

【工业设计】

虚拟现实多通道信息呈现方式的认知负荷定量研究

罗世怀, 吕健, 刘翔

(贵州大学 现代制造技术教育部重点实验室, 贵阳 550025)

摘要: **目的** 针对虚拟现实体验系统中的不同信息呈现方式, 对使用者交互效率影响难以确定及量化的问题, 进行定量研究。**方法** 通过搭建虚拟现实场景, 以信息呈现通道的类别与数量作为变量, 展开跟踪-检测响应双任务实验; 通过记录任务行为数据中的跟踪误差和响应时间, 以及生理数据中的瞳孔直径大小, 分析并讨论在不同通道刺激影响因素下的实验中, 任务绩效及眼动生理反应变化规律; 同时结合主观负荷评价数据, 建立了基于BP神经网络的多通道认知负荷模型, 以认知负荷作为交互效率的综合评价指标, 量化任务执行效率。**结果** 信息呈现的通道类别及其数量对任务效率均有显著的影响。**结论** 信息呈现通道数量与任务绩效及生理反应呈一定程度的正相关; 多个任务使用相同通道呈现信息会损害所有任务的绩效, 增加认知负荷。模型输出的负荷值与主观认知负荷评估值吻合较好, 相对误差为8.2%, 验证了其有效性。

关键词: 认知负荷模型; 多通道交互; 交互效率; 虚拟现实

中图分类号: TB472 **文献标识码:** A **文章编号:** 1001-3563(2023)04-0069-08

DOI: 10.19554/j.cnki.1001-3563.2023.04.009

Quantitative Study on Cognitive Load of Multi-channel Information Presentation in Virtual Reality

LUO Shi-huai, LYU Jian, LIU Xiang

(Key Laboratory of Advanced Manufacturing Technology of the Ministry of Education, Guizhou University, Guiyang 550025, China)

ABSTRACT: Aiming at the problem that different information presentation methods in the virtual reality experience system are difficult to determine and quantify the impact on user interaction efficiency, the work aims to conduct quantitative research. By building a virtual reality scene, the category and number of information presentation channels were used as variables to carry out the track-detection response dual-task experiment. By recording the tracking error and response time in the task behavior data, as well as the pupil diameter in the physiological data, the changes in task performance and eye movement physiological response in experiments under different channel stimulation factors were analyzed and discussed. At the same time, combined with subjective load evaluation data, a model of multi-channel cognitive load based on BP neural network was established. The cognitive load was used as a comprehensive evaluation index of interaction efficiency to quantify the task execution efficiency. The results showed that the type and number of channels for information presentation had a significant impact on task efficiency. The number of channels for information presentation is positively correlated with task performance and physiological response to a certain extent. Using the same channel for multiple tasks to present information will harm all the performance of all tasks and increase the cognitive load. At the same time, the cognitive load value of the model is in good agreement with the evaluation value of subjective cognitive load, with an absolute error of 8.2%, which verifies its effectiveness.

KEY WORDS: cognitive load model; multi-channel interaction; interaction efficiency; virtual reality

收稿日期: 2022-09-25

基金项目: 国家自然科学基金项目(52065010); 贵州省自然学科基金(ZK2021341); 贵州省科技项目(黔科合支撑[2021]一般397; 黔科合支撑[2022]一般008)

作者简介: 罗世怀(1996—), 男, 硕士生, 主攻智能设计、虚拟现实研究。

通信作者: 吕健(1983—), 男, 博士, 副教授, 主要研究方向为工业设计、智能设计研究。

虚拟现实技术作为智能工业制造的一部分,通过其高度逼真感的三维交互式虚拟可视化环境在工业制造、培训领域中得到了广泛应用。在模拟驾驶室、设备监控室等环节复杂、需求变化多的虚拟操作任务中,操作员在操控设备时,通常还需要处理来自人机界面的其他信号^[1]。以往的研究证实,相对于单一感官的刺激信号,多通道信号可以通过减少感知模糊性、降低反应时间等方式来提高操作员任务执行的认知绩效^[2]。根据认知负荷理论,认知负荷是交互难易程度重要评价指标之一,且认知负荷的高低直接影响任务执行效率^[3]。因此,虚拟现实下信息呈现方式的影响和构建认知负荷量化模型是一个值得深入研究的主题。

1 相关工作

1.1 多通道信息呈现对任务绩效影响

与单一任务相比,在多任务情况下,额外的任务会增加工作负载,这种负载会对整体任务性能产生影响^[4]。Huang 等^[5]通过模拟驾驶员执行并行驾驶任务,证实听觉信息比视觉信息能让被试者更好地理解指令。这种呈现跨模态信息优于模态内信息的优势已经由多资源理论解释,该理论提出视觉处理与听觉处理使用不同的资源^[6],可减少对公共处理资源的竞争,从而为双任务处理提供更好的认知绩效。

而多通道交互(或显示、输出)是指将信息同时呈现给用户的多个感官通道,而通过多种通道呈现相同的信息,即冗余多通道输出,有可能提高信息传递的能力,减少认知负荷^[7]。Pitts 等^[8]发现冗余的视觉-听觉和视觉-触觉双通道刺激比单独使用视觉刺激减少了反应的时间,这种现象被称为冗余目标效应(RTE)。Sella 等^[9]等在模拟击球游戏中,三通道线索比双通道线索检测得更快更准确,而双通道线索比单通道线索检测得更快、更准确。

1.2 认知负荷评估

Calandra 等^[10]提出多通道交互对认知负荷的影响,主要体现在提供高效的人机交互方面。其中认知负荷是任务复杂性与用户完成任务所需的认知能力之比。目前,认知负荷评估方法主要有主观评价、绩效评估和生理三大类评估指标^[11]。

主观评价法是被试者根据任务中的自身主观感受,对负荷进行评价,包括美国航天局负荷指数(NASA-TLX)、主观认知负荷评价(SWAT)等成熟方法^[12]。绩效评估法通过被试者参加任务的绩效评估认知负荷,在双任务范式中要求被试者在规定时间内同时进行两项任务,通过被试者任务完成的准确率和响应时间长短评估认知负荷的高低^[13]。生理实验测量通过获取被试者在任务过程中不同的生理反应来评估认知负荷。Wang L 等^[14]通过设置三种不同时间压

力的不同任务发现,主观认知负荷与瞳孔直径和时间压力呈线性关系。柯青等^[15]通过研究眼动跟踪技术和认知负荷相关的生理指标,通过心理学理论评估用户的认知负荷。

在构建认知负荷评估模型领域,黄晶等^[16]模拟驾驶实验,通过获取各种情绪下的心电生理信号,提出基于个性化敏感生理特征的驾驶负荷评价方法。BP神经网络通过模拟人脑对复杂信息的处理机制,能够建立输入与输出之间复杂的非线性关系的预测模型^[17]。在认知负荷模型研究中,研究者通过信号采集设备,提取被试者人眼动、脑电等生理信号,运用支持向量机及神经网络^[18]等算法,建立认知负荷评估模型。由于生理信息测量的局限性以及特征选取的单一性,而导致其分类评估精度不高^[11]。

因此,本文在多通道交互的基础上,通过虚拟现实场景呈现多种通道刺激,搭建跟踪-检测响应双任务范式,研究在呈现不同通道刺激信号及通道数量等因素下,任务绩效及眼动生理指标的影响。使设计者更好地利用多通道信息呈现的特性来改进虚拟现实场景的设计。同时,通过受试者眼动生理信号和任务绩效作为评价指标,结合主观评价数据,基于BP神经网络构建多通道认知负荷评估模型,提升认知负荷的评估准确性,为后续虚拟现实多任务场景提供设计方法和优化途径。

2 方法

2.1 多通道认知负荷质量评价模型的构建

人通过触觉、听觉和视觉等多种感觉信息来感知物理环境。这些不同形式的信息需要从有效的合并中得到一个增强和连贯的信号。Ohshiro 等^[19]构建多感官的归一化模型,成功地用它解释了包括空间原理、逆效应等多感融合中许多重要的原理。

如图1所示,展示了两个不同的单通道刺激在多感神经元中进行融合处理,并产生输出的过程。可简单理解为,单通道输入 q_1 和 q_2 为对应通道响应的外界刺激引起的神经反应,随后在对应的初级感觉层中

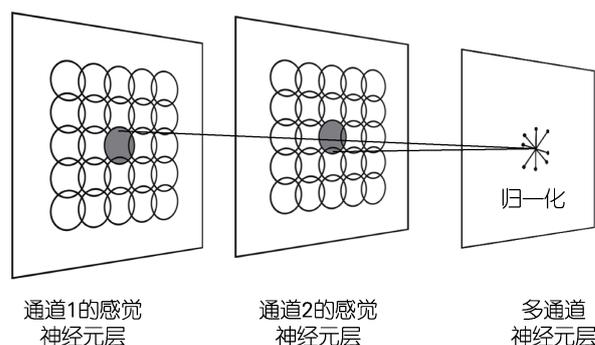


图1 多感融合的归一化模型总体架构
Fig.1 Overall structure of normalized model for multi-sensory integration

转化成多感神经元的输入, 多感神经元 Q 对来自不同通道输入的集合, 见式 (1)。

$$Q = f(q_1, q_2, \dots, q_n) \quad (1)$$

本文对刺激所引起的神经响应不做深入研究, 旨在构建信息识别结果、交互行为与认知负荷的对应关系。杨贤等^[20]采用概念外延的方式对用户认知负荷进行表达, 把用户认知负荷这个复杂对象转换成研究用户认知负荷表现因素的集合。

在虚拟环境中用户所感受的通道刺激及其用户认知表现的因素较多, 用 U 表示用户的认知负荷, 该认知负荷由刺激信号 Q 和在该刺激下的用户认知表现因素 a 组成, 其中 a 为用户的任务绩效和生理指标。因此, 多通道用户认知负荷模型如下:

$$U = [Q \ a_1 \ a_2 \ a_3 \ \dots \ a_n] \quad (2)$$

以实验中交互行为参数为输入层, 以认知负荷为输出层, 构建工作负荷量化模型 (见图 2)。

输入层: 指刺激信号及任务绩效与生理指标等活动数据。包括跟踪误差、反应时间、平均瞳孔直径等。

隐含层: 将获取的数据纳入神经网络认知负荷量化模型中进行数据处理。

输出层: 数据融合处理后的最终输出值, 即在一定条件通道下被测试者量化的认知负荷值。

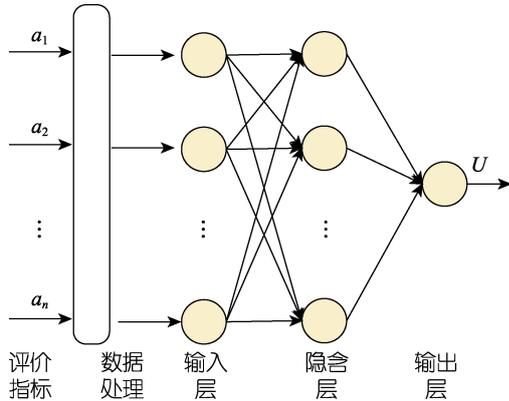


图 2 多通道交互认知负荷评价模型

Fig.2 Multi-channel interactive cognitive load evaluation model

假设每个测试值中有 n 个评价指标。则多任务矩阵如下:

$$W' = [w'_1 \ w'_2 \ w'_3 \ \dots \ w'_n] \quad (3)$$

矩阵中的值代表实验获得指标数据。各评价指标的性质不同, 所以指标之间具有不同的属性和量级, 无法直接用原始指标进行分析。为保证结果有效性, 采用 max-min 标准化, 对原始数据进行归一化处理, 使数据落到 $[0, 1]$ 区间内。转换函数如下:

$$w = \frac{w' - w_{\min}}{w_{\max} - w_{\min}} \quad (4)$$

式中, w_{\max} 为同一指标数据中最大值, w_{\min} 为最小值, w' 为原始样本值, W 为标准化结果。标准化后的矩阵如下:

$$W = [w_{11} \ w_{12} \ \dots \ w_{1n}] \quad (5)$$

标准化后的主观负荷 W 为期望输出, 本文的目标是寻找估计函数 \hat{W} 使式 (6) 表示的误差最小。

$$E_{\text{error}} = \frac{1}{2} \sum_{k=1}^n (W_k - \hat{W}_k)^2 \quad (6)$$

假设输入层指标数为 n , 隐含层节点个数为 l , 输出层认知负荷计算如式 (7) 所示, 其中 H_j 为隐含层输出值, 如公式 (8) 所示, w_{jk} 和 w_{ij} 分别为连接各层之间的权重值。

$$\hat{W}_k = \sum_{j=1}^l H_j w_{jk} \quad (7)$$

$$H_j = f \left(\sum_{i=1}^n w_{ij} x_i \right) \quad (8)$$

在本文算法中, 经过前期试验对比, 在隐含层使用 Sigmoid 函数作为激活函数收敛速度略快, 最后训练及测试结果较好, 其公式如下:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (9)$$

得到输出后, 利用期望和实际两个输出指标, 计算误差函数对输出层的各神经元的偏导数。

$$\begin{aligned} \frac{\partial E}{\partial w} &= \frac{1}{\partial w} \times \frac{1}{2} \sum_{k=1}^n (W_k - \hat{W}_k)^2 \\ &= (W_k - \hat{W}_k) \times f(x_i) \times (1 - f(x_i)) \times \frac{\partial S}{\partial w} = \delta \times x_i \end{aligned} \quad (10)$$

$$\delta = (W_k - \hat{W}_k) \times f(x) \times (1 - f(x)) \quad (11)$$

综合式 (10) ~ (12) 可利用各神经元的 δ 和下一层各神经元的输出来修正连接权值 w_{ij} 使其逼近期望输出值, 其中 η 是设置的学习率。

$$\Delta w = -\eta \cdot \frac{\partial E}{\partial w} = -\eta \cdot \delta \cdot x_i \quad (12)$$

2.2 实验设计

实验在 Unreal Engine 4 引擎搭建的场景中输出通道刺激和收集实验数据。采用 HTC Vive Pro Eye 作为输入设备。在实验中, 需要参与者去完成同时出现的两种不同任务。第一组任务为视觉跟踪任务: 一个直径为 10 mm 的圆形目标在 200 mm (宽) × 200 mm (高) 的域内随机移动。参与者用手柄发射的一个 10 mm × 10 mm 的“十字”光标来追踪圆形目标, 见图 3。

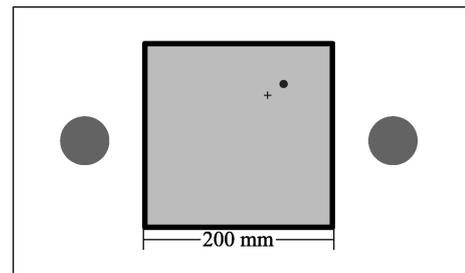


图 3 跟踪-视觉检测响应双任务实验示例

Fig.3 Example of tracking-visual detecting response dual-task experiment

第二组任务为检测响应任务^[21]。视觉,听觉和触觉刺激分别由虚拟现实头盔,耳机和手柄控制器提供。为了确保视觉、听觉、触觉刺激进行实验时,被试者感知的刺激强度是相等的,依据 Riggs 等^[22]在跨通道强度匹配研究中得到的结论,并假设其强度刺激都为 1。视觉输入为直径 30 mm 的圆形图标,位于跟踪场的左右,间距为 50 mm;从微软视窗声音库中选择的 75 分贝的单声道“叮”声作为听觉信号;使用者通过将大拇指放于手柄控制器的触摸板上感受 6.5 Hz 的振动频率作为触觉信号。被试者需要在感受到刺激后,扣动与刺激方位相同的手柄扳机作为响应,见表 1。

表 1 检测响应实验不同通道特征
Tab.1 Characteristics of different channels in detection response experiment

通道	刺激	方位
视觉 (V)	圆形图标	跟踪场的左侧;右侧
听觉 (A)	单声道声音	左耳;右耳
触觉 (T)	手柄震动	左手;右手

如表 2 所示,第二组实验的自变量为不同通道的刺激信号以及它们的组合,测试者通过不同的刺激通道完成实验。

表 2 跟踪-检测响应实验自变量
Tab.2 Independent variable of the track-detection response experiment

通道数量	任务刺激通道
单通道	视觉 (V)、听觉(A)、触觉 (T)
双通道	视听觉 (VA)、视触觉 (VT)、听触觉 (AT)
三通道	视听触觉 (VAT)

整个实验因变量包括均方根跟踪误差、响应时间和平均瞳孔直径和主观认知负荷。均方根跟踪误差(以下简称跟踪误差)通过计算被跟踪目标和跟随光标之间的像素差来衡量跟踪任务的精度,其公式如下:

$$X_{\text{RMSTE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\sqrt{(X_{\text{mod}} - X_{\text{ob}})^2 + (Y_{\text{mod}} - Y_{\text{ob}})^2})^2} \quad (13)$$

响应时间是从刺激开始到测试者做出反应的持续时间;准确性被定义为在特定条件下正确反应的数量与所有刺激数量的比率;本文使用美国国家航空 TLX 调查问卷测量认知负荷,并以 0~100 的加权平均值进行计算;实验通过 HTC 公司最新带有眼球跟踪的虚拟现实装置 HTC Vive Pro Eye 来获取任务过程中的平均瞳孔直径的变化数据。

2.3 实验流程

实验招募了 22 名参与者,其中 12 名女性和 10 名男性,年龄范围为 20~25 岁 ($M = 23.2$, $SD = 3.5$)。所有人报告听力正常,触觉没有受损,视力正常或矫

正至正常。

参与者单独进入实验室。坐在显示屏正前方 500 mm 处,佩戴集成耳机的虚拟现实显示头盔,并握住两个手柄控制器,将肘部放于椅子扶手上,大拇指放于手柄触控板上接受触觉刺激,食指在扳机上做出反应。参与者阅读并观看了任务演示之后,随机选取一个练习模块。练习后,脱下头盔休息 1 min,然后开始正式实验。

实验开始时,跟踪目标和手柄射线的十字光标位于跟踪区域中心,参与者点击手柄控制器,跟踪目标开始随机移动,实验者需要移动手柄将十字光标射线尽可能地叠加到跟踪目标上。在每个实验中随机延迟 1~3 s 后,左右两侧会给出不同通道的刺激信号,参与者用左/右食指按下对应手柄控制器扳机响应。从刺激呈现开始到反应所经过的时间被作为反应时间。1 s 后做出的反应被认为是失误。每位被试者需完成 8 次子实验,分别对应于视觉、听觉、触觉三个通道和它们之间的不同组合。实验中各子实验按随机顺序出现,抵消顺序带来的影响。在每个子实验结束后,参与者脱下虚拟现实显示头盔,在计算机上填写认知负荷调查问卷,休息 30 s,开始下一个子实验,实验场景见图 4。



图 4 跟踪-检测响应实验场景
Fig.4 Scene of track-detection response experiment

3 实验结果分析

实验一共收集了来自 22 名实验参与者的表现,包括主观认知负荷、跟踪误差、瞳孔直径和反应时间与准确率。具体数据见表 3。

3.1 跟踪误差分析

本文选用均方根跟踪误差来衡量跟踪任务的精度,跟踪误差计算被跟踪目标和跟随光标之间的像素差。总共有 176 个 (22 个参与者×8 个条件) 值,从 34.81~111.43 px 不等。

表 3 跟踪-检测响应交互实验结果
Tab.3 Results of track-detection response interaction experiment

刺激通道	跟踪误差	准确率/%	响应时间	瞳孔大小	认知负荷
无	45.85				31.82
视觉	77.21	97.83	607.50	3.73	64.13
听觉	69.39	96.12	535.39	3.67	62.32
触觉	64.97	98.12	498.12	3.65	55.10
视听	57.87	98.15	495.93	3.62	63.70
视触	55.65	97.75	469.48	3.62	61.12
触听	52.32	98.75	458.47	3.58	58.01
视听触	50.70	99.15	424.14	3.56	48.12

如图 5 所示, 无响应任务的单独跟踪误差值最低 (44.72 px), 随后是三通道响应任务的跟踪误差 (50.36 px)。单通道视觉响应任务条件下的跟踪误差最高 (77.46 px)。跟踪误差数据的方差齐性检验 ($P=0.182, P>0.05$) 显著性大于 0.05, 即方差齐性假设成立, 因此可以对数据进行下一步方差分析。

对通道进行单因素方差分析发现, 响应任务的通道变化对跟踪绩效有显著性影响 ($F=13.72, P<0.001$), 这意味着参与者的跟踪精度受到响应任务刺激变化的明显影响。对不同通道进行最小显著性差异 (Least-significant Difference, LSD) 验后多重比较检验发现, 无响应任务的跟踪误差显著小于有响应任务, 而在三种单通道条件之间, 视觉条件下的跟踪误差显著大于听觉与触觉 ($P<0.05$), 双通道和三通道的跟踪误差显著小于所有单通道跟踪误差 ($P<0.010$), 且彼此之间没有差异。

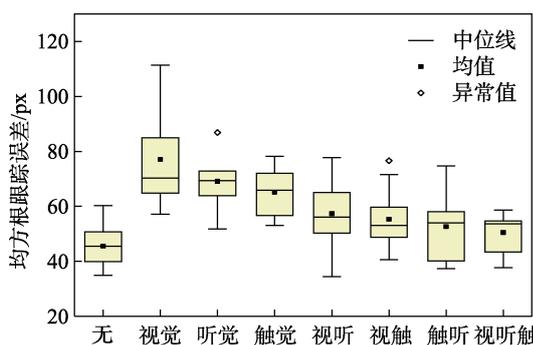


图 5 不同变量因素的均方根跟踪误差比较

Fig.5 Comparison of root mean square tracking error of different variable factors

3.2 响应时间分析

响应时间是从刺激出现到第一次按下扳机之间的时间。实验总共收集了 1 540 个反应 (22 名参与者 \times 7 个条件 \times 10 个试验)。所有反应时间均在平均反应时间的 3σ 内, 即在 166~1 495 ms, 没有要丢弃的异常值。

与跟踪误差类似, 三通道的响应任务的平均反应

时间最短 (424.14 ms), 随后为包含触觉刺激的双通道触觉单通道的反应时间也小于其他单通道反应时间。响应时间的方差齐性检验 ($P=0.284, P>0.05$) 显著性大于 0.05, 因此对数据进行下一步方差分析。

通道对反应时间有着显著的影响 ($F=6.742, P<0.001$), 视觉条件下的反应时间明显慢于其他通道 ($P<0.05$), 三通道的反应时间明显快于未包含触觉通道即视觉、听觉, 以及视听觉条件下的反应时间 ($P<0.05$), 见图 6。

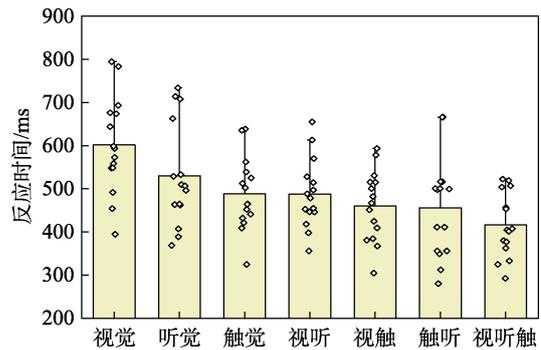


图 6 不同变量因素的反应时间比较

Fig.6 Comparison of reaction time of different variable factors

3.3 准确率分析

准确性是通过包含受试者刺激的全部试验中正确反应 (命中) 的比例来衡量的。准确性在所有条件下都非常高 ($>96\%$) (见图 7), 总的来说, 错误响应发生率为 2.01%, 两两比较显示通道之间没有差异 ($P>0.120$)。

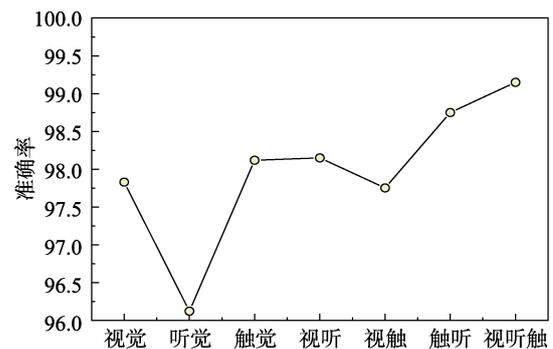


图 7 不同变量因素的准确率比较

Fig.7 Comparison of accuracy rates of different variable factors

3.4 瞳孔直径分析

瞳孔直径是用户认知负荷指标之一, 其变化幅度与实验者的努力程度相关。当认知负荷较大时, 瞳孔直径较大^[23]。通过获取受试者在任务过程中双眼瞳孔的平均直径作为分析数据。如图 8 所示, 随着通道的增加, 瞳孔直径逐渐减小, 其方差齐性检验 ($P=0.114, P>0.05$) 显著性大于 0.05, 因此对数据进行方差分析。对不同变量因素的平均瞳孔直径进行单因素方差分

析,得出不同通道对瞳孔大小有显著影响 ($F=2.82$, $P=0.036$, $P<0.05$)。

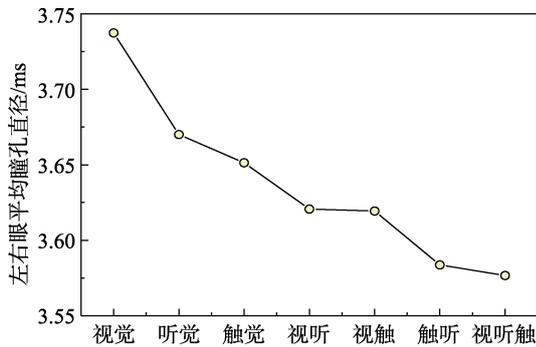


图8 不同变量因素的平均瞳孔直径比较
Fig.8 Comparison of mean pupil diameter with different variable factors

3.5 多通道认知负荷量化模型训练

表4为不同通道组合下实验指标数据(部分),数据已进行归一化处理。

表4 神经网络模型输入层数据(部分)
Tab.4 Input layer data of neural network model (partial)

样本编号	认知通道	要素				主观负荷
		通道刺激	跟踪误差	响应时间	瞳孔直径	
1		1,0,0	0.497	0.370	0.809	0.487
2	视觉	1,0,0	0.505	0.687	0.520	0.573
3		1,0,0	0.681	0.749	0.885	0.892
4		0,1,0	0.367	0.200	0.686	0.328
5	听觉	0,1,0	0.475	0.451	0.217	0.476
6		0,1,0	0.445	0.918	0.697	0.819
7		0,0,1	0.332	0.292	0.245	0.362
8	触觉	0,0,1	0.403	0.176	0.666	0.517
9		0,0,1	0.477	0.814	0.772	0.646
10		1,1,0	0.196	0.184	0.740	0.145
11	视听觉	1,1,0	0.262	0.626	0.417	0.333
12		1,1,0	0.369	0.495	0.558	0.594
13		1,0,1	0.238	0.293	0.214	0.300
14	视触觉	1,0,1	0.173	0.641	0.847	0.422
15		1,0,1	0.304	0.307	0.512	0.646
16		0,1,1	0.233	0.270	0.523	0.262
17	听触觉	0,1,1	0.236	0.344	0.504	0.496
18		0,1,1	0.285	0.692	0.327	0.706
19		1,1,1	0.239	0.142	0.180	0.095
20	视听触	1,1,1	0.239	0.399	0.237	0.222
21		1,1,1	0.178	0.291	0.665	0.379

神经网络输出隐含层神经元数量理论上没有明确的规定。多数情况下使用试凑法,公式如下:

$$m = \sqrt{n+l} + a \quad (14)$$

式中, m 为隐含层节点数, l 为输出层节点数, n 为输入层节点数, a 为 1~10 的常数。依据前期测试,当模型的节点数为 4 时网络开始收敛,代入公式(14)确定隐含层的节点数为 4~14,隐含层的神经元个数,通过试凑法确定为 8 个。选择对数 S 型函数(Log-Sigmoid)作为激活函数,训练最大次数为 2000。

BP 神经网络模型评估的负荷与实验测试主观负荷的对比,如图 9 所示,认知负荷模型评估结果和实验主观数据较为吻合,说明该模型具有一定的有效性。

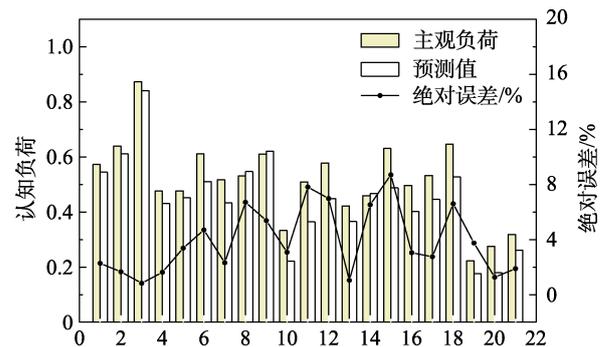


图9 BP神经网络预测值与实验主观负荷对比
Fig.9 Comparison of predicted value and expected output of BP neural network model

4 实验结果讨论

4.1 不同通道呈现信息性能讨论

现有的多重资源理论或听觉先占理论,能够解释视觉或听觉等单通道刺激对正在进行的视觉跟踪任务的干扰效果^[2]。多重资源理论认为多个视觉任务会共享相同的视觉通道资源,因此视觉刺激对进行中的跟踪任务的干扰应大于听觉刺激。而听觉先占理论认为,听觉刺激会将注意力从进行中的视觉任务上转移,增加认知负荷。因此,听觉刺激的干扰效应大于视觉刺激。

本实验通过对比不同通道信息刺激下的跟踪误差,发现任何形式的检测响应任务都会干扰同时进行的跟踪任务的表现,证明额外的任务会增加认知负载,这符合 Wickens 的多重资源理论^[6],其中视觉通道的刺激误差显著大于听觉和触觉。表明在执行视觉跟踪任务时,与视觉提示相比,使用听觉、触觉提示具有更好的性能。与跟踪误差相似,实验中的反应时间和瞳孔直径也支持多重资源理论,表明当多任务使用相同通道时,会同时损害两种任务的表现,并增加工作负荷。说明在设计虚拟现实系统中类似人机交互任务时,应避免使用同一通道进行交互,以消除或减少反应干扰。

4.2 多通道交互增强效应讨论

多通道输出在自然人机交互中是必不可少的,尤

其是在虚拟现实^[18]。本实验中观察到通过多个感官通道来传递相同的信息, 显著提高了信息传递能力, 这与 Petts 等^[8]、Sella 等^[9]研究的结论类似。相对于单通道刺激, 双通道刺激对跟踪任务干扰及反应时间和瞳孔平均直径大小显著降低, 三种通道信息呈现的组合则更低。将单通道中最优的实验数据平均值, 减去包含其多通道的平均数据时, 再除以单通道最优值, 将该值乘以 100 可以得到多通道信息组合反应增强百分比。

表 5 多通道信息增强百分比

Tab.5 Multi-channel information enhancement percentage

通道	增强百分比/%		
	跟踪误差	反应时间	瞳孔直径
视听觉	17.52	7.85	0.82
视触觉	14.24	5.03	1.89
听触觉	19.18	6.79	1.42
视听触	21.14	14.71	2.08

如表 5 所示, 虽然各项指标的增强效应不同, 但在同一指标下实验能够观察到与此前研究普遍观点相一致的现象, 即多通道的冗余信息整合能够提升用户任务表现, 降低认知负荷。

4.3 多通道认知负荷量化模型讨论

相较于^[13]提出的多维综合评估模型, 本文通过加入任务绩效、多通道信号等指标, 基于 BP 神经网络构建了多通道认知负荷量化模型。模型输出结果与实验主观负荷具有高度相关性, 表明得到的预测值是有效的。可针对影响虚拟现实多任务交互系统进行优化改进。

同时, 为了检验模型的准确性, 采用相对误差 E_1 和相对均方误差 E_2 来评价模型输出精度, 公式如下:

$$E_1 = \left| \frac{WI_k - WI_k^*}{WI_k} \right| \quad (15)$$

$$E_2 = \sqrt{\frac{1}{k} \sum_{k=1}^k \left(\frac{WI_k - WI_k^*}{WI_k} \right)^2} \quad (16)$$

其中, k 为实验人数, 评价结果见表 6。

表 6 神经网络训练结果评价

Tab.6 Evaluation of neural network training results

通道	各通道相对误差	平均相对误差	各通道均方根误差	平均均方根误差
视觉	0.036		0.091	
听觉	0.091		0.087	
触觉	0.106		0.129	
视听	0.109	0.082	0.129	0.114
视触	0.062		0.128	
触听	0.103		0.144	
视听触	0.071		0.134	

总的来说, 平均相对误差为 8.2%, 平均均方根误差为 11.4%。同时, 从评价表中可以看出, 最大相对误差为 10.9%, 最小相对误差为 3.6%, 最大均方误差为 14.4%, 最小均方误差为 8.7%。这说明本文提出的认知负荷模型具有较高的准确性。

5 结语

本文基于虚拟现实下的多通道交互和认知负荷理论, 通过分析不同信息呈现的方式对任务绩效及眼动生理反应的变化, 运用 BP 神经网络构建多通道认知负荷量化模型, 旨在为虚拟现实系统中提高交互、认知效率的设计方案提供科学的设计和评价依据。

研究发现, 信息呈现的通道类别及其数量对任务效率均有显著的影响, 信息呈现通道数量与任务绩效及生理反应呈一定程度的正相关; 多个任务使用相同通道呈现信息会损害所有任务的绩效。通过对比主观工作负荷和量化后的认知负荷, 证明本文提出的神经网络认知负荷评估模型具有一定的可靠性, 可以作为评价虚拟现实多任务交互系统的检测标准。

参考文献:

- [1] 滕健, 黄佳慧, 宫凯. 基于 BCI 的虚拟现实模拟驾驶教学系统设计[J]. 图学学报, 2020, 41(2): 217-223.
TENG Jian, HUANG Jia-hui, GONG Kai. Design of Virtual Reality Simulated Driving Teaching System Based on BCI[J]. Journal of Graphics, 2020, 41(2): 217-223.
- [2] 赵剑, 王柳, 史丽娟, 等. 面向言语康复的多模态人机交互系统[J]. 吉林大学学报(工学版), 2020, 50(4): 1478-1486.
ZHAO Jian, WANG Liu, SHI Li-juan, et al. Multimodal Human-Computer Interaction System for Speech Rehabilitation[J]. Journal of Jilin University (Engineering and Technology Edition), 2020, 50(4): 1478-1486.
- [3] 赵欣, 丁怡, 侯文军, 等. 复杂信息系统界面可用性指标体系研究[J]. 图学学报, 2018, 39(4): 716-722.
ZHAO Xin, DING Yi, HOU Wen-jun, et al. On the Usability Evaluation Index System of Complex Information System Interface[J]. Journal of Graphics, 2018, 39(4): 716-722.
- [4] STEFANIE X Q K, TSANG S N H, CHAN A H S, et al. Hand- and Foot-Controlled Dual-Tracking Task Performance Together with a Discrete Spatial Stimulus-Response Compatibility Task[J]. International Journal of Human-Computer Interaction, 2017, 33(1): 21-34.
- [5] HUANG Gao-jian, STEELE C, ZHANG Xin-rui, et al. Multimodal Cue Combinations: A Possible Approach to Designing In-Vehicle Takeover Requests for Semi-Autonomous Driving[J]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2019, 63(1): 1739-1743.

- [6] WICKENS C D. Multiple Resources and Mental Workload[J]. *Human Factors*, 2008, 50(3): 449-455.
- [7] PETIT O, VELASCO C, SPENCE C. Digital Sensory Marketing: Integrating New Technologies into Multisensory Online Experience[J]. *Journal of Interactive Marketing*, 2019, 45: 42-61.
- [8] PITTS B J, SARTER N. What You Don't Notice can Harm You: Age-Related Differences in Detecting Concurrent Visual, Auditory, and Tactile Cues[J]. *Human Factors*, 2018, 60(4): 445-464.
- [9] SELLA I, REINER M, PRATT H. Natural Stimuli from Three Coherent Modalities Enhance Behavioral Responses and Electrophysiological Cortical Activity in Humans[J]. *International Journal of Psychophysiology*, 2014, 93(1): 45-55.
- [10] CALANDRA D M, CASO A, CUTUGNO F, et al. CoWME: A General Framework to Evaluate Cognitive Workload During Multimodal Interaction[C]// *Proceedings of the 15th ACM on International Conference on Multimodal Interaction*. New York: ACM, 2013: 111-118.
- [11] 郑玲, 乔旭强, 倪涛, 等. 基于多维信息特征分析的驾驶人认知负荷研究[J]. *中国公路学报*, 2021, 34(4): 240-250.
ZHENG Ling, QIAO Xu-qiang, NI Tao, et al. Driver Cognitive Loads Based on Multi-Dimensional Information Feature Analysis[J]. *China Journal of Highway and Transport*, 2021, 34(4): 240-250.
- [12] 查先进, 黄程松, 严亚兰, 等. 国外认知负荷理论应用研究进展[J]. *情报学报*, 2020, 39(5): 547-556.
ZHA Xian-jin, HUANG Cheng-song, YAN Ya-lan, et al. Progress of Foreign Cognitive Load Theory Application Research[J]. *Journal of the China Society for Scientific and Technical Information*, 2020, 39(5): 547-556.
- [13] 卫宗敏. 面向复杂飞行任务的脑力负荷多维综合评估模型[J]. *北京航空航天大学学报*, 2020, 46(7): 1287-1295.
WEI Zong-min. A Multi-Dimensional Comprehensive Evaluation Model of Mental Workload for Complex Flight Missions[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(7): 1287-1295.
- [14] WANG Li-jing, HE Xue-li, CHEN Ying-chun. Quantitative Relationship Model between Workload and Time Pressure under Different Flight Operation Tasks[J]. *International Journal of Industrial Ergonomics*, 2016, 54: 93-102.
- [15] 柯青, 丁松云, 秦琴. 健康信息可读性对用户认知负荷和信息加工绩效影响眼动实验研究[J]. *数据分析与知识发现*, 2021, 5(2): 70-82.
KE Qing, DING Song-yun, QIN Qin. Health Information Readability Affects Users' Cognitive Load and Information Processing: An Eye-Tracking Study[J]. *Data Analysis and Knowledge Discovery*, 2021, 5(2): 70-82.
- [16] 黄晶, 杨梦婷. 考虑初始情绪的个性化驾驶负荷状态评价[J]. *中国公路学报*, 2021, 34(1): 167-176.
HUANG Jing, YANG Meng-ting. Initial Emotion-Based Evaluation of the Personalized Driving Load State[J]. *China Journal of Highway and Transport*, 2021, 34(1): 167-176.
- [17] 王欢欢, 初胜男, 顾经纬. 基于神经网络的汽车侧面造型评价方法[J]. *图学学报*, 2021, 42(4): 688-695.
WANG Huan-huan, CHU Sheng-nan, GU Jing-wei. Evaluation Method of Vehicle Side Modeling Based on Neural Network[J]. *Journal of Graphics*, 2021, 42(4): 688-695.
- [18] 徐小萍, 吕健, 金昱潼, 等. 用户认知驱动的VR自然交互认知负荷研究[J]. *计算机应用研究*, 2020, 37(7): 1958-1963.
XU Xiao-ping, LYU Jian, JIN Yu-tong, et al. Research on Natural Interactive Cognitive Load of Virtual Reality Driven by User Cognition[J]. *Application Research of Computers*, 2020, 37(7): 1958-1963.
- [19] OHSHIRO T, ANGELAKI D E, DEANGELIS G C. A Normalization Model of Multisensory Integration[J]. *Nature Neuroscience*, 2011, 14(6): 775-782.
- [20] 杨贤. 虚拟现实自然交互环境下用户认知的数学表达及其可视化评估[D]. 广州: 广东工业大学, 2018.
YANG Xian. Mathematical Expression and Visual Assessment of User Cognition in Natural Interaction of Virtual Reality Environment[D]. Guangzhou: Guangdong University of Technology, 2018.
- [21] THORPE A, INNES R, TOWNSEND J, et al. Assessing Cross-Modal Interference in the Detection Response Task[J]. *Journal of Mathematical Psychology*, 2020, 98: 102390.
- [22] RIGGS S L, SARTER N. Crossmodal Matching: The Case for Developing and Employing a Valid and Feasible Approach to Equate Perceived Stimulus Intensities in Multimodal Research[J]. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, 2019, 61(1): 29-31.
- [23] 陈晓皎, 薛澄岐, 陈默, 等. 基于眼动追踪实验的数字界面质量评估模型[J]. *东南大学学报(自然科学版)*, 2017, 47(1): 38-42.
CHEN Xiao-jiao, XUE Cheng-qi, CHEN Mo, et al. Quality Assessment Model of Digital Interface Based on Eye-Tracking Experiments[J]. *Journal of Southeast University (Natural Science Edition)*, 2017, 47(1): 38-42.