

DOI:10.3969/j.issn.1673-4785.201312035

网络出版地址: <http://www.cnki.net/kcms/doi/10.3969/j.issn.1673-4785.201312035.html>

# 一种基于情感智能的机器人自主趋光行为研究

阮晓钢<sup>1</sup>, 庞涛<sup>1,2</sup>, 张晓平<sup>1</sup>, 王尔申<sup>2</sup>

(1. 北京工业大学 电控学院, 北京 100124; 2. 沈阳航空航天大学 电信学院, 辽宁 沈阳 110136)

**摘要:**针对移动机器人的自主趋光行为问题,提出了一种基于情感智能的内发动机仿生学习机制。该学习机制以生物体感觉运动系统的学习机制为基础,包括评价环节、行为选择环节和取向环节,采用模糊神经网络构建情感模型,情感模型的输出作为评价环节的内部奖赏信号。该学习机制能够使机器人在未知环境下通过自主的学习和训练逐渐形成、发展和完善趋光行为技能,通过情感智能的作用可以增加试探成功次数和减小学习步数,仿真实验证明了该方法的有效性。

**关键词:**认知机器人学;内发动机;移动机器人;趋光技能;情感智能;感觉运动系统;模糊推理方法;取向性

**中图分类号:** TP242.6 **文献标志码:** A **文章编号:** 1673-4785(2015)01-0097-06

中文引用格式:阮晓钢,庞涛,张晓平,等.一种基于情感智能的机器人自主趋光行为研究[J].智能系统学报,2015,10(1):97-102.

英文引用格式:RUAN Xiaogang,PANG Tao,ZHANG Xiaoping,et al. Research on the autonomous phototaxis behavior of a robot based on emotion intelligence[J]. CAAI Transactions on Intelligent Systems, 2015, 10(1): 97-102.

## Research on the autonomous phototaxis behavior of a robot based on emotion intelligence

RUAN Xiaogang<sup>1</sup>, PANG Tao<sup>1,2</sup>, ZHANG Xiaoping<sup>1</sup>, WANG Ershen<sup>2</sup>

(1. College of Electronic and Control Engineering, Beijing University of Technology, Beijing 100124, China; 2. College of Electronic Information Engineering, Shenyang Aerospace University, Shenyang 110136, China)

**Abstract:** In this paper, a kind of bionic learning mechanism of intrinsic motivation on the basis of emotion intelligence is proposed for the autonomous phototaxis behavior of a mobile robot. The learning mechanism is on the basis of the learning mechanism of the sensorimotor system of a living body, including the links of evaluation, behavioral choices and orientation. The emotion model is constructed by the fuzzy neural network and the output of emotion model is taken as the internal reward signal of the evaluation link. The learning mechanism may make a robot gradually form, develop and perfect the phototaxis skills by autonomous learning and training in an unknown environment. The emotion intelligence may increase the number of exploratory successes and reduce the learning steps. The simulation results demonstrated the effectiveness of this method.

**Keywords:** cognitive robotics; intrinsic motivation; mobile robot; phototaxis skill; emotion intelligence; sensorimotor system; fuzzy reasoning method; orientation

内发动机机制是认知的动力和源泉,是机器人最基本的认知发育机制。通过内发动机机制的不断完善,使生物体逐渐形成、发展和完善控制技

能<sup>[1-2]</sup>。认知科学研究表明,情感智能是内发动机机制不可或缺的组成部分,在生物自主行为决策中起着至关重要的作用<sup>[1-2]</sup>。随着科学研究的不断深入,越来越多的研究者开始关注人类的情感智能,尝试采用多种方法将生物情感智能抽象化建模,指导机器人行为决策取得更佳的控制效果。2008年Ahn等<sup>[3]</sup>提出了基于情感和认知强化学习模型,考虑了来自情感的内在奖励和来自认知的外部奖励,

收稿日期:2013-12-22. 网络出版日期:2015-01-13.

基金项目:国家973计划资助项目(2012CB720000);国家自然科学基金资助项目(61375086; 61101161);北京市自然科学基金资助项目/北京市教育委员会科技计划重点资助项目(KZ201210005001).

通信作者:庞涛. E-mail: pangtao163@126.com.

研究表明将来自情感的内在奖励融入到学习和决策中,能提高学习速度和决策能力。2011年,Malfaz等<sup>[4]</sup>将害怕情感因数与Q学习结合起来,通过增加Q函数的权重,有效降低了危险的行为决策。2011年,Sequeira等<sup>[5]</sup>提出了基于情感内发动机机制的强化学习模型,将奖赏信号分为内部和外部2部分,细化为好奇心、动机、价值和控制部分,实验证明了智能体可以克服环境的缺点获得更好的决策。2013年,Abdi等<sup>[6]</sup>利用情感经验获得奖赏信号,利用TDQ学习算法,实现优化多智能体的行为决策。2003年刘明等<sup>[7]</sup>提出一种基于模糊逻辑的情感模型,将情感对环境和agent自身状态的评估的变化作为再励信号,用于引导agent的行为选择策略的学习。2008年,张惠娣等<sup>[8]</sup>将基于情感和认知的学习与决策模型引入到基于行为的移动机器人控制体系中,设计了一种自主导航控制系统,提高了基于行为的移动机器人在未知环境中的自主导航能力。2010年,胡云斗等<sup>[9]</sup>提出一种基于任务的机器人情感决策模型,建立从多种感知输入到多种行为输出的映射,以“福娃”机器人为平台验证了该方法的正确性和实用性。2011年,祝宇虹等<sup>[10]</sup>对害怕进行了人工情感建模,提出了带情感权重的Q学习加权策略,并将该策略应用于虚拟机器人的行为决策,取得了较高的成功率。目前,针对机器人的基于情感的内发动机仿生自主学习问题,很多学者做了大量的工作,但对未知环境的自主学习方法,仍需不断探索。

近年来,机器人的趋光特性作为认知领域的范例已经得到广泛重视,通过模拟生物的趋光行为,研究机器人内发动机机制的渐近过程,并把这些原理正确运用在机器人制造中,对认知科学的发展和應用起到了重要作用。针对机器人仿生自主学习问题,构建了基于情感因素的内发动机学习模型,包括评价、行为选择和取向环节,通过“感知—行动”的往复学习,使机器人逐渐形成、发展和完善趋光行为技能,采用模糊推理方法构建情感模型,将情感因素作为内发动机机制的内部奖赏信号,通过情感智能的作用,增加机器人趋光学习试探成功次数和减少学习步数。

### 1 机器人模型

采用三轮式圆盘状移动机器人模型,结构如图1所示,左右2个驱动轮用来控制速度与方向,后面方向轮用来保持机器人平衡。机器人圆盘前端有6个位置可以安装传感器,每个位置安装一个光敏传

感器和一个超声波测距传感器,将6个位置的传感器分为左右2组,分别测定左侧和右侧的光强度和障碍物距离,每组3个传感器中测定光强度最大值做为该组传感器测定的光强度值,左侧光强度值为 $o_l$ ,右侧光强度值为 $o_r$ 。每组3个传感器中测定与障碍物距离最小值做为该组传感器测定的与障碍物距离值,与左侧障碍物距离值为 $d_l$ ,与右侧障碍物距离值为 $d_r$ 。

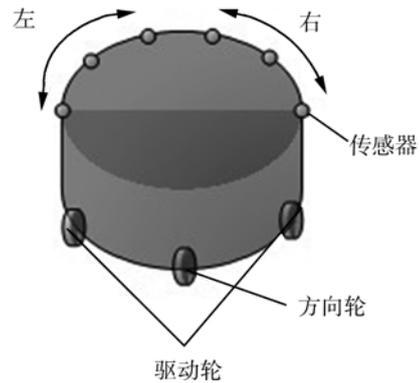


图 1 机器人结构

Fig.1 Robot structure

## 2 情感智能的机器人趋光仿生模型

### 2.1 感觉运动系统的内发动机机制

根据神经生理学<sup>[11-13]</sup>,生物体利用感受器、运动神经系统和效应器,通过自学习和自组织过程,逐渐完善运动技能的学习机制,如图2所示。感受器是感觉器官,其功能是感知环境(或客体)状态,运动神经系统是调节特定技能的神经细胞群,根据环境(或客体)状态产生动作决策,效应器是机体的运动器官,其功能是根据运动神经系统动作决策对环境(或客体)施加操作。

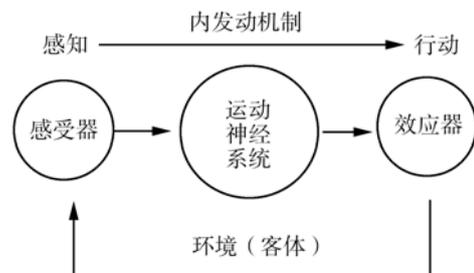


图 2 基于感觉运动系统的内发动机机制

Fig.2 Intrinsic motivation mechanism based on sensorimotor system

生物学领域的研究表明,在生物体感觉运动系统中存在动机与内在目标和目的相关联的机制<sup>[1]</sup>,即内发动机机制,这种机制是一种以感觉运动系统

为基础,受生存、取向性、好奇心以及情感和信念驱使的学习机制<sup>[14-15]</sup>。内发动机机制<sup>[1-2]</sup>是生物体在不需环境(或客体)数学模型的情况下,将环境(或客体)状态映射为对环境(或客体)的操作,并获得奖惩信号,成功的操作会获得奖赏信号,失败的操作则获得惩罚信号,生物体将奖惩信息转换成评价指标并生成取向信息,调整对环境的操作,这一过程在“感知-行动”环中周而复始地进行,生物体逐渐形成、发展和完善运动控制技能。其间情感在内发动机机制中起着重要的作用。

2.2 趋光情感模型

情感是生物体一种特殊意识形式,是生物的认知过程中行为决策的关键因素,对情感进行科学分析和精确计算,就能够建立情感的数学模型。目前,以心理学、生物学和人工智能等学科为基础,研究人员采用多种方法构建情感模型。文中针对机器人模拟趋光行为问题,采用模糊推理方法构建了人工情感模型<sup>[16-17]</sup>,模糊逻辑更能体现情感状态推理的复杂性和不确定性的本质特征。人工情感模糊推理模型,如图 3 所示。

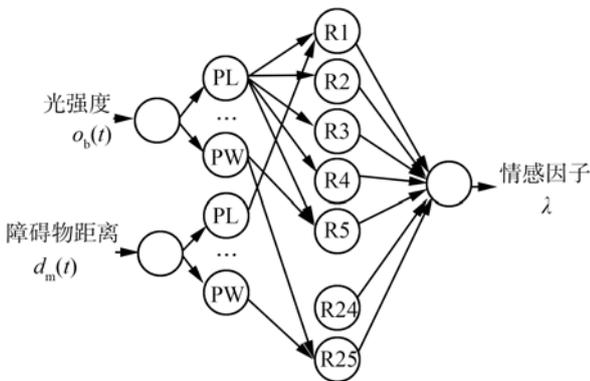
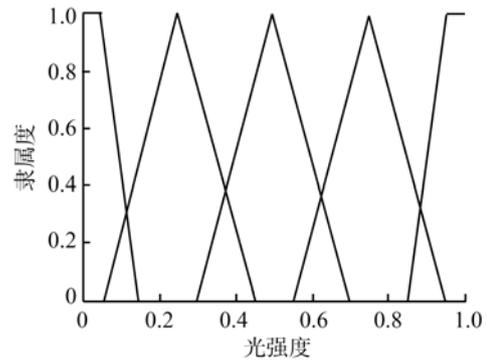


图 3 基于模糊推理的情感模型

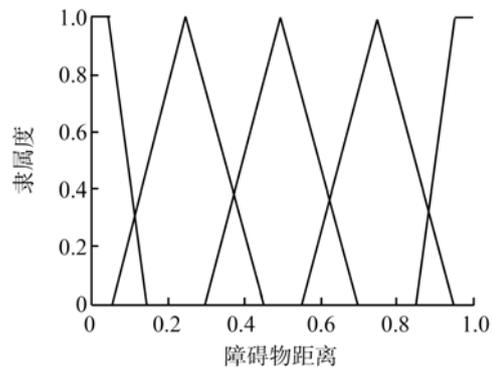
Fig.3 The emotion model based on fuzzy reasoning

图 3 中输入是机器人所在位置光强度  $o_b(t)$  和机器人与障碍物距离  $d_m(t)$ ,各划分为 5 个等级,隶属度函数如图 4 所示,机器人所在位置光强度  $o_b(t)$  采用很高(PL)、高(PB)、中(PM)、低(PS)和很低(PW)5 个等级,机器人与障碍物距离  $d_m(t)$ ,采用很远(PL)、远(PB)、中(PM)、近(PS)和很近(PW)5 个等级。人工情感模糊推理模型输出是机器人在趋光过程中产生的情感因子  $\lambda$ 。趋光过程中情感状态的设定,根据 OCC 情感模型<sup>[18]</sup>中的 22 种情感状态,根据趋光具体问题,考虑趋光过程中机器人操作结果对机器人情感的影响,选取 6 类情感  $E = \{Pride, Joy, Satisfaction, Disappointment, distress, Shame\}$ ,Pride 为骄傲情感, Joy 为高兴情感, Satisfaction 为满意情感, Disappointment 为失望情感,

distress 为忧伤情感, Shame 为羞愧情感。人工情感模型中模糊规则设计依据 OCC 模型中情感产生规则确定情感状态,如表 1, OCC 模型中情感产生规则依据机器人对趋光操作结果的期望度确定,趋光操作结果包括趋光操作后机器人与光源距离  $o_b(t)$  和机器人与障碍物距离  $d_m(t)$ ,机器人每操作一次,需要重新确定情感状态。



(a) 光强度  $o_b(t)$  隶属度函数



(b) 与障碍物距离  $d_m(t)$  隶属度函数

图 4  $o_b(t)$  和  $d_m(t)$  的隶属度函数

Fig.4 Membership function of  $o_b(t)$  and  $d_m(t)$

表 1 OCC 模型情感产生规则

Table 1 Emotion generation rule of OCC model

Emotion	Rule
Pride	The action is done by the agent and is approved by standards
Joy	The occurrence of a desirable event
Satisfaction	The occurrence of a confirmed desirable event
Disappointment	The occurrence of a unconfirmed desirable event
Distress	The occurrence of a undesirable event
Shame	The action is done by the agent and is unapproved by standards

模糊规则共 25 条,设计如下:

- 1) 如果光强度是很高,且障碍物距离是很远,则骄傲;
- 2) 如果光强度是很高,且障碍物距离是远,则

高兴;

3) 如果光强度是很高,且障碍物距离是中,则满意;

4) 如果光强度是很高,且障碍物距离是近,则不满意;

5) 如果光强度是很高,且障碍物距离是很近,则忧伤;

6) 如果光强度是高,且障碍物距离是很远,则高兴;

7) 如果光强度是高,且障碍物距离是远,则高兴;

8) 如果光强度是高,且障碍物距离是中,则满意;

……

文中采用了重心模糊推理方法得出精确化后的输出情感因子  $\lambda$ , 公式为

$$\lambda = \frac{\sum_{i=1}^9 \eta_i \lambda_i}{\sum_{i=1}^9 \eta_i} \quad (1)$$

$\eta_i$  是通过代数乘积方法得到的每一条规则真实值。

### 2.3 基于情感智能的机器人趋光仿生模型设计

依据感觉运动系统的内发动机机制<sup>[2,19-20]</sup>, 设计了包含 3 个模块的机器人模型, 如图 5 所示, 分别是评价模块、行为选择模块、取向模块。评价环节根据输入状态  $s(t)$ 、奖惩信号  $r(t)$  和动作量  $a(t)$ , 对动作量  $a(t)$  进行评价输出评价值  $V_e(t)$ , 评价函数为

$$V_e(t) = \sum_{j=1}^{\infty} \gamma^{j-1} r(j) = r(t) + \gamma V_e(t+1) \quad (2)$$

式中:  $\gamma$  是折扣因子, 表示近期和远期回报预测的重要程度。奖赏信号  $r(t)$  包括 2 部分: 环境给予的外部奖赏信号  $r_e(t)$  和情感智能给予的内部奖赏信号  $r_c(t)$ , 情感模型输出的情感因子  $\lambda$  ( $0 < \lambda < 1$ ) 作为内部奖赏信号  $r_c(t)$ 。评价网络采用前馈神经网络构成, 采用时间差分函数对评价函数进行逼近。取向环节根据评价值  $V_e(t)$  生成取向信息, 判定该动作在以后出现时选择的概率。行为选择环节根据状态量  $s(t)$ 、评价值  $V_e(t)$  和取向信息输出动作量  $a(t)$ 。行为选择环节采用前馈神经网络实现, 应满足以较大概率从行为集合中选择具有最大行为值函数的元素。机器人对环境(或客体)实施操作, 得到奖惩信号  $r(t+1)$ , 环境并转变为新的状态, 然后重复感知到行动这一过程, 最终完成机器人的自主趋光学习。仿生模型见图 5。

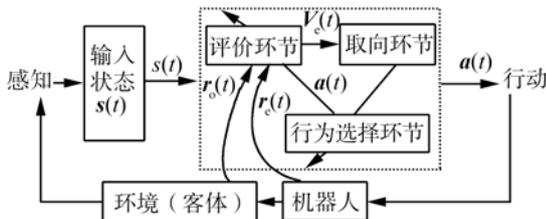


图 5 基于情感智能的机器人趋光仿生模型

Fig.5 The robot phototaxis biomimetic model based on emotion intelligence

### 2.4 算法流程

移动机器人基于内发动机自主学习算法步骤为

1) 初始化行为选择和评价环节中各初始值, 及各神经元连接权值; 初始化机器人的各状态变量, 并归一化;

2) 根据环境状态量  $s(t)$ , 计算奖赏信号  $r_e(t)$ 、 $r_c(t)$ ;

3) 行为选择环节根据状态量  $s(t)$  计算动作量  $a(t)$ ;

4) 评价环节根据状态量  $s(t)$  和动作量  $a(t)$  对动作量评价, 输出评价值  $V_e(t)$ ;

5) 取向环节根据评价值  $V_e(t)$ , 计算取向信息, 利用取向信息, 调整动作量  $a(t+1)$  输出;

6) 将动作量  $a(t+1)$  对环境实施操作, 环境产生下一时刻的状态量  $s(t+1)$ ;

7) 返回 2), 直到产生的各状态量满足智能体的期望性能指标, 则执行 8);

8) 行为决策模块对动作量加权并对客体实施操作, 客体产生下一时刻的状态量;

9) 结束。

## 3 仿真实验

### 3.1 趋光实验

为验证算法的有效性, 采用 MATLAB 仿真软件, 仿真环境中圆形表示光源位置, 方形表示障碍物, 环境中分布着有限个静态障碍物, 以机器人的趋光学习为目标进行研究, 光源位置和移动机器人起点位置可以任意设定, 移动机器人运动起点位置注明“S”, 终点位置注明“G”, 设定光源所在位置光强是 1 坎德拉 (Candela), 与光源越远的位置光强越弱, 按 e 指数规律减弱。机器人对环境是完全未知的, 可以自定位, 采用在线学习方式。机器人匀速移动, 设定线速度  $v$  为 0.5 m/s, 采样时间是  $T=1$  s。移动机器人进行试探学习, 每次学习步数不超过 1 000 步, 每次学习移动机器人从起始位置出发, 到达光源附近位置停止, 学习结束, 准备进行下一次学习, 如果学习步数超过 1 000 步仍未到达光源附近位置, 则终止这次学习, 进行下一次学习。趋光学习评价环节趋光输入域是  $x(t) = [o_l(t), o_r(t), d_l(t), d_r(t), \theta(t)]^T$ , 其中  $\theta(t)$  为机器人当前方向与目标点的夹角。趋光学习行为选择环节输入域是  $s(t) = [o_l(t), o_r(t), d_l(t), d_r(t)]^T$ 。

图 6 是机器人在有障碍物的环境下趋光学习情况。图 6(a) 是学习初期试探结果, 机器人从起始位置出发并不向光源方向运动, 说明机器人原本不具备趋光能力。图 6(b) 是学习中期试探结果, 机器人起始位置出发到达接近光源附近位置停止, 但并未始终按光源所在方向移动, 说明通过学习并未完全掌握趋光能力。6(c) 是机器人从起始位置出发到达接近光源附近位置停止, 始终按光源所在方向移

动,说明已经掌握趋光能力。图 6(d)是移动机器人在趋光移动中感知光强的变化曲线,机器人在起点位置时,光强较弱,随着与光源距离的接近,光强逐渐增加,最后到达光源附近,感知的光强最强。

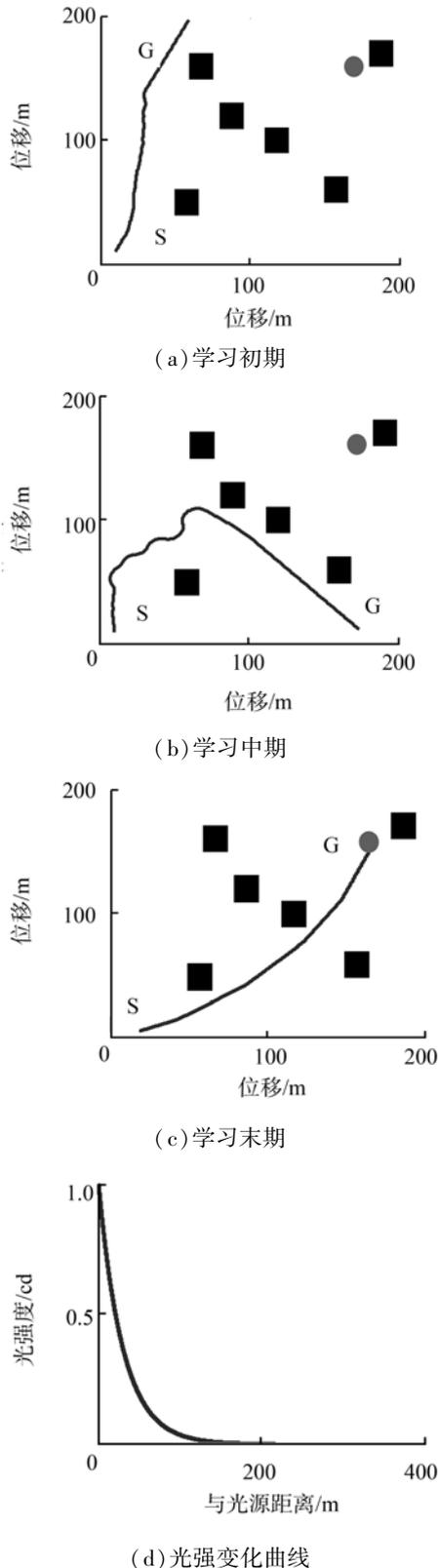


图 6 机器人趋光学习过程曲线

Fig.6 The robot curve of phototactic learning

图 7 是机器人在趋光学习过程中的相邻动作评价价值变化的收敛曲线。从曲线上可以看出,在学习初期、中期和末期,曲线波动幅度由大到小,说明机器人最初缺乏先验知识,学习随机性较大,动态性能较差,随着学习的进行,随机性减小,动态性能有所提高,机器人自组织、自学习的能力逐步完善,机器人逐渐学会了趋光移动技能,可以较好地模拟生物的趋光行为。将学习实验进行 100 次,统计学习成功的仿真实验步数并计算平均值,机器人平均需要 420 步的学习,逐渐学会了趋光行为。将该方法与没有情感智能作用的方法进行对比实验,在相同实验条件下,没有情感智能作用时,平均需要 550 步的学习。说明情感因数在机器人趋光行为决策中可以减少学习步数。

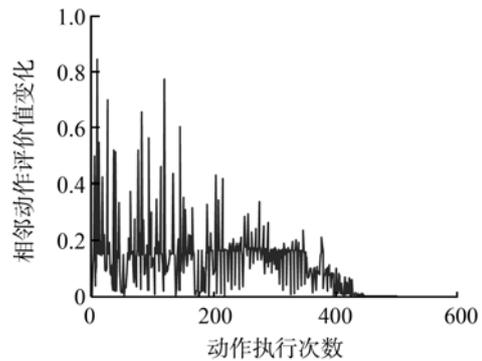


图 7 机器人趋光仿真收敛曲线

Fig.7 The robot convergence curve of phototaxis simulation

### 3.2 趋光对比实验

为了进一步验证算法的有效性,在相同实验条件下,将有无情感智能的 2 种情况做对比实验,改变实验参数折扣因子  $\gamma$ , 2 种方法分别进行 100 次试探学习,分别统计两者学习的试探成功次数。通过实验数据结果可以看出,在折扣因子选取不同的数值时,有情感智能作用下的试探成功次数多,如图 8 所示。说明在机器人的趋光学习过程中,情感智能的引入能够提高机器人学习的试探成功率。

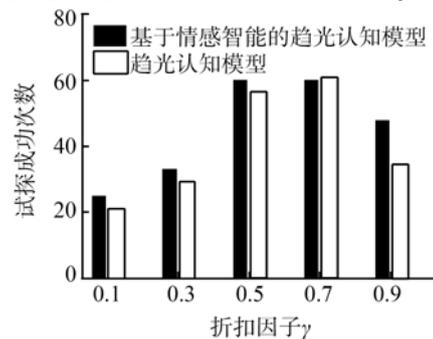


图 8 机器人趋光试探成功次数对比实验

Fig.8 The comparative experiments of robot phototaxis training success rate

## 4 结束语

通过对基于情感的内发动机机器人自主趋光行为的研究,使机器人通过自主学习逐渐形成、发展并完善了趋光能力。通过模拟生物的情感智能,采用模糊推理方法设计情感模型,提高了机器人趋光学习的试探成功率,减少了学习步数,实验表明情感智能可以增强机器人的行为决策能力。

## 参考文献:

- [1] OUDEYER P, KAPLAN F. What is intrinsic motivation? a typology of computational approaches[J]. *Frontiers in Neurobotics*, 2007, 1(6): 1-14.
- [2] SINGH S. Intrinsically motivated reinforcement learning: an evolutionary perspective[J]. *IEEE Transactions on Autonomous Mental Development*, 2010, 2(2): 70-82.
- [3] AHN H, PICARD R W. Affective cognitive learning and decision making: the role of emotions [C]//The 18th European Meeting on Cybernetics and Systems Research. [s.l.], 2006: 1-6.
- [4] MALFAZ M, SALICHS M A. Learning to avoid risky actions [J]. *Cybernetics and Systems*, 2011, 42(8): 636-658.
- [5] SEQUEIRA P, MELO F S, PAIVA A. Emotion-based intrinsic motivation for reinforcement learning agents[C]//Affective Computing and Intelligent Interaction. Berlin: Springer, 2011: 326-336.
- [6] ABDI J, MOSHIRI B, ABDULHAI B, et al. Short-term traffic flow forecasting: parametric and nonparametric approaches via emotional temporal difference learning [J]. *Neural Computing and Applications*, 2013, 23: 141-159.
- [7] 刘明, 许力. 人工情感在 Agent 行为选择策略中的应用 [J]. *江南大学学报*, 2003, 2(6): 564-568.  
LIU Ming, XU Li. Study on artificial emotion and its application in the behavior selection strategy of agents [J]. *Journal of Southern Yangtze University*, 2003, 2(6): 564-568.
- [8] 张惠娣, 刘士荣. 基于情感与环境认知的移动机器人自主导航控制 [J]. *控制理论与应用*, 2008, 25(6): 995-1000.  
ZHANG Huidi, LIU Shirong. Autonomous navigation control for mobile robots based on emotion and environment cognition [J]. *Control Theory & Applications*, 2008, 25(6): 995-1000.
- [9] 胡云斗, 胡丹丹, 高庆吉. 一种基于任务的机器人情感决策模型的构建 [J]. *机器人*, 2010, 32(4): 464-477.  
HU Yundou, HU Dandan, GAO Qingji. Construction of a robot emotion decision-making model based on tasks [J]. *Robot*, 2010, 32(4): 464-477.
- [10] 祝宇虹, 毛俊鑫. 基于人工情感与 Q 学习的机器人行为决策 [J]. *智能工程*, 2008, 61(7): 61-65.  
ZHU Yuhong, MAO Junxin. Robot's behavioral decision-making strategy based on artificial emotion and Q-learning [J]. *Intelligent Engineering*, 2008, 61(7): 61-65.
- [11] ASADA M, UCHIBE E, HOSODA K. Cooperative behavior acquisition for mobile robots in dynamically changing real worlds via vision-based reinforcement learning and development [J]. *Artificial Intelligence*, 1999, 110(2): 275-292.
- [12] HUMBERT J S, HYSLOP A M. Bioinspired visuomotor convergence [J]. *IEEE Transactions on Robotics*, 2010, 26: 121-30.
- [13] ZOLLO L, ESKIZMIRLILER S, TETI G, et al. An anthropomorphic robotic platform for progressive and adaptive sensorimotor learning [J]. *Advanced Robotics*, 2008, 22: 91-118.
- [14] BARANES A, OUDEYER P Y. Active learning of inverse models with intrinsically motivated goal exploration in robots [J]. *Robotics and Autonomous Systems*, 2013, 61(1): 49-73.
- [15] VIGORITO C M, BARTO A G. Intrinsically motivated hierarchical skill learning in structured environments [J]. *IEEE Transactions on Autonomous Mental Development*, 2010, 2(2): 132-143.
- [16] EI-NASR M S, YEN J. Agents, emotion intelligence and fuzzy logic [C]//Proceedings of the North American Conference on the Fuzzy Information Processing Society. Pensacola Beach, 1998, 1: 135-140.
- [17] DEGAONKAR V N, APTE S D. Emotion modeling from speech signal based on wavelet packet transform [J]. *International Journal of Speech Technology*, 2013, 16(1): 1-5.
- [18] ORTONY A. The cognitive structure of emotions [M]. Cambridge: Cambridge University Press, 1990: 35-77.
- [19] OUDEYER P Y, KAPLAN F, HAFNER V V. Intrinsic motivation systems for autonomous mental development [J]. *IEEE Transactions on Evolutionary Computation*, 2007, 11(2): 265-286.
- [20] ZORAN M, MARKO M, MIHAILO L, et al. Neural network reinforcement learning for visual control of robot manipulators [J]. *Expert Systems with Applications*, 2013, 40(5): 1721-1736.

### 作者简介:



阮晓钢,男,1958年生,博士生导师,主要研究方向为机器人学、自动控制与人工智能技术。先后主持和承担国家自然科学基金重点项目、国家 973 计划重点项目以及省部级项目 30 余项,发表学术论文 200 余篇。



庞涛,女,1976年生,讲师,主要研究方向为认知机器人学。



张晓平,女,1991年生,博士研究生,主要研究方向为机器人学、人工智能技术。