2905–2924.
LIANG Tianxin, YANG Xiaoping, WANG Liang, et al. Review on research and development of memory neural networks[J]. Journal of software, 2017, 28(11): 2905–2924.

[19] TANG Duyu, QIN Bing, LIU Ting. Aspect level sentiment classification with deep memory network[C]//Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. Austin, USA, 2016.

[20] GRAVES A, WAYNE G, REYNOLDS M, et al. Hybrid computing using a neural network with dynamic external memory[J]. Nature, 2016, 538(7626): 471–476.

[21] YANG Feng, ZHANG Shiyue, ZHANG Andi, et al. Memory-augmented neural machine translation[C]//Proceedings of the 2017 Conference on Empirical Methods in

Natural Language Processing. Copenhagen, Denmark, 2017.

[22] CINGILLIOGLU N, RUSSO A. DeepLogic: towards end-to-end differentiable logical reasoning[C]//Proceedings of the AAAI 2019 Spring Symposium on Combining Machine Learning with Knowledge Engineering. Palo Alto, USA, 2019.

[23] 廖祥文, 林威, 吴运兵, 等. 基于辅助记忆循环神经网络的视角级情感分析 [J]. 模式识别与人工智能, 2019, 32(11): 987–996.
LIAO Xiangwen, LIN Wei, WU Yunbing, et al. Aspect level sentiment analysis based on recurrent neural network with auxiliary memory[J]. Pattern recognition and artificial intelligence, 2019, 32(11): 987–996.

[24] DAI Hanjun, KHALIL E B, ZHANG Yuyu, et al. Learning combinatorial optimization algorithms over graphs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA, 2017: 6351−6361.

[25] PARISI G I, KEMKER R, PART J L, et al. Continual lifelong learning with neural networks: a review[J]. Neural networks, 2019, 113: 54–71.

[26] PARISOTTO E, SALAKHUTDINOV R. Neural map: Structured memory for deep reinforcement learning[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[27] PRITZEL A, URIA B, SRINIVASAN S, et al. Neural episodic control[C]//Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia, 2017: 2827−2836.

[28] BENTLEY J L. Multidimensional binary search trees used for associative searching[J]. Communications of the ACM, 1975, 18(9): 509–517.

[29] ZHANG Jingwei, TAI Lei, BOEDECKER J, et al. Neural SLAM: learning to explore with external memory[J]. arXiv:1706.09520v4, 2017.

[30] CHAPLOT D S, PARISOTTO E, SALAKHUTDINOV R. Active neural localization[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[31] CHUNG J, GÜLCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. arXiv:1412.3555v1, 2014.

[32] TAMAR A, WU Yi, THOMAS G, et al. Value iteration networks[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia, 2017: 4949−4953.

[33] GUPTA S, DAVIDSON J, LEVINE S, et al. Cognitive mapping and planning for visual navigation[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 7272−7281.

[34] KHAN A, ZHANG C, ATANASOV N, et al. Memory augmented control networks[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[35] SUTTON R S, PRECUP D, SINGH S. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning[J]. Artificial intelligence, 1999, 112(1/2): 181–211.

[36] RABINER L R. A tutorial on hidden markov models and selected applications in speech recognition[J]. Proceedings of the IEEE, 1989, 77(2): 257–286.

[37] KALMAN R E. A new approach to linear filtering and prediction problems[J]. Journal of basic engineering, 1960, 82(1): 35–45.

[38] CHUNG J, KASTNER K, DINH L, et al. A recurrent latent variable model for sequential data[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal, Canada, 2015: 2962−2970.

[39] KRISHNAN R G, SHALIT U, SONTAG D. Deep Kalman filters[J]. arXiv preprint arXiv:1511.05121v2, 2015.

[40] GEMICI M, HUNG C C, SANTORO A, et al. Generative temporal models with memory[J]. arXiv preprint arXiv:1702.04649v2, 2017.

[41] FRACCARO M, REZENDE D J, ZWOLS Y, et al. Generative temporal models with spatial memory for partially observed environments[C]//Proceedings of the 35th International Conference on Machine Learning. Stockholm, Sweden, 2018.

[42] WAYNE G, HUNG C C, AMOS D, et al. Unsupervised predictive memory in a goal-directed agent[J]. arXiv preprint arXiv:1803.10760v1, 2018.

[43] BEATTIE C, LEIBO J Z, TEPLYASHIN D, et al. Deepmind lab[J]. arXiv:1612.03801v2, 2016.

[44] JOHNSON M, HOFMANN K, HUTTON T, et al. The Malmo platform for artificial intelligence experimentation[C]//Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. New York, American, 2016.

[45] KEMPKA M, WYDMUCH M, RUNC G, et al. Vizdoom: A doom-based AI research platform for visual reinforcement learning[C]//Proceedings of IEEE Conference on Computational Intelligence and Games. Santorini, Greece, 2016.

[46] WU Yi, WU Yuxin, GKIOXARI G, et al. Building gener-

alizable agents with a realistic and rich 3D environment[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[47] SONG Shuran, YU F, ZENG A, et al. Semantic scene completion from a single depth image[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 190−198.

[48] BRODEUR S, PEREZ E, ANAND A, et al. Home: a household multimodal environment[C]//Proceedings of the 6th International Conference on Learning Representations. Vancouver, Canada, 2018.

[49] SAVVA M, CHANG A X, DOSOVITSKIY A, et al. MINOS: Multimodal indoor simulator for navigation in complex environments[J]. arXiv:1712.03931v1, 2017.

[50] KOLVE E, MOTTAGHI R, HAN W, et al. AI2-thor: An interactive 3D environment for visual AI[J]. /arXiv:1712.05474v3, 2019.

[51] YANG Wei, WANG Xiaolong, FARHADI A, et al. Visual semantic navigation using scene priors[C]//Proceedings of the 7th International Conference on Learning Representations. New Orleans, USA, 2019.

[52] MIROWSKI P, GRIMES M K, MALINOWSKI M, et al. Learning to navigate in cities without a map[C]// Proceedings of the 32nd International Conference on Neural Information Processing Systems. Montreal, Canada, 2018: 2424−2435.

作者简介：

王作为，副教授，主要研究方向为智能机器人与智能控制、机器学习与人工智能。主持省部级、局级基金项目 3 项。发表学术论文 20 余篇。

徐征，副教授，主要研究方向为电机控制与运动控制系统。获天津市科技进步二等奖 1 项。主持和参与省部级基金项目 5 项。发表学术论文 8 篇。

张汝波，教授，博士生导师，主要研究方向为智能机器人与智能控制、机器学习与计算智能、智能信息处理。主持完成国防 973、国家 863、国家自然科学基金项目、省自然科学基金项目和国防预研项目 20 余项，获国家科学技术进步二等奖 1 项、国防科学技术奖 3 项、中国船舶工业总公司科技进步奖 2 项。发表学术论文 200 余篇。