

引入深度强化学习思想的脑-机协作精密操控方法

张腾¹, 张小栋^{1,2}, 张英杰^{1,2}, 陆竹风¹, 朱文静¹, 蒋永玉¹

(1. 西安交通大学机械工程学院, 710049, 西安; 2. 西安交通大学陕西省智能机器人重点实验室, 710049, 西安)

摘要: 针对精密操控中人-机之间缺乏信息双向交互, 以及人的精神状态变化会严重影响肢体操控的精度和安全性等问题, 提出了一种引入深度强化学习思想的脑-机协作精密操控方法。首先, 结合人在上层规划与机器在精细控制上的各自优势, 建立由主动操控和被动调控组成的双环路人-机信息交互机制; 其次, 引入深度强化学习思想, 从蒙特卡罗采样原理出发, 以表征精神状态的脑电信号(EEG)作为模型的输入, 以机器人速度指令作为模型的输出, 推导出脑-机协作方法的数学模型; 再次, 建立具有3个全连接层的精神状态感知网络, 从脑-机接口系统的实时监测计算机内存中提取最后1 000 ms的EEG作为输入信号, 设计开发脑-机协作精密操控算法。最后, 创建轨迹跟踪虚拟环境和任务场景, 对脑-机协作精密操控方法进行了实验验证, 结果表明: 该方法在轨迹跟踪任务的控制精度和完成时间指标上均得到了提高; 相较于传统方法, 该方法的平均轨迹跟踪精度和完成时间指标分别提高了36.55%和22.81%。

关键词: 脑-机协作; 深度强化学习; 脑-机接口; 轨迹跟踪; 精密操控

中图分类号: TP242 **文献标志码:** A

DOI: 10.7652/xjtuxb202102001 **文章编号:** 0253-987X(2021)02-0001-09



OSID 码

A Precise Control Method for Brain-Computer Cooperation with Deep Reinforcement Learning

ZHANG Teng¹, ZHANG Xiaodong^{1,2}, ZHANG Yingjie^{1,2}, LU Zhufeng¹,
ZHU Wenjing¹, JIANG Yongyu¹

(1. School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China;

2. Shaanxi Key Laboratory of Intelligent Robots, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: A precise control method for brain-computer cooperation with deep reinforcement learning is proposed to solve the problem that the lack of bidirectional information interaction between human and computer and the change of mental state in precise control seriously affect the precision and safety of limb control. First of all, combining the advantages of human in global planning and machine in fine control, a 'double-loop' information interaction mechanism composed of active control loop and passive control loop is established. Secondly, the idea of deep reinforcement learning is introduced, and a mathematical model of brain-computer cooperation is derived based on the Monte Carlo sampling principle, with the electroencephalogram (EEG) representing mental state feature as the input of the model and the robot speed instruction as the output. Thirdly, a mental state perception network with three fully connected layers is established, and the EEG of the last 1 000 ms in the real-time monitoring computer memory of brain-machine interface system is extracted as input signal, then a precise brain-computer

cooperation algorithm is designed and developed. Finally, a virtual environment and task scene for trajectory tracking is created, and the precise brain-computer cooperation method is experimentally verified. Results and a comparison with a traditional method show that the proposed method improves both the accuracy and completion time of trajectory tracking control task by 36.55% and 22.81%, respectively.

Keywords: brain-computer cooperation; deep reinforcement learning; brain-machine interface; trajectory tracking; precise control

随着人工智能技术的发展,对具有仿人高级认知能力、能在复杂环境中执行非设定任务的智能机器人的需求日趋紧迫,但是单凭人工智能技术尚无法满足。因此,研究者提出了人-机智能融合方法,该方法结合了人的直观推理、自组织学习以及处理非结构化信息的能力,同时兼具机器强大的计算能力、存储容量和不知疲倦的特性,目的在于充分发挥人和机器两者的优势。随着人-机智能融合系统所面临任务和场景的复杂化,研究者在人-机智能融合的基础之上,提出了一系列脑-机智能融合增强技术,例如脑控汽车^[1]、脑控无人机^[2]、脑控外骨骼机器人^[3]、脑控轮椅^[4]以及脑控假肢^[5-6]等。然而,在精密操控领域(例如医疗、军事和太空探索等),与肢体操控方式相比,脑控方式在稳定性和安全性上还存在风险。因此,目前在精密操控领域依然以人的肢体操控方式为主,例如手术机器人^[7]、排爆机器人^[8]、武器操控系统^[9]以及在轨对星球表面遥操作系统^[10]等。经过研究发现,在人-机交互的精密操控领域,由于缺乏操控者和机器之间信息的双向交互,无法实现对操控者意图的精密感知;同时,由于人脑精神状态(例如疲劳、专注度和脑力负荷等)的变化会影响肢体操控的精度和安全性。

对此,有关学者在精密操控领域引入脑-机接口技术,在不改变肢体操控方式的前提下,使用表面脑电信号(EEG)检测操控者的精神状态,并根据操控者的精神状态对机器人的指令进行补偿调控,以实现精密操控。目前国内外相关的研究主要分为两个方面:操控者精神状态检测研究和基于精神状态的补偿控制研究。

在操控者精神状态检测研究方面:Dijksterhuis等要求驾驶员在驾驶任务中,执行不同难度等级的驾驶任务,利用 EEG 识别驾驶员的脑力负荷,平均识别精度达到 95%^[11];朱成杰等提出在飞行任务中 EEG 的各种节律波会随着飞行员疲劳状态和脑力负荷的变化而改变,因此利用 EEG 检测技术可有效识别飞行员的疲劳状态和脑力负荷^[12];Wang 等在

多属性任务组(MATB)中,要求受试者执行 3 种难度等级的 MATB 任务,以 2~100 Hz 的 EEG 傅里叶变换系数为特征,对受试者脑力负荷的识别精度进行测试,结果表明其识别精度可达 80%左右^[13]。上述研究均证明了 EEG 可有效地反映操控者的精神状态,但是目前的研究还仅停留在精神状态的检测和识别方法上,并未考虑精神状态和操控品质之间的联系。

在基于精神状态的补偿控制研究方面:Wilson 提出了一种自适应自动化系统,该系统可根据人的脑力负荷动态分配机器和人之间的任务属性和等级,目的在于提高操控品质^[14];Jia 等在遥操作任务下,根据操控者的精神状态实时调控机器人的速度和响应时间参数,目的在于提高操控的精度和安全性^[15];杨少增采用模糊建模方法建立了人的精神状态估计和预测模型,以使操作员所承担的任务与其当前的状态两者相匹配^[16]。然而,目前大多数的研究,多预先设定所谓的精神状态“好与坏”的表现特征,从而主观认为当检测到精神状态“好”的特征时就增加任务难度,反之则降低任务难度,忽略了精神状态的多样性(尤其是在跨个体和跨时间角度上精神状态的多样性更强),从而导致精神状态和控制指令之间失匹配,无法有效提高操控品质 and 安全性。

综合上述两方面的研究现状,发现两个方面多单独研究,没有形成一体化的脑-机协作模型;同时,基于精神状态的补偿控制方面主观性强,未考虑到精神状态的多样性。因此,针对人-机交互精密操控领域亟待解决的这两大基础共性问题,本文以机器人最基本的轨迹跟踪任务为应用目标,提出一种引入深度强化学习思想的脑-机协作精密操控方法。首先结合人在上层规划与机器在精细控制上的优势,提出双环路的人-机之间信息交互机制,进而建立一种基于深度强化学习的一体化的脑-机协作方法模型;然后设计相应的精神状态实时监控方法,开发一套具有工程应用价值的脑-机协作精密操控算法;最后拟搭建具有 3 种难度等级的轨迹跟踪虚拟

环境,并设计训练实验、验证实验和对照实验,以验证脑-机协作精密操控方法的有效性。

1 双环路人机信息交互机制

本文融合人在上层规划与机器在精细控制上的各自优势,提出了由主动操控环路和被动调控环路组成的双环路人机信息交互机制,如图 1 所示。考虑到人在上层决策、突发情况处理方面的优势,因此在轨迹跟踪任务中,使操控者对机器人方向指令的控制具有优先权。在主动操控环路中,操控者通过操控装置发送方向指令给机器人,同时通过视觉等信息对机器人的运行状态进行监督,实时地调整方向指令,并对突发的错误进行纠正;考虑到机器在精细控制上的优势,使计算机对机器人速度指令的控制具有优先权。在被动调控环路中,引入深度强化学习思想^[17],创新性的将操控者大脑作为环境对象,将控制算法作为智能体对象,建立一体化脑-机协作方法模型,其以反应操控者精神状态 EEG 微分熵特征作为输入,以机器人速度指令作为输出。模型经过多次自主训练,将多样性的精神状态和机器人的控制指令相匹配,从而促进人-机之间相互适应和监督,实现人脑和计算机协同合作(简称脑-机协作)执行精密操控任务。

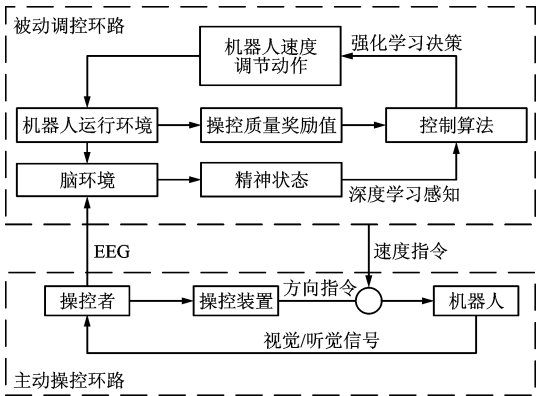


图 1 双环路人机信息交互机制

在一体化脑-机协作模型训练阶段:操控者通过肢体操控方式发送方向指令的同时,脑电设备采集大脑 EEG 并传给控制算法,控制算法根据当前 EEG 生成相应的速度指令,机器人根据方向指令和速度指令执行相应任务。此外,操控者在监视和操控机器人时,会引发精神状态变化(例如:机器人犯错会引发大脑警觉;操控任务复杂且长时间执行任务会由于高脑力负荷而引发大脑疲劳;操控任务过于单一会引发大脑专注度下降等)。精神状态的变

化会影响肢体的操控的精度和安全性,因此脑电采集设备将 EEG 实时输入控制算法,控制算法及时决策机器人的速度指令。机器人每执行一次完整实验,控制算法会根据任务完成精度和时间两个指标计算奖励值并更新模型中网络参数,直到控制算法中网络模型收敛并达到奖励最大化。在一体化脑-机协作模型验证阶段:将训练好的控制算法参数导入到机器人的控制器中,通过所建立深度神经网络感知操控者精神状态,利用强化学习方法根据精神状态决策机器人的速度指令,从而实现脑-机协作精密操控。此模型利用深度强化学习理论建立脑-机之间一体化的架构,真正实现了双环路的交互机制。

2 脑-机协作方法数学模型

本文所建立的脑-机协作方法模型属于“免模型强化学习”范畴,因此,本文根据蒙特卡罗采样原理^[18],从任意起始精神状态 s_1 出发,使用某种策略 G 进行采样,执行该策略 i 步并获得轨迹 τ ,详细的采样过程如图 2 所示。获得采样轨迹 τ 的概率可由下式表示

$$p(\tau) = p(s_1)p_\theta(a_1 | s_1)p(s_2 | s_1, a_1)p_\theta(a_2 | s_2)p(s_3 | s_2, a_2)\cdots p_\theta(a_i | s_i)p(s_{i+1} | s_i, a_i) = p(s_1) \prod_{i=1}^k p_\theta(a_i | s_i)p(s_{i+1} | s_i, a_i) \quad (1)$$

式中: $s_i(i=1, \cdots, k)$ 表示第 i 时刻的精神状态(以下简称状态); $a_i(i=1, \cdots, k)$ 表示第 i 时刻的机器人速度调节动作(以下简称动作); $p_\theta(\tau)$ 是指给定策略神经网络参数 θ 的情况下,出现采样轨迹 τ 的概率; $p(s_1)$ 是指初始状态 s_1 出现的概率; $p_\theta(a_i | s_i)$ 是指给定当前状态 s_i ,采取动作 a_i 的概率; $p(s_{i+1} | s_i, a_i)$ 是指采取当前状态 s_i 和动作 a_i 之后,基于该条件概率返回下一个状态 s_{i+1} 的概率。

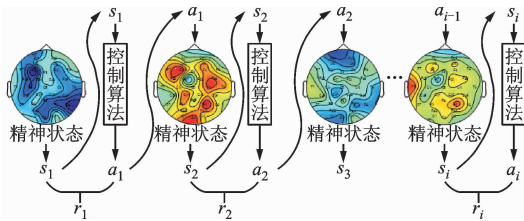


图 2 脑-机协作方法模型蒙特卡罗采样示意图

对于某一个采样轨迹 τ ,可以得到其对应的奖励,通过优化控制算法,可以得到不同的奖励。由于控制算法采取的动作以及出现某一个状态是随机的,最终的目标是找到一个具有最大期望奖励的策略神经网络参数,因此目标函数如下

$$\bar{R}_\theta = \sum_{\tau} R(\tau) p_\theta(\tau) \quad (2)$$

式中: $\bar{R}_\theta$ 表示采样轨迹 τ 的累积期望奖励; $R(\tau)$ 表示采样轨迹 τ 获得的奖励。

得出目标函数之后,需要求解目标函数的最大值以及最大值对应的神经网络参数 θ 。本文采用梯度下降的方法,即需要求出目标函数的梯度 $\nabla \bar{R}_\theta$ 。因此,将目标函数对参数 θ 求导,具体梯度公式如下

$$\begin{aligned} \nabla \bar{R}_\theta &= \sum_{\tau} R(\tau) \nabla p_\theta(\tau) = \sum_{\tau} R(\tau) p_\theta(\tau) \frac{\nabla p_\theta(\tau)}{p_\theta(\tau)} = \\ &= \sum_{\tau} R(\tau) p_\theta(\tau) \nabla \ln p_\theta(\tau) = E[R(\tau) \nabla \ln p_\theta(\tau)] \approx \\ &= \frac{1}{N} \sum_{n=1}^N R(\tau^{(n)}) \nabla \ln p_\theta(\tau^{(n)}) \end{aligned} \quad (3)$$

式中: n 为采样次数; N 为总采样数。由式(1)和式(3)可得

$$\nabla \bar{R}_\theta \approx \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^k R(\tau^{(n)}) \nabla \ln p_\theta(a_i^{(n)} | s_i^{(n)}) \quad (4)$$

为了使奖励值 $R(\tau)$ 不受采样随机性的影响,这里引入一个基准线 b ,因此梯度公式优化为

$$\begin{aligned} \nabla \bar{R}_\theta &\approx \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^k (R(\tau^{(n)}) - b) \nabla \ln p_\theta(a_i^{(n)} | s_i^{(n)}) \\ b &\approx E[R(\tau)] \end{aligned} \quad (5)$$

式中:精神状态 s_i 由所采集到的 EEG 的微分熵特征来表示;机器人速度指令 a_i 由神经网络的输出值而获得;奖励值 R 根据机器人执行任务品质评分获得,详细介绍见第3节。首先将原始 EEG 进行滤波和降采样预处理;其次进行小波分解与重构;最后计算微分熵特征。其中,小波分解与重构过程如下

$$x_j = A_j + \sum_{j=1}^L D_j \quad (6)$$

式中: x_j 表示第 j 个频带的 EEG; L 表示分解层数; A_j 表示近似分量; D_j 表示不同尺度的细节分量^[19]。

对于固定长度的脑电序列可进行如下近似处理计算微分熵^[20]

$$\left. \begin{aligned} s(x) &= - \int_x f(x) \ln(f(x)) dx \approx \\ &= - \int_{-\infty}^{\infty} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \cdot \right. \\ &\quad \left. \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\right) \right\} dx = \\ &= \frac{1}{2} \ln(2\pi\sigma^2) + \frac{1}{2} \\ \sigma^2 &= \frac{1}{L} \sum_{j=1}^L x_j^2 \end{aligned} \right\} \quad (7)$$

式中: $s(x)$ 表示微分熵特征值; $f(x)$ 表示时间序列的概率密度函数; μ 和 σ 分别表示高斯分布的均值和标准差。

最后,利用梯度下降法更新策略神经网络的参数 θ ,直到网络模型收敛,具体算法如下

$$\theta = \theta + \nabla \bar{R}_\theta \quad (8)$$

3 脑-机协作精密操控算法

为了将反映精神状态的实时 EEG 输入给控制算法的模型,每次实验总是提取计算机内存中最后 1 000 ms 长的 EEG 作为输入信号。对输入信号的处理主要分为 3 个阶段:①预处理,采用 4 阶巴特沃兹带通滤波器处理 EEG,保留 0.5~45.0 Hz 的频带信号,然后进行降采样处理;②小波变换处理,采用 5 阶 Daubechies 小波基函数,从 EEG 中分解并重构出 5 种节律波,小波分解原理如图 3 所示;③特征提取,分别计算 32 个通道的 5 种节律波微分熵特征,形成 160 维的特征数据矩阵 $\mathbf{S}$, $\mathbf{S}$ 即作为反映操控者精神状态的特征输入给脑-机协作方法模型。

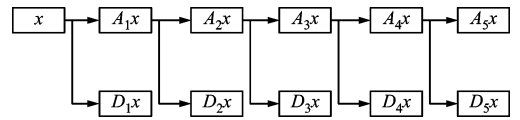


图3 5层小波分解原理图

在脑-机协作方法模型中为了兼顾算法精度和响应速度,本文建立了3层全连接神经网络,详细的网络拓扑结构如图4所示。网络输入为精神状态(即表征精神状态的特征矩阵 $\mathbf{S}$),输出为机器人速度指令 a_i 。其中输入层有 160 个神经元,隐含层有 80 个神经元,输出层有 4 个神经元,分别表示 4 种无量纲的速度等级(0.5、1、2 和 2.5)。隐含层采用 tanh 激活函数^[21],输出层经过 softmax 函数^[22]处理之后,利用随机策略选择速度指令 a_i ,其中概率值高的被选中几率高,反之,被选中的几率低,目的在于符合“探索”和“利用”平衡的原则(EEb),从而获得最大的奖励。奖励值 R 由轨迹跟踪精度和完成时间两项指标组成,详细表达式如下

$$R = \frac{1}{\sum_{m=1}^M |Y_m - O_m|} + \frac{g}{T} \quad (9)$$

式中: Y 表示机器人的行走轨迹; O 表示目标轨迹; T 表示每实验一次完成的时间; g 表示时间系数; M 表示整个轨迹的总步数。将数据组 (s, a, R) 输入给目标函数,根据自适应矩估计梯度下降法(ADAM)^[23]更新神经网络参数,其中学习率(r_L)设置为

0.001。训练阶段每执行一次完整实验,模型更新一次,直到模型收敛为止。

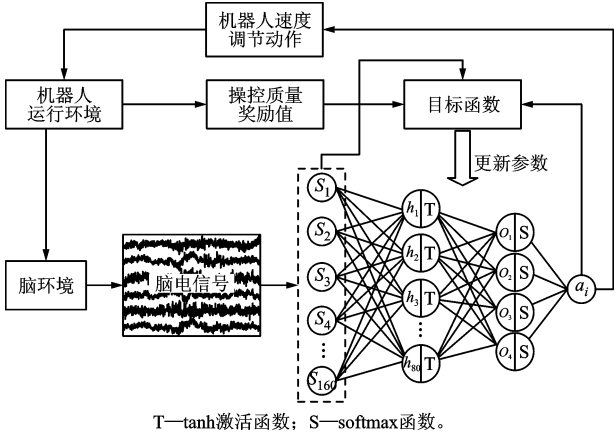
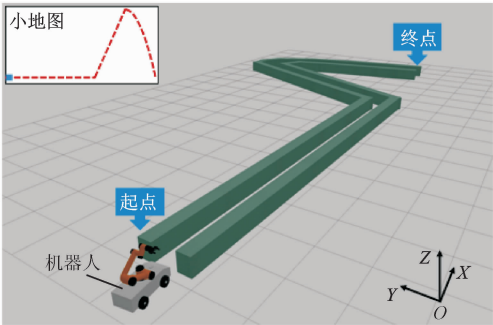


图 4 脑-机协作方法模型参数更新原理图

4 实验验证分析

4.1 环境平台及实验系统创建

为了验证脑-机协作方法的有效性,本文以精密操控机器人执行轨迹跟踪任务为例。此任务作为远程操控特种机器人排爆、消防、救援等领域关键的任务之一受到广泛关注。本实验搭建了一个脑-机协作操控的轨迹跟踪环境平台,具体如图 5a 所示,操控者通过鼠标控制机器人执行轨迹跟踪任务。为了增加实验的多样性,设计了水平直线、斜线和曲线 3



(a) 轨迹跟踪环境

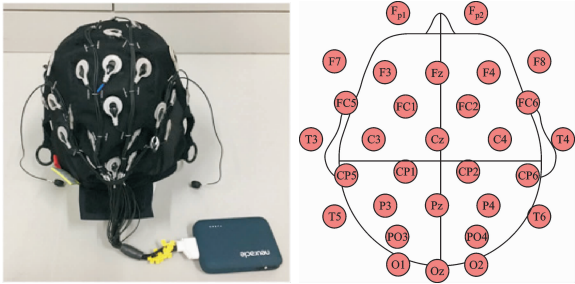


(b) 实验场景

图 5 轨迹跟踪环境及受试者实验场景

种难度等级的轨迹。实验场景如图 5b 所示,其中机器人的方向指令由操控者通过鼠标控制,速度指令由计算机中的控制算法根据操控者精神状态的变化不断地调节,机器人结合方向和速度两项指令,执行规定的任务。在轨迹跟踪任务中,操控者通过观察机器人的运行状态,不断调整方向指令,同时计算机中的控制算法通过检测人脑精神状态实时调整机器人的速度指令,从而通过脑-机协作,实现对机器人的精密操控。每完成一次完整实验,控制算法会记录机器人行走的轨迹和完成时间,并根据此两项指标计算奖励值,具体计算方法如式(9)。控制算法中的神经网络模型根据奖励值更新参数,直到神经网络模型收敛。

实验采用博睿康公司(Neuracle)开发的 32 通道的脑电采集设备,如图 6a 所示,其采样频率为 1 000 Hz,通过无线路由与电脑相连。本实验按照国际 10-20 系统选择脑电通道,具体电极分布位置如图 6b 所示。



(a) 采集设备 (b) 电极分布位置

图 6 脑电采集设备及电极分布位置

4.2 受试者及实验步骤

本次实验共有 5 名受试者(标记为 Sub1 ~ Sub5,其中 1 名为女性),年龄均在 23~30 岁,无精神疾病史,视力或矫正视力在 1.0 以上。实验过程中要求受试者静坐在电脑屏幕前,通过鼠标控制屏幕中的机器人执行轨迹跟踪的任务。每完成 3 种预设轨迹(水平直线、斜线及曲线)的跟踪任务即为完成一次实验的全过程。实验共分为 3 个阶段,分别为训练阶段、验证阶段和对照阶段。训练阶段用来训练脑-机协作方法模型参数;验证阶段是将训练好的脑-机协作方法模型输入到机器人的控制器中,进行验证实验;对照阶段不使用脑-机协作方法(即传统方法),其他设置与验证阶段相同。训练阶段实验执行 65 次,前 5 次用来练习轨迹跟踪任务,防止因为操作的熟练度不同而影响实验结果,后 60 次为正式实验。每 20 次中间有 1 min 的休息时间。对照

阶段实验和验证阶段实验各执行 25 次,前 5 次均用来练习轨迹跟踪任务,后 20 次为正式实验,具体实验步骤如图 7 所示。

考虑到精神状态的多样性,每名受试者在一天之中不同的时间段进行实验(例如 Sub1 的实验时间为 20:00~22:00;Sub2 和 Sub5 的实验时间为 9:00~11:00;Sub3 的实验时间为 14:30~16:30;Sub4 的实验时间为 16:30~18:30),适当增加了表征精神状态的 EEG 数据的多样性。考虑到每名受试者实验的舒适度,实验时长为 2 h。

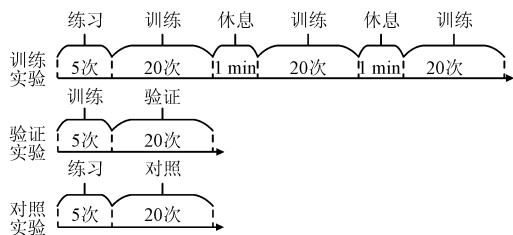


图 7 实验步骤示意图

4.3 ADAM 优化器学习率分析

通过研究发现 ADAM 优化器中学习率会影响脑-机协作方法模型的训练效果。分析其原因是由于在模型收敛过程中发生了梯度消失或梯度爆炸的现象。因此为了验证此猜测,这里进行一个离线的测试,选择受试者 Sub1 的对照组实验数据,以图 4 中的 3 层全连接神经网络为基础,建立精神状态影响下的操控品质预测网络,分别设置 6 种 ADAM 优化器学习率 r_L 参数,分析学习率与损失函数值之间的关系,具体如图 8 所示。损失函数值越小,代表神经网络收敛性越好,训练效果越好,反之损失函数值越大,则神经网络收敛性差,训练效果差。当 $r_L \geq 0.1$ 时,由于梯度爆炸原因使误差曲线出现了二次震荡的现象,并且 r_L 值越大,震荡越剧烈,波峰的横坐标值越大,从而造成收敛速率慢,训练效果差。当 $r_L \leq 0.01$ 时,由于避免了梯度爆炸而使震荡现象消失。在 $0.001 \sim 0.01$ 之间, r_L 值越小,初始损失函数值越低,整体收敛效果越好,但是当 r_L 值降低至 0.0001 时,由于梯度消失的原因使收敛速率骤然变慢,当迭代 50 次时,损失函数值是 $r_L = 0.001$ 时的 15 倍,是 $r_L = 0.01$ 的 28 倍;当迭代 100 次时,损失函数值是 $r_L = 0.001$ 时的 5.8 倍,是 $r_L = 0.01$ 时的 5 倍。实验结果显示:当 r_L 在 0.001 附近时,既可以避免由于梯度爆炸造成的训练模型鲁棒性差的问题,又可以避免由于梯度消失造成的训练模型收敛过缓的问题,因此可达到最优的训练效果。

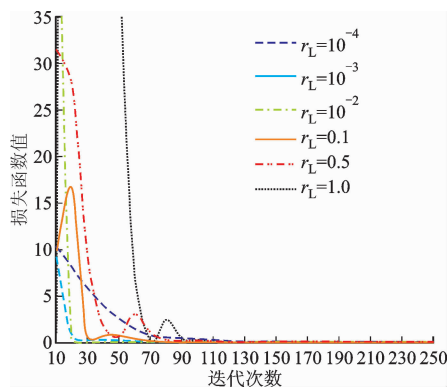


图 8 学习率与损失函数值间关系

4.4 操控品质分析

为了验证面向精密操控轨迹跟踪任务的脑-机协作方法的有效性,这里主要通过脑-机协作方法模型收敛性和轨迹跟踪任务完成品质两个方面分析。选择具有代表性的受试者 Sub1 为例进行讨论,该受试者一共进行了 8 组训练实验,其模型收敛曲线如图 9 所示,纵坐标为奖励值,代表操控品质。图中置信区间的上限和下限分别为 95%。前 22 次实验曲线呈现剧烈震荡状态,并且置信区间范围较大,说明数据概率分布不稳定。之后模型逐渐稳定,并达到收敛,奖励值稳定在相对较高的 1.92 水平,其置信区间范围变窄,说明数据概率分布稳定,操控品质维持在相对较高水平。但是,其中仍然存在轻微波动,例如第 39 和第 49 次实验的奖励值出现明显的降低,这是因为在训练过程中,脑-机协作方法模型在输出动作指令时按照 EEB 原则,存在输出非最优速度的概率,会导致受试者的操控品质降低。

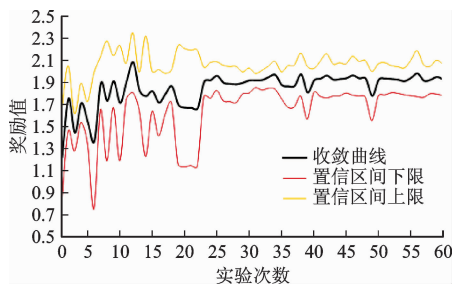


图 9 训练模型收敛曲线图

将受试者 Sub1 训练实验中的脑-机协作方法模型参数导入到验证实验中,对操控品质进行分析,验证实验与对照实验对比曲线如图 10 所示。由图 10 可以看到,对照实验曲线在 20 次实验中剧烈振荡,这是由于对照实验中机器人的速度无法与操控者的精神状态匹配,导致奖励值变化剧烈,且大多数奖励值很低,操控品质很差。相反,验证实验曲线在 20

次实验中始终保持在一个相对较高的奖励值水平上轻微浮动,说明操控品质良好。原因在于脑-机协作方法可根据操控者的精神状态匹配最优的机器人速度,从而有效提高了操控品质。

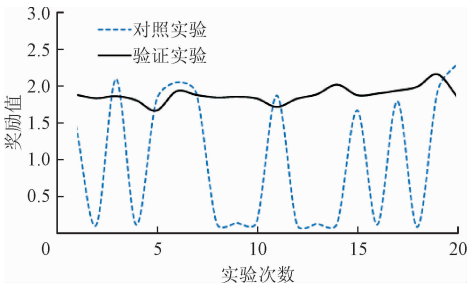
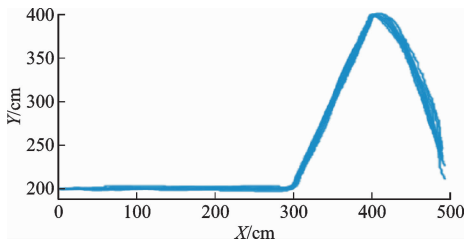
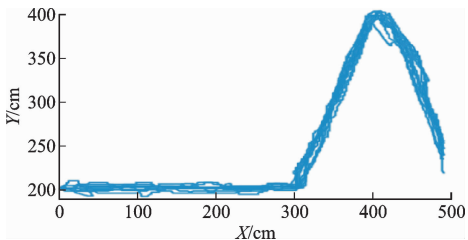


图 10 操控品质曲线对比图

为了更直观的讨论脑-机协作方法在轨迹跟踪任务中的有效性,分别记录了受试者 Sub1 在验证实验和对照实验中的 10 次轨迹,具体如图 11 所示。从对照实验相对波动的轨迹上可知,由于其频繁出现目标轨迹缺失现象,从而操控者需要不断地调整方向指令,这不仅导致机器人与障碍物碰撞的次数增加、整个任务所花费的时间变长,而且增加了脑力负荷和负面情绪。而对于使用了该方法的验证实验,其轨迹跟踪精度要优于对照实验,目标轨迹缺失现象较少,不需要操控者频繁调整方向指令,操作更精准和高效。



(a) 验证实验结果



(b) 对照实验结果

图 11 验证和对照实验轨迹曲线图

为了进一步证明脑-机协作方法的普遍有效性,对所有受试者的操控品质进行了对比分析,结果如表 1 所示。表 1 中各指标计算公式如下

$$\left. \begin{aligned} R_h &= \frac{1}{\sum_{m=1}^H |Y_m - O_m|} \\ R_b &= \frac{1}{\sum_{m=1}^B |Y_m - O_m|} \\ R_c &= \frac{1}{\sum_{m=1}^C |Y_m - O_m|} \\ R_t &= \frac{g}{T} \end{aligned} \right\} \quad (10)$$

式中: H 表示水平直线轨迹的步数; B 表示斜线轨迹的步数; C 表示曲线轨迹的步数; R_h 、 R_b 、 R_c 和 R_{ave} (为 3 者之和)分别代表水平直线、斜线、曲线以及整条轨迹的归一化奖励值,其值越大代表轨迹跟踪精度越好,反之,轨迹跟踪精度越差; R_t 代表时间指标奖励值,其值越大代表完成时间越短,反之,完成时间越长。因此, $R_{ave} + R_t$ 的值越大表示操控品质越好,反之,操控品质越差。表 1 结果显示:相比较于对照实验,5 名受试者验证实验的平均操控品质提高了 59.36%,证明了脑-机协作方法的普遍有效性。其中平均轨迹跟踪精度和完成时间两项指标分别提高了 36.55%和 22.81%,说明此方法不仅提高了轨迹跟踪任务的控制精度,而且缩短了操控的时间。

表 1 轨迹跟踪任务操控品质对比表

受试者	实验项目	R_h	R_b	R_c	$R_{ave} + R_t$
Sub1	验证	0.53	0.21	0.27	1.92
	对照	0.45	0.15	0.18	0.84
	提高比例/%	17.78	40.00	50.00	128.57
Sub2	验证	0.50	0.20	0.28	0.99
	对照	0.33	0.24	0.28	0.92
	提高比例/%	51.00	-16.00	0.00	7.61
Sub3	验证	0.46	0.21	0.23	1.46
	对照	0.28	0.18	0.18	0.73
	提高比例/%	64.29	16.67	27.78	100.00
Sub4	验证	0.62	0.16	0.19	1.01
	对照	0.40	0.10	0.12	0.70
	提高比例/%	55.00	60.00	58.33	44.29
Sub5	验证	0.63	0.26	0.24	1.14
	对照	0.36	0.18	0.16	0.98
	提高比例/%	75.00	44.44	50.00	16.33
平均提高率/%		52.69	29.12	37.21	59.36

注:黑体数据代表最优值。

5 结 论

本文面向人-机交互的精密操控领域,针对人、机之间缺乏信息双向交互,以及操控精度和安全性受操控者精神状态影响的两大问题,通过引入脑-机接口技术,提出了一种脑-机协作精密操控方法研究。通过人机信息交互机制创建、脑-机协作精密操控数学模型推导与算法研究,以及实验验证分析,得出如下主要结论:①结合人在上层规划与机器在精细控制上的各自优势,可以创建一种双环路人-机之间信息交互机制;②通过引入深度强化学习思想,以表征操控者精神状态的 EEG 微分熵特征作为模型的输入,以机器人速度指令作为模型的输出,可以获得一体化的脑-机协作方法模型;③基于精神状态实时监控,采用 3 层全连接神经网络感知模型,可以实现脑-机协作精密操控算法;④通过轨迹跟踪虚拟环境和任务场景创建,能够实现对脑-机协作方法的实验验证分析。实验结果表明:本文方法明显提高了轨迹跟踪任务的控制精度,缩短了任务执行时间。本文方法不仅实现了脑-机协作精密操控轨迹跟踪任务,而且借助此项任务的研究,探讨了一体化的脑-机协作模型的构建,促进了脑-机之间的信息双向、实时交互,实现人-机交互系统的互适应、互监督和互增长。

参考文献:

- [1] BI Luzheng, WANG Huikang, TENG Teng, et al. A novel method of emergency situation detection for a brain-controlled vehicle by combining EEG signals with surrounding information [J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2018, 26(10): 1926-1934.
- [2] DUAN Xu, XIE Songyun, XIE Xinzhou, et al. Quadcopter flight control using a non-invasive multi-modal brain computer interface [J]. Frontiers in Neurorobotics, 2019, 13: 23.
- [3] 唐智川, 张克俊, 李超, 等. 基于深度卷积神经网络的运动想象分类及其在脑控外骨骼中的应用 [J]. 计算机学报, 2017, 40(6): 1367-1378.
TANG Zhichuan, ZHANG Kejun, LI Chao, et al. Motor imagery classification based on deep convolutional neural network and its application in exoskeleton controlled by EEG [J]. Chinese Journal of Computers, 2017, 40(6): 1367-1378.
- [4] 徐光华, 张锋, 谢俊, 等. 稳态视觉诱发电位的脑-机接口范式及其信号处理方法研究 [J]. 西安交通大学

学报, 2015, 49(6): 1-7.

- XU Guanghua, ZHANG Feng, XIE Jun, et al. Brain-computer interface paradigms and signal processing strategy for steady state visual evoked potential [J]. Journal of Xi'an Jiaotong University, 2015, 49(6): 1-7.
- [5] 张小栋, 李睿, 李耀楠. 脑控技术的研究与展望 [J]. 振动、测试与诊断, 2014, 34(2): 205-211.
ZHANG Xiaodong, LI Rui, LI Yaonan. Research on brain control technology [J]. Journal of Vibration, Measurement & Diagnosis, 2014, 34(2): 205-211.
- [6] ZHANG Jinhua, WANG Baozeng, ZHANG Cheng, et al. An EEG/EMG/EOG-Based multimodal human-machine interface to real-time control of a soft robot hand [J]. Frontiers Neurorobotics, 2019, 13: 7.
- [7] 李治非, 杨阳, 苏月, 等. 我国外科手术机器人研究应用现状与思考 [J]. 中国医学装备, 2019, 16(11): 177-181.
LI Zhifei, YANG Yang, SU Yue, et al. Current status and thinking of research and application of surgical robots in China [J]. China Medical Equipment, 2019, 16(11): 177-181.
- [8] 陈涛. 遥操作排爆机器人研制及其运动控制研究 [D]. 济南: 山东大学, 2019: 6-12.
- [9] 唐道光, 张公平, 杜肖, 等. 人在回路对导弹制导性能的影响研究 [J]. 航空兵器, 2018(3): 24-30.
TANG Daoguang, ZHANG Gongping, DU Xiao, et al. Research on missile guidance performance of soldier-in-the-loop [J]. Aero Weaponry, 2018(3): 24-30.
- [10] BURNS J O, MELLINKOFF B, SPYDELL M, et al. Science on the lunar surface facilitated by low latency telerobotics from a lunar orbital platform-gateway [J]. Acta Astronautica, 2019, 154: 195-203.
- [11] DIJKSTERHUIS C, WAARD D D, BROOKHUIS K A, et al. Classifying visuomotor workload in a driving simulator using subject specific spatial brain patterns [J]. Frontiers in Neuroscience, 2013, 7: 149.
- [12] 朱成杰, 陈继龙, 施斌斌. 飞行疲劳检测技术的研究进展 [J]. 航空航天医学杂志, 2019, 30(5): 543-546.
ZHU Chengjie, CHEN Jilong, SHI Binbin. Research progress of flight fatigue detection technology [J]. Journal of Aerospace Medicine, 2019, 30(5): 543-546.
- [13] WANG Ziheng, HOPE R M, WANG Zuoguan, et al. Cross-subject workload classification with a hierarchical Bayes model [J]. NeuroImage, 2012, 59(1): 64-69.

- [14] PARASURAMAN R, WILSON G F. Putting the brain to work: neuroergonomics past, present, and future [J]. *Human Factors*, 2008, 50(3): 468-474.
- [15] JIA Y, XI N, LIU S, et al. Quality of teleoperator adaptive control for telerobotic operations [J]. *The International Journal of Robotics Research*, 2014, 33(14): 1765-1781.
- [16] 杨少增. 基于多源生理数据与模糊建模方法的操作员功能状态预测与调节 [D]. 上海: 华东理工大学, 2014: 72-134.
- [17] 赵星宇, 丁世飞. 深度强化学习研究综述 [J]. *计算机科学*, 2018, 45(7): 1-6.
ZHAO Xingyu, DING Shifei. Research on deep reinforcement learning [J]. *Computer Science*, 2018, 45(7): 1-6.
- [18] 刘雪晨, 李宝瑜, 张晰. 蒙特卡洛特征样本采样方法研究 [J]. *统计与信息论坛*, 2019, 34(1): 3-12.
LIU Xuechen, LI Baoyu, ZHANG Xin. The research on Monte Carlo feature sample generation method [J]. *Statistics & Information Forum*, 2019, 34(1): 3-12.
- [19] ZHANG T, ZHANG X, ZHANG Y, et al. Effects of user fatigue mental state on the facial-expression paradigm of BCI [C]//2nd WRC Symposium on Advanced Robotics and Automation. Piscataway, NJ, USA: IEEE, 2019: 394-400.
- [20] 石立臣. 基于脑电信号的警觉度估计研究 [D]. 上海: 上海交通大学, 2012: 45-48.
- [21] 刘坤华, 钟佩思, 徐东方, 等. 基于双曲正切函数的修正线性单元 [J]. *计算机集成制造系统*, 2020, 26(1): 145-151.
LIU Kunhua, ZHONG Peisi, XU Dongfang, et al. Tangent-based rectified linear unit [J]. *Computer Integrated Manufacturing Systems*, 2020, 26(1): 145-151.
- [22] 章东平, 陈思瑶, 李建超, 等. 基于改进型加性余弦间隔损失函数的深度学习人脸识别 [J]. *传感技术学*

报, 2019, 32(12): 1830-1835.

ZHANG Dongping, CHEN Siyao, LI Jianchao, et al. Deep learning face recognition based on improved additive cosine margin softmax loss function [J]. *Chinese Journal of Sensors and Actuators*, 2019, 32(12): 1830-1835.

- [23] KINGMA D P, BA J L. Adam: a method for stochastic optimization [C]//3rd International Conference on Learning Representations. Montreal, Canada: ICLR, 2015: 20193407343622.

[本刊相关文献链接]

仲文远, 李大海, 张进华, 等. 可穿戴式干电极脑机接口系统设计. 2020, 54(6): 066-74. [doi:10.7652/xjtuxb202006009]

杨磊, 杨帆, 何艳. 采用样本熵自适应噪声完备经验模态分解的脑电信号眼电伪迹去除算法. 2020, 54(8): 177-184. [doi:10.7652/xjtuxb202008023]

张克旭, 杜昌旺, 赵浩淇, 等. 针对局灶性癫痫患者的脑电微状态分析. 2020, 54(9): 157-163. [doi:10.7652/xjtuxb202009018]

张彦军, 谢俊, 薛涛, 等. 稳态视觉诱发电位脑机接口的现场可编程逻辑门阵列实现. 2020, 54(2): 158-165. [doi:10.7652/xjtuxb202002020]

杜光景, 谢俊, 张玉彬, 等. 用于稳态视觉诱发电位脑机接口目标识别的深度学习方法. 2019, 53(11): 042-48. [doi:10.7652/xjtuxb201911006]

吴敏, 王化宁, 张羿阳, 等. 情绪面孔识别状态下抑郁症患者大脑微状态的差异性研究. 2019, 53(4): 143-149. [doi:10.7652/xjtuxb201904021]

王栋, 李扩, 刘晓芳, 等. 利用高频脑电的局灶性癫痫患者癫痫发作检测. 2018, 52(2): 148-154. [doi:10.7652/xjtuxb201802023]

赵广宾, 王会友, 秦歌, 等. 脑深部刺激电极表面多壁碳纳米管的电泳沉积工艺研究. 2019, 53(2): 142-149. [doi:10.7652/xjtuxb201902019]

(编辑 刘杨)