

一种基于情感智能的无人机自主决策方法

任天助¹, 周锐¹, 李浩²

(1. 北京航空航天大学自动化科学与电气工程学院, 北京 100191; 2. 中航工业成都飞机设计研究所, 成都 610091)

摘要: 无人机在各领域的不断深化应用使得人们对无人机的自主性提出了更高的要求。通过模拟人的情感决策过程建立一种基于情感智能的决策方法进行人工情感与智能决策设计, 并将其应用于无人机航路规划与威胁规避的决策问题。仿真结果表明, 基于情感智能的决策方法在应用中具有较传统方法更优的效果。

关键词: 无人机; 情感智能; 智能决策

中图分类号: V279 **文献标志码:** A **文章编号:** 1671-637X(2017)03-0011-05

An Emotional Intelligence Based Autonomous Decision-Making Method for UAVs

REN Tian-zhu¹, ZHOU Rui¹, LI Hao²

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;
2. Chengdu Aircraft Design Research Institute of AVIC, Chengdu 610091, China)

Abstract: The continuous deepening of the application of UAVs in all fields raises higher requirements to the autonomy of the UAVs. Through the simulation to human emotion decision-making process, we established a decision-making method based on emotional intelligence, and used it in the decision-making problems of UAV path planning and threat avoidance. The simulation results show that the method of decision-making based on emotional intelligence has better effect than the traditional method.

Key words: UAV; emotion intelligence; intelligent decision

0 引言

近年来, 军用无人机在现代战争中崭露头角, 成为许多军事强国的重要装备; 民用无人机在航拍、物流等领域亦有了越来越多的应用, 逐渐给人们生活带来重大的影响。然而当前无论是军用固定翼无人机还是民用多旋翼无人机, 其自主性都很低^[1]。无人机对于复杂环境下的任务, 往往难以做到自主决策, 还需要人的介入来完成。现实中的任务实时性很高, 而通讯链路不能保证每时每刻都可靠, 智能自主决策的缺失对无人机的进一步发展形成制约。从目前的研究和技术水平来看, 真正实现非结构化环境下无人机的自主飞行控制还是一项具有挑战性的技术难题。从人类的认知决策过程环 (Observation-Oriented-Decision-Action, OODA) 来看^[2],

实现无人机自主飞行的关键是实现以“态势感知为中心”而不是“以人为中心”的智能决策与管理, 也就是说, 让无人机根据其自身外部信息进行决策, 而不是过多的人为干预。

研究智能体决策的一个重要方向就是模拟人的认知-决策-行为过程, 建立仿人的智能决策结构, 实现无人机的智能决策。根据当代认知心理学研究的成果, 人在做决策的过程中, 一部分依赖于对过往知识的学习水平, 另一部分依赖于当前的情感状态^[3]。因此把对知识学习与情感相结合运用于智能体决策是一种有效的方案。其中, 研究的热点集中在探讨智能体 (Agent) 在外界激励情况下的情感改变和行为特征。文献[4]提出一种嵌入式情感 Agent 结构, 论述了大脑的情感产生机制, 通过距离费用函数量化情感的影响, 并使用复杂度和准确性不同的两种变量刻画外界激励; 同时, 文献[5]论述了一种 Greta's mind, 通过 Agent 的 BDI 模型 (目标、信念、事件) 刻画面部表情的情感表达; 国内的研究中, 文献[6-7]通过建立交互式的情感决策模型得到一套多属性决策方法, 并在工

收稿日期: 2016-03-29

修回日期: 2016-04-15

基金项目: 国家自然科学基金 (61273349, 61573042)

作者简介: 任天助 (1992 —), 男, 北京人, 硕士生, 研究方向为无人机航路规划算法。

程控制中得到应用。然而上述的文献并没有根据无人机自主决策的特点进行尝试与应用,缺乏对决策结构与具体问题的详细阐述。

本文通过借鉴前人经验提出一套基于情感智能的无人机决策结构,采用情感模型作为感知的融合,强化学习作为经验的积累,以无人机航路规划与威胁规避的具体问题进行理论分析与仿真实验对这个方法的有效性进行验证。

1 无人机的决策结构

1.1 结合情感的无人机决策框架

在传统的无人机决策的框架下,影响无人机决策的因素只有环境,无人机感知外界的环境变化,将这些信息反馈到控制系统中然后做出行为选择,这个过程忽视了无人机内部的一些信息,如无人机的“内在需求”和“身体状况”,将它们都认为处于理想的状况。当无人机处于一些极端的环境或者复杂多变的环境下,无人机自身的续航能力、各零部件状态、性能也是很重要的因素,必须将这样的“身体状况”考虑进来,这时候传统的无人机决策算法(如图 1 所示)就不能满足对决策任务的要求。基于人工情感的无人机决策框架(如图 2 所示)就是基于这样一个思路,将无人机的内部状态或者一些内在的需求信息和外界的环境信息先经过情感模型处理,得到无人机目前所处状态的综合评价,然后反馈到无人机的控制系统中,使无人机能够更加全面地考虑当前无人机所处的实际状态,进而做出更加正确的行为选择。除了上面情感因素对无人机决策的促进方面,另一方面,情感因素的作用体现在无人机对学习经验的积累上,也就是说无人机在学习的过程中会不断地积累面对各种环境的经验,当无人机再次处于类似的环境状态时,就可以利用情感的因素直接指导无人机的控制系统根据以前的经验做出行为选择,而不必通过逻辑推理等过程后再进行决策。这个作用类似于人类的直觉反应,在一些复杂多变的环境下是很有用的,能够使无人机快速做出应对措施,避免对其自身的损害,使无人机在这些复杂环境下的生存能力增强。

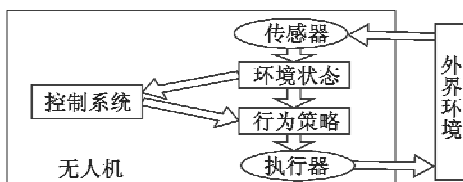


图 1 传统无人机决策框架

Fig. 1 Traditional decision-making framework for UAV

相对于传统的无人机决策框架,基于人工情感的

无人机决策框架在决策的过程中更加综合地考虑了环境和自身状态的一些因素,并且能够利用之前已经学习到的经验,在复杂情况下帮助无人机更好地做出行为选择,提高学习的速度,并且增强学习的性能。本文借鉴文献[8]的思路,可以通过融合强化学习与情感智能的方法进行决策的设计。

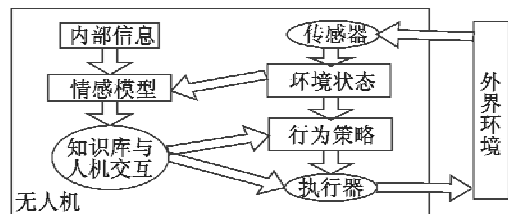


图 2 基于人工情感的无人机决策框架

Fig. 2 Decision-making framework for UAV based on artificial emotion

1.2 无人机情感智能决策模型

让无人机在决策过程中拥有类似人类的智能,一个比较直接的方法就是类比人类的决策方法,使得无人机具有综合自身条件与周围环境的决策能力。现代认知心理学研究表明,人类的决策与自身情感状态有着显著的关系,其过程可由一种多级情感的模型表示。

如图 3 所示,人的情感对于决策的影响分为两级,当人受到外界刺激时,首先会产生类似条件反射式的一级情感,这种情感与对周围情况进行态势感知之后的结果与大脑中的经验知识建立联系,生成更加成熟的二级情感,进而生成正确的决策判断。将类似的决策机制应用在无人机上,以无人机遇到威胁进行规避和重规划为例。无人机发现周围有威胁靠近时,如果其拥有类似人类的情感智能,首先应该出于自身安全考虑,在检测到有威胁靠近时,应对威胁进行规避。同时,如果无人机有需要按时到达的目标,无人机应该在躲避威胁的同时尽可能地向目标靠近,这就需要经过多次训练后无人机在检测到威胁靠近时,根据先验知识,向靠近目标点的方向进行规避。如果说一级情感就是一种“消极的恐惧”,而二级情感则是基于先验知识的一种“积极的恐惧”,因为其在应对无人机自身的情况同时,更多地考虑了它要完成的目标,在智能水平上更近一步。针对这种思路,可以设计如图 4 所示的无人机决策流程。

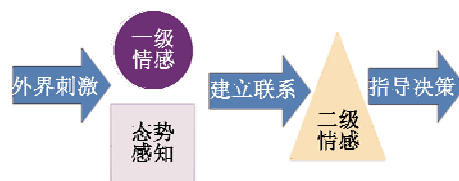


图 3 多级情感决策模型

Fig. 3 Multi-level emotion decision-making model

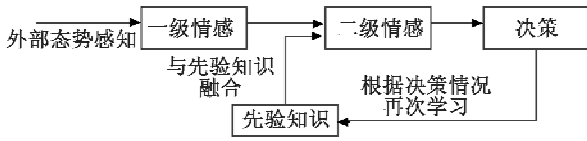


图4 基于情感智能的无人机决策流程

Fig.4 UAV decision-making process based on emotion intelligence

2 情感决策的模型与算法

2.1 情感模型的建立

针对上文提到情感决策流程,有必要建立一种用于决策的情感模型。对于情感模型,根据文献[9]的描述,可以用一个“情感空间”进行向量的描述,图5所示是一种常见的二维情感模型。

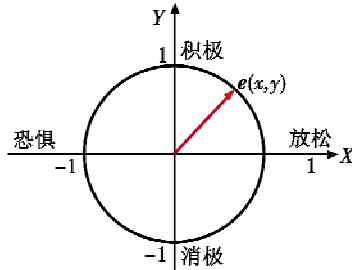


图5 二维情感模型

Fig.5 Two-dimensional emotion model

根据基本情绪的研究,将讨论最常见的4种互为反向的情绪:恐惧、放松、积极和消极。即基本情绪空间 $S = \{A, B, C, D\}$, 其中:A表示放松情绪;B表示恐惧情绪;C表示积极情绪;D表示消极情绪。放松和恐惧互为反向情绪,代表横向坐标轴的两端;积极和消极互为反向情绪,代表纵向坐标轴的两端。所有的情绪的值都被限制在圆心在原点、半径为1的单位圆之中。在此情绪空间中,某一情绪的值是由一个向量表示的,通过向量的模与向量的方向来代表一个情绪。用数学式可以表示情绪为

$$e_a = (x_a, y_a) \quad (1)$$

而情绪强度为

$$|e_a| = \sqrt{x_a^2 + y_a^2} \quad (2)$$

情绪方向为

$$\theta = \arctan \frac{y_a}{x_a} \quad (3)$$

式中, x_a 和 y_a 分别对应图中的 x 轴与 y 轴的4种基本情感的数值。

2.2 情感向量值的确定

对于情感数值的确定,在许多文献中多使用 OCC 模型进行确定。然而 OCC 模型具有复杂多样的特点,对于决策属性比较单一的问题,诸如无人机在威胁环境

中躲避障碍可以采用更简单具体的方法进行设计。如图6所示,图中,五角星代表无人机所在的位置,其与威胁 i 的距离可以用 d_i 来表示。对于不同种类的威胁,还可以对其威胁程度进行加权,其与威胁 i 的威胁权重用 c_i 来表示。认为当大量威胁距离无人机很近时,无人机的情感处于“恐惧”状态,当周围威胁数量很少且距离无人机很远,认为无人机处于“放松”状态。具体计算如下

$$T_a = k \cdot (c_1/d_1 + c_2/d_2 + \dots + c_n/d_n) \quad (4)$$

式中, k 为归一化因子。情感 e_a 的分量 x_a 可以根据 T_a 确定为

$$x_a = \begin{cases} 1 & T_a \leq T_m \\ 0 & T_m < T_a < 2T_m \\ -1 & T_a \geq 2T_m \end{cases} \quad (5)$$

式中, T_m 为选择情感值的阈值,根据实验的结果以及对威胁的敏感程度决定。

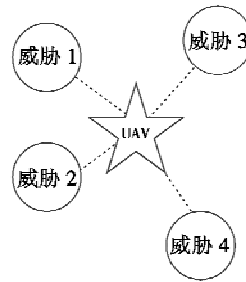


图6 威胁与无人机关系图

Fig.6 The relationship between threat and UAV

情感 e_a 的分量 y_a 的值则由无人机当前的学习情况而定,如果无人机处于相对熟悉的环境,则处于“积极”的情感,如果处于陌生的环境,则处于“消极”的情感。无人机对环境的熟悉程度由学习策略的收敛情况决定。学习策略可采用强化学习、神经网络、专家系统等方式进行离线与在线学习。本文采用 Q 学习方法,旨在对无人机威胁规避问题中一些基本威胁情况进行学习,逐步形成对相对复杂的环境的应对机制。

Q 学习是一类重要的增强学习算法,它属于“非监督学习”,不通过正确的学习样本采取何种行为,而是通过不断试错的方法来逐步发现最优策略。由于无人机往往应用于未知环境的探索,这种算法在各种新环境中更具有适应性。假定 Q 学习系统接受环境状态的输入为 s ,根据学习决策机制,系统输出的行为动作为 a 。此时环境状态在动作 a 作用下,“转移”到新的状态 s' 。系统接受环境中新的状态输入,同时得到由环境对于系统的奖惩反馈 r 。对于 Q 学习系统而言,其目标是学习一个行为策略,使系统选择的动作能够获得环境奖赏的累计值最大。

转移规则可以写为

$$Q(s, a) = r(s, a) + \gamma \cdot \max \{Q(s', a')\} \quad (6)$$

式中, γ 为学习参数, $0 \leq \gamma < 1$ 。

Q 学习算法的运行步骤如图 7 所示。

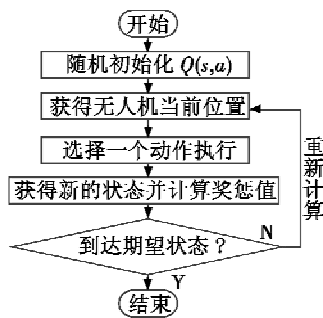


图 7 Q 学习算法流程图

Fig. 7 Q learning algorithm flow

此外, 要对 Q 学习的熟练程度进行描述, 这里以到达目标步数 n 作为指标。设定一个阈值 $M1$, $M1$ 为直接使用贪婪算法无碰撞到达目标的总步数。当 $n > M1$ 时, 认为无人机尚未学习熟练, 此时对应的情感状态认定为消极, 即 $y_a = -1$ 。当 $n < M1$ 时, 认为无人机已经学习熟练, $y_a = 1$ 。此外, 还可以在实际应用中加入对无人机能量的约束, 由于进行学习算法的应用需要进行大量迭代计算以及数据存取, 同时将增加机载计算机的功耗。当无人机自身电量(或者燃料)低于一定程度时, 可以自动关闭 Q 学习计算机, 将情感状态设计为消极。这种设计体现了更贴近于人类情感的思想, 人在疲惫时往往不愿意仔细思考而是就近休息。同时这种设计更加贴近实际中提高无人机航时保证安全的要求。

2.3 基于情感的行为选择策略

根据之前所描述的情感计算方法, 得到人工情感以后, 根据表 1 的方式进行策略选择。在无人机情感处于胆怯状态时, 首先要保证无人机的安全, 此时应该选择紧急避障策略, 无人机不再向目标前进而是向没有威胁的位置进行规避。当无人机能保证自身安全时, 则根据当前掌握先验知识的情况进行选择, 此时如果无人机尚未熟练掌握在当前环境中进行寻路的能力, 则进行 Q 学习, 并采用贪婪算法完成决策, 如果已经掌握, 则直接用 Q 学习的知识进行决策。

表 1 不同情感状态下行为选择策略

Table 1 The behavior selection strategy

X 情感类型	Y 情感类型	
	积极 ($y_a = 1$)	消极 ($y_a < 1$)
放松 ($x_a > -1$)	基于 Q 学习方法	贪婪算法
胆怯 ($x_a = -1$)	紧急避障	紧急避障

3 仿真实验

为了验证方法的可行性, 在仿真环境下对其进行验证。分别在简单静态威胁场景和同时具有动、静态

威胁的复杂场景下, 将本文方法与人工势场法^[10]进行对比。人工势场法实现较为简单, 是一种具有代表性的无人机路径规划与避障算法。仿真实验环境为软件 Matlab 7.0; 计算机配置 Windows XP 操作系统、CPU 为 Inter Core i5、主频 3.3 GHz。

3.1 简单静态威胁场景的学习训练

图 8 中, 蓝色圆形代表无人机所面对的威胁, 无人机的起点由一个红叉表示, 终点以绿色星形来表示, 学习的目标是让无人机在避开威胁的情况下从起点移动到终点。首先让无人机在一些简单环境下进行学习, 场景大小 20×20 , 起点 (1, 1), 终点 (10, 10)。设所有威胁与无人机所在位置均为整数点, 威胁数目为 10, 训练次数为 900 次, 其训练效果如下。

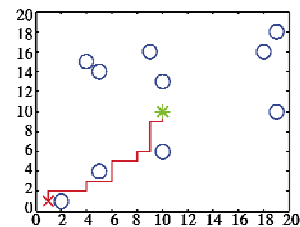


图 8 简单静态威胁场景

Fig. 8 The simple static threat scenarios

这里只是为了验证算法的有效性, 而使用相对简单的质点模型。对于无人机决策的执行手段, 假设无人机的决策集为 $\{(0, 1), (0, -1), (1, 0), (-1, 0)\}$, 即在地图上只能向正北、正南、正东、正西 4 个方向进行移动, 对应实际情况中多旋翼无人机的一般情况。用从起点到终点的总路程作为学习的指标, 其学习过程收敛曲线如图 9 所示。

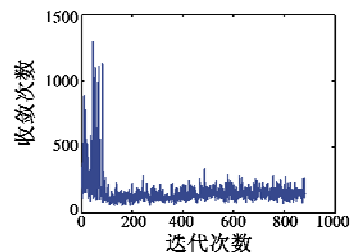


图 9 学习过程收敛曲线

Fig. 9 The convergence curve of learning

3.2 在复杂静态场景中的应用

在简单环境中获取了一定的经验知识以后, 无人机在更复杂的静态环境中进行应用, 此时就需要之前提到的情感智能。如图 10 所示, 场景大小 150×150 , 起点 (10, 10), 终点 (150, 150)。无人机已经对一些简单样本完成训练, 无人机能在复杂的环境中智能地找到一条路径。然而由于有些环境是无人机不熟悉的, 所以飞行中的路径有毛刺现象, 在加入平滑算法之后问题会得到改善。

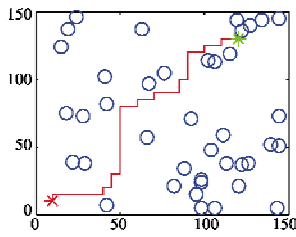


图 10 复杂静态威胁场景

Fig. 10 The complicated static threat scenarios

3.3 在动态环境中的仿真实验

本文为了验证无人机情感智能算法的有效性,设计了如下的一个简单的仿真程序,并采用人工势场法进行对比。场景大小规定为 2500×2500 , 起点可以在程序中设定, 终点假设为一个移动目标, 绕一固定点做半径为 5 的圆周运动。无人机绕过障碍物对目标进行捕捉。障碍物设计为上下各 4 排彼此交错的威胁, 随时间进行上下往复运动, 无人机从起始点以一个固定速度移动到目标位置。两种方法的效果如图 11 所示。

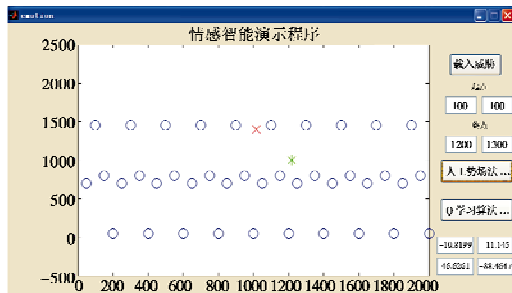


图 11 仿真程序界面

Fig. 11 Simulation program interface

由表 2 的结果可以看出, 两种算法在实时性上不分伯仲, 但是在最优性与安全性上, 经过学习训练之后的情感智能算法有着很大的优势。当然这只是一次实验的值, 多次实验的值在总的寻优时间上有比较大的差异, 但在训练程度相同而情感的时间比例不同的各次实验的差别不大, 其分布如图 12 所示。

表 2 动态场景下仿真结果对比

Table 2 Simulation results

算法	到达时间	路径长度	与威胁碰撞次数	“放松”情感时间
情感智能	28.4	2620	0	11.2
人工势场法	30.1	3478	3	N/A

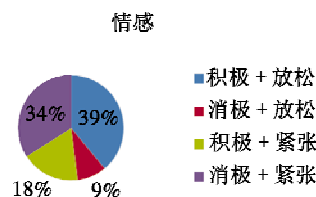


图 12 各情感比例分布图

Fig. 12 The distribution of emotion state

然而即使对于仿真中这样有规律的动态威胁, 无人机依然不能在一半时间都以最佳的状态飞行, 说明多种情感状态均起到了作用, 同时也说明该算法还有很大的优化空间。

4 结束语

本文运用情感智能的理念, 将人工情感模型与无人机决策方法相结合, 提出了基于情感智能的无人机自主决策方法。不同于传统的输入-决策方法, 该方法通过 Q 学习训练使无人机获得基本经验, 根据环境与无人机自身感知决定情感属性, 应用经验与感知的融合进行决策。仿真结果表明, 基于情感智能的无人机自主决策方法在静态航路规划与动态威胁规避的决策中具有良好的实时性。本文的主要工作:

- 1) 提出了结合情感的无人机决策框架与决策流程, 通过将外界信息与自身经验形成高级情感融合, 从而形成从感知到执行的完整决策流程;
- 2) 建立适用于无人机航路规划与威胁规避决策问题的二维情感模型, 并通过 Q 学习对无人机决策过程进行在线学习和训练、实现, 以兼顾在熟悉环境中的决策效率与陌生环境中的适应性。

本文的不足之处在于, 该方法考虑的模型比较简单, 应用到真正的无人机上还存在一定的问题。另外, 在适用于具体问题时往往需要进行大量的学习才能对决策效果有大幅度提高, 本文所有的仿真均假设威胁与无人机坐标均在所设计网格的整数点, 因而要考虑的情况大幅度减少, 很容易找到匹配的学习方案, 而实际情况则复杂得多, Q 学习的经验很难直接进行应用。因此, 接下来的工作希望在经验学习方面提高学习的效率与质量, 并找到在更大范围内应用的具体方法。

参考文献

- [1] Office of the Secretary of Defense. Unmanned aircraft systems roadmap 2005-2030 [R]. Washington: DOD, 2005.
- [2] BREHMER B. The dynamic OODA loop: amalgamating body's OODA loop and the cybernetic approach to command and control [C]//The 10th International Command and Control Research and Technology Symposium (IC-CRTS), 2005:126-139.
- [3] 江道平, 班晓娟, 尹怡欣, 等. 情感理论及基于情感的决策理论与模型研究 [J]. 计算机科学, 2007, 34 (4): 154-157.
- [4] VENTURA R, FERREIR C P. Responding efficiently to relevant stimuli using an affect-based agent architecture [J]. Neuro Computing, 2009, 72(13):923-930.

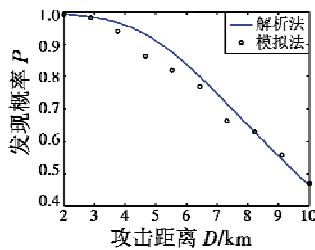


图4 精确信息源下攻击距离与发现概率的关系

Fig.4 Attack distance vs detection probability based on accurate information source

3.2 模糊信息源下仿真结果分析

结合图4、图5可知,空投鱼雷的发现概率随着攻击距离的增大而减小,攻击距离相同时,精确信息源下的发现概率大于模糊信息源下的发现概率,且攻击距离越远优势就越明显。

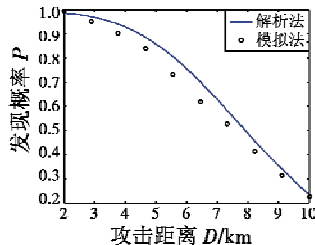


图5 模糊信息源下攻击距离与发现概率的关系

Fig.5 Attack distance vs detection probability based on fuzzy information source

此外,还可以看出,在攻击距离为4~8 km时,解析法的计算结果要略大于模拟法的仿真结果,而在2~4 km和8~10 km时,两者则较为接近,推测是因为解析法中未考虑鱼雷水下运动散布。近距离时,发现概率较大,鱼雷水下运动散布的影响有限;远距离时,空投鱼雷飞行时间较长,鱼雷水下运动散布相较其入

水点散布来说影响较小。

4 结束语

本文就影响空投鱼雷发现概率的3个要素,即鱼雷搜索区域、空投鱼雷入水点散布和潜艇散布进行了分析,在此基础上构建了精确信息源和模糊信息源下的发现概率解析模型。解析法的计算结果与模拟法的仿真结果一致性较好,从而验证了所建解析模型的正确性。该解析模型可快速高效地给出各种态势下空投鱼雷发现概率的精确值,为战场决策提供依据。此外,研究成果完善了空投鱼雷作战效能的相关理论研究。但本文尚未考虑潜艇规避和使用水声对抗器材等情况,在今后的工作中将进一步完善。

参考文献

(上接第15页)

- [5] ROSIS F D, PELACHAUD C, POGGI I, et al. From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent[J]. *International Journal of Human-Computer Studies*, 2003, 59(1): 81-118.
- [6] SU C, LI H. Approaches to affective computing and learning towards interactive decision making in process control engineering[J]. *Acta Automatica Sinica*, 2013, 39(5): 617-625.
- [7] 宿翀,李宏光. 基于情感学习智能体的交互式进化计算方法[J]. *计算机集成制造系统*, 2012, 18(3): 617-625.
- [8] MATSUDA A, MISAWA H, HORIO K. Decision making based on reinforcement learning and emotion learning for social behavior[C]//2011 IEEE International Conference on Fuzzy Systems (FUZZ), IEEE, 2011: 2714-2719.
- [9] 王志良,乔向杰,王超,等. 基于自定义空间和 OCC 模型的情绪建模研究[J]. *计算机工程*, 2007, 33(4): 189-192.
- [10] KOREN Y, BORENSTEIN J. Potential field methods and their inherent limitations for mobile robot navigation [C]//Proceedings of the IEEE Conference on Robotics and Automation, 1991: 1394-1404.