

混合现实指控界面人机工效综合评价方法研究

王 巍¹, 张靖韬^{1,2}, 陈 同¹, 朱野山^{1,2}, 瞿 珏¹

(1. 空军工程大学防空反导学院, 西安, 710051; 2. 空军工程大学研究生院, 西安, 710051)

摘要 混合现实作为一种新的人机交互方式具有广阔的应用前景, 但将混合现实技术应用于指控系统能否有效提高其人机工效水平还有待进一步研究。为了有效评估混合现实指控界面的人机工效水平, 鉴于现有界面评价方法的局限性, 提出了多指标综合评价法, 通过设计实验, 对比传统二维与混合现实显示交互系统下的任务绩效、生理反应及主观感受。通过实验采集任务完成时间、眼动数据、生理指标、脑电数据及主观评分, 对主客观数据进行归一化加权处理, 实现了对两类人机交互系统的界面人机工效的量化评估, 验证了多指标综合评价法在评估混合现实人机界面中的有效性, 为混合现实指控系统人机界面的科学评估与优化提供了实验依据与理论支持。

关键词 混合现实; 认知负荷; 界面评价; 指控系统; 操作绩效

DOI 10.3969/j.issn.2097-1915.2025.03.009

中图分类号 TP391.9 **文献标志码** A **文章编号** 2097-1915(2025)03-0069-10

Research on Comprehensive Evaluation Method of Ergonomics of Mixed Reality Charging Interface

WANG Wei¹, ZHANG Jingtao^{1,2}, CHEN Tong¹, ZHU Yeshan^{1,2}, QU Jue¹

(1. Air Defense and Antimissile School, Air Force Engineering University, Xi'an 710051, China;

2. Graduate School, Air Force Engineering University, Xi'an 710051, China)

Abstract Mixed reality, as a new human-computer interaction method, has broad application prospects, but whether the mixed reality technology can effectively improve the human-computer effectiveness level in command system remains to be researched. In order to effectively evaluate the human-computer effectiveness level of mixed reality command interface in consideration of the limitations in existing human-computer interface evaluation methods, a multi-indicator comprehensive evaluation method is proposed by designing experiments to compare the task performance, physiological responses, and subjective feelings between the traditional two-dimensional and mixed reality display and interaction systems. By collecting task completion time, eye movement data, physiological indicators, EEG data, and subjective ratings, normalization and weighted processing of both objective and subjective data are carried out to achieve quantitative evaluation of the interface human-computer effectiveness of the two human-computer interaction systems. The multi-indicator comprehensive evaluation method in evaluating the mixed reality human-computer in-

收稿日期: 2024-09-11

基金项目: 国家自然科学基金(52175282)

作者简介: 王 巍(1974—), 男, 陕西西安人, 教授, 博士生导师, 研究方向为人机工程、机器视觉。E-mail: wangwei_lhww11@163.com

通信作者: 张靖韬(2000—), 男, 山西运城人, 硕士生, 研究方向为机械工程。E-mail: zjt001003@163.com

引用格式: 王巍, 张靖韬, 陈同, 等. 混合现实指控界面人机工效综合评价方法研究[J]. 空军工程大学学报, 2025, 26(3): 69-78. WANG Wei, ZHANG Jingtao, CHEN Tong, et al. Research on Comprehensive Evaluation Method of Ergonomics of Mixed Reality Charging Interface[J]. Journal of Air Force Engineering University, 2025, 26(3): 69-78.

terface is valid, providing an experimental basis and theoretical support for the scientific evaluation and optimization of the mixed reality command system human-computer interface.

Key words mixed reality; cognitive load; interface evaluation; accusation system; operational performance

近年来,随着科技的飞速发展,混合现实(mixed reality, MR)技术正逐步扩展到军事、工业、教育等多个领域,其中混合现实指控系统作为新兴技术的代表,在提升作战效能、优化任务执行方面展现出巨大潜力^[1-3]。例如,在军事领域,美国陆军已部署 MR 战术指挥系统,实时叠加战场关键信息于士兵视野,显著提升态势感知与协同作战效率。航空航天领域,波音公司利用 MR 优化飞机装配线,实现无图纸直接操作,缩短生产周期并降低错误率,生动展现 MR 技术在提升任务效率与用户体验方面的巨大潜力。人机界面作为混合现实指控系统与操作人员之间的桥梁,其设计质量直接决定了系统的易用性、操作效率及用户满意度^[4],进而影响到整体任务的完成质量与效率^[5]。因此,对混合现实指控系统的人机界面进行科学有效的评价显得尤为重要。

当前,人机界面评价方法大致可划分为主观评价法和客观评价法^[6]。例如,陈玉等^[7]通过整合层次分析法与基于端点混合可能度函数的灰色聚类分析,构建了针对煤矿机械人机界面的综合安全评价模型。Haibo 等^[8]在 2010 年提出,结合眼动追踪数据与定量研究方法,可有效评估并优化战斗机信息界面的设计。孙林辉等^[9]创新性地采用主观量表与客观眼动数据相结合的方式,对英语学习类手机 APP 界面进行了全面评价。王大颜^[10]进一步融合了脑电数据和行为数据,实现了对指控系统人机界面的深入评估。文献^[11]提出了界面设计的布局方法,规范显示边界并验证元素项数合理性,为 MR 指控界面设计提供科学依据,有助于本研究评价框架构建,其侧重设计层面,对用户反馈与体验评价涉及不足。文献^[12]聚焦于工业机器人在教育领域的标准化教学模式,其中所涉及的 MR 指控界面设计及应用效果评价,对本研究亦具备启发性意义。

然而,尽管上述研究在人机界面评价领域取得了一定成果,但多数研究仍聚焦于传统的二维人机界面,且评价指标相对单一,难以全面反映混合现实人机界面的复杂性与多维度特性。混合现实人机界面融合了物理世界与数字信息的无缝交互,其深度沉浸感、交互自然度及多维度信息处理能力,均超越了传统二维界面的范畴,这对现有的评价体系带来了新的挑战。因此,如何对混合现实人机界面人机工效进行科学全面的评价,成为当前亟待解决的重要问题。

鉴于上述分析,本文提出融合多指标的综合评价法,将主观评价法与客观评价法的优势进行融合,通过实验实现了对 2 个人机交互系统人机界面人机工效的量化评估。研究表明,混合现实显示交互系统在提高任务完成效率、优化视觉注意分配、降低交互成本、提高操作的直观性和准确性方面均优于传统的二维显示交互系统。但混合现实指控界面带来的认知负荷也相对较高。实验结果验证了多指标综合评价法在评估混合现实指控界面人机工效中的有效性,还为混合现实指控系统人机界面的科学评估与优化提供坚实的实验证据与理论支撑。

1 实验方案

本文采用的多指标综合评价法融合了主观评价和客观评价。主观评价采用美国 AMES 研究中心 Hart 等^[13]建立的 NASA-TLX 量表评价法,这一方法应用广泛,可以实现对混合现实系统的多维度主观评估。客观评价法采用量化分析被试者在执行任务过程中的任务测试成绩及人体生理指标的方式,可以系统地分析总结评价规律,为混合现实界面优化设计提供有效的指导^[14-15]。针对混合现实指控界面沉浸感强的特点,在评价指标体系中加入了眼动数据,以直观反映操作者在视觉层面的互动情况与注意力分配;同时,为了全面把握操作者的生理与心理状态,本文还引入了脑电数据和外围生理数据,这些深度生理指标的加入,使我们能够更深层次地理解用户在混合现实环境中的真实感受与体验,从而为提升人机工效、优化界面设计提供更加精准、全面的参考依据。

为确保主观评价与客观评价的有效融合,在权重分配方面,通过邀请人机工效、人机界面评估及 MR 技术领域具备深厚经验的专家采用层次分析法(analytic hierarchy process, AHP)对各数据指标进行权重评估,通过比较各指标间的相对重要性,科学合理地确定了各指标的权重,从而确保了评价体系的严谨性。

为了验证多指标综合评价法评估混合现实指控界面人机工效水平的有效性,本文以地空导弹指控系统的人机交互过程为实验背景,设计对比实验,对操作人员在使用传统二维显示交互系统与混合现实显示交互系统完成既定任务时的主客观数据进行分

析研究。

1.1 被试者的选取

被试选取了 20 名有相关业务知识的在校研究生作为初步探索的样本,身体健康,年龄在 20~30 岁($SD=2.4$),视力或矫正视力都为 5.0,无色盲色弱。将这 20 名在校研究生按随机原则分成 A、B 组,每组 10 人。A 组为传统二维显示交互系统组,B 组为混合现实显示交互系统组,实验开始前分别对 A、B 组的每名实验人员进行 1 个小时的二维、混合现实环境适应和操作交互训练。

1.2 实验设备和场地

实验场地选址为专门的 MR 训练室,场地面积约 40 m²,场地宽敞。室内温度和湿度均恒定。

MR 显示设备选用微软的 HoloLens2 代,二维显示设备选择的是 AOC27 寸显示器,图 1 为实物照片。实验测试装置有脑电仪和眼动仪,脑电仪型号为 Nuamps7181,眼动仪型号为 Tobii Pro Glasses 2,其采样率为 50 Hz 或 100 Hz,图 2 为实物图片。

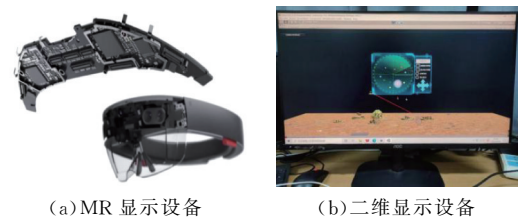


图 1 显示设备实物
Fig. 1 Display the device

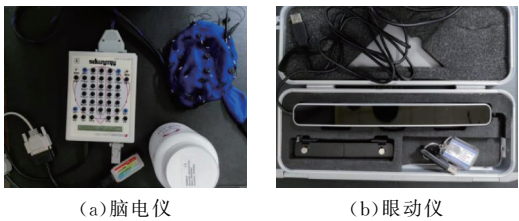


图 2 实验测试装置
Fig. 2 Experimental test equipment

1.3 实验任务场景设计

本实验参考已有文献中的防空反导作战场景^[16-18],设计了防空作战的 6 种典型任务场景,见表 1,旨在模拟指挥员在防空任务中的指挥操作流程。

为确保实验结果的可靠性和有效性,实验充分考虑了任务场景的复杂度和任务难度。具体而言,首先设定了一个包含红蓝双方兵力部署、地形地貌等关键要素的作战沙盘作为基础场景,并在此基础上根据防空作战的不同阶段和指挥需求细化了任务场景。任务设计遵循由简至繁的原则,逐步增加操作复杂度和信息干扰,以考察实验者的操作精度、反应速度以及在复杂信息环境下的判断和决策能力。同时,采用随机分组、任务标准化以及详细的数据记

录与分析方法,以确保 2 组实验任务在难度和复杂度上的一致性。具体的实验操作任务涵盖观察整体与局部态势、探测装备的交互操作与态势显示、武器装备的架设与发射以及信息显示状态的选择等多个方面,可以代表防空任务中指挥员指挥操作的一般过程。

表 1 实验操作分任务界面示意图

Tab. 1 Diagram of the task-based interface for experimental operation

任务场景	传统二维显示交互系统	MR 显示交互系统
任务 1		
任务 2		
任务 3		
任务 4		
任务 5		
任务 6		

具体的实验操作任务为:

任务 1:观察整个指挥沙盘;

任务 2:打开防空区域的雷达;

任务 3:观察红方无人机群发现蓝方飞机后颜色和队形变化的全过程;

任务 4:控制红方防空区域的导弹车进行架设准备;

任务 5:控制红方防空区域的导弹车发射导弹击落临近的低空突袭的蓝方飞机;

任务 6:勾选加载红方地形、显示蓝方情报、加载地标、雷达显示。

1.4 实验操作步骤

被试者准备完成后按如下步骤开始实验:

步骤 1 被试者熟悉整个实验的操作流程和操

作方法,并于实验前进行预演,避免实验过程中出现学习效应。

步骤 2 实验前确保被试者精神状态放松,操作状态良好。

步骤 3 将被试者分成 A、B 组,按照任务 1~任务 6 的顺序分别在 2 个不同的显示交互系统上进行实验。

步骤 4 被试者在工作人员引导下进行实验,被试人员完成后举手示意,工作人员记录被试者从开始操作任务到举手的间隔时间,将其作为该项操作任务的完成时间,以此循环直到整个实验任务全部完成。

实验过程在严格控制任务流程和效果一致性的条件下进行,通过收集并分析 2 种显示交互系统在执行相同任务时的绩效与认知负荷差异,便可清晰揭示二者在同一任务尺度下的性能优劣及其对用户认知负担的影响。

1.5 实验数据采集

采集的实验数据指标包括:任务完成时间、眼动数据、外围生理数据、脑电数据、主观评价量表打分。其中眼动数据包含了平均注视时间、瞳孔直径、眨眼频率、注视点个数、首次注视时间,外围生理数据有心率、血压和血氧饱和度,脑电数据采集的是被试者完成任务时 P300 的波幅与潜伏期。

采集数据时二维显示设备选择的是 AOC27 寸显示器,分辨率为 $1\,920 \times 1\,080$ dpi,屏幕刷新率为 75 Hz,屏幕比例为 16:9,亮度调整为 200 cd/m^2 ;脑电仪模拟输入为 40 导单极,采样频率为 100、250、500、1 000 Hz,输入范围为 $\pm 130\text{ mV}$,输入阻抗不低于 $80\text{ M}\Omega$,共模抑制比为 100 dB (50/60 Hz 时),输入噪声为 $0.7\text{ }\mu\text{V}$;眼动仪采样率为 50 Hz 或 100 Hz。

主观数据通过主观量表法进行测量,主观评价量表涵盖了脑力需求、体力需求、时间需求、努力程度、绩效水平及受挫程度 6 大维度,每个维度均设有 20 个等距刻度,每刻度对应 5 分,刻度数量的增加直接反映了该维度负荷的加剧。最终,量表所得的认知负荷值是通过将各维度分数与其相应权重相乘后求和而得出的结果。权重的确定过程涉及被试者对 6 个维度进行重要性的两两比较,记录每次比较中被试者认为更为重要的维度。经过总计 15 轮次成对比较后,统计各维度被选为更重要的次数,并将该次数除以 15,所得结果即为该维度的权重。此方法确保了权重分配的客观性与合理性,为量表认知负荷值的计算提供了科学依据。

则 2 种显示交互系统第 i 个维度的 NASA-

TLX 评价分数如下:

$$W_i = \frac{N_i}{N}$$

$$N = \sum_{i=1}^6 N_i, i=1,2,\dots,6 \quad (1)$$

式中: N_i 为第 i 个维度被实验者选择为重要维度的次数。

2 种显示交互系统的 NASA-TLX 评价总分数如下:

$$W = \sum_{i=1}^6 W_i R_i \quad (2)$$

式中: R_i 为第 i 维度的评分; W_i 为第 i 维度的权重。

被试者使用 2 个显示交互系统完成任务所采集的主客观实验数据分别如表 2~表 6 所示。

数据采集完成后,首先对原始数据进行预处理,包括去除噪声、滤波、基线校正等步骤,以确保数据的准确性和可靠性。然后,利用数据分析软件对处理后的数据进行深入分析,提取关键特征和信息。

2 实验数据分析

为全面评估 2 个人机交互系统的界面人机工效,需对实验数据进行分析处理。对数据的处理分析包含对单个指标的分析和对主客观整体数据的归一化处理分析。数据归一化处理的步骤包括:首先,对主客观数据归一化处理以确保数据的一致性和可比性;其次,运用层次分析法精确界定各数据指标的权重^[19],依据所确定的权重对 2 个不同显示交互系统实验的各项数据指标进行加权评分;最后,将各系统所有数据指标的加权得分进行累加,得出各系统的加权总分,从而实现对 2 个人机交互系统人机界面性能的全面量化评估。

2.1 任务完成时间

如图 3 所示,在任务 1(即观察整个沙盘环境)的执行过程中,2 个显示交互系统之间的差异并不显著,然而,从任务 2 至任务 6 的数据图中,可以明确观察到 MR 显示交互系统在时间绩效上展现出了显著优势,其提升幅度最低达到了 77.4%,而最高则实现了 293.3% 的显著提升。可能的原因为任务 1 相对简单,主要是观察整个沙盘环境,不涉及复杂的交互或决策过程。然而,从任务 2 至任务 6,随着任务复杂度的增加,MR 显示交互系统展现出了显著的时间绩效优势,这种优势可能源于 MR 系统提供的沉浸式三维环境,使得操作人员能够更直观地理解任务环境,减少了对二维平面信息的转换和解释时间。

由表 2 数据可知,传统二维显示交互系统完成任务的总平均时间 194.7 s 高于使用 MR 显示交互

系统完成任务的总平均用时 86.2 s。对 2 种显示交互系统完成任务的平均用时进行配对 T 检验结果为 $t(9)=20.096, p<0.001$;传统二维显示交互系统单样本 T 检验结果为 $t(9)=42.638, p<0.001$,MR 显示交互系统单样本 T 检验结果 $t(9)=54.322, p<0.001$ 。由检验结果可知,不同的显示交互系统对于任务完成时间有显著影响,MR 显示交互系统相较于传统二维显示交互系统在提升任务完成速度方面有显著优势。

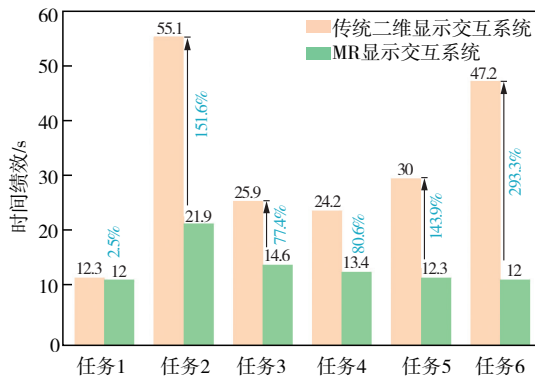


图 3 基于表 2 的传统二维与 MR 显示交互系统任务完成平均用时柱状图

Fig. 3 Bar chart of average task completion time for traditional 2D and MR display interaction systems based on Table 2

表 2 使用传统二维显示交互系统与 MR 显示交互系统完成任务的平均用时

Tab. 2 The average time taken to complete tasks using traditional 2D display interactive systems versus MR display interactive systems

任务	传统二维显示交互系统/s	MR 显示交互系统/s
任务 1	12.3	12.0
任务 2	55.1	21.9
任务 3	25.9	14.6
任务 4	24.2	13.4
任务 5	30.0	12.3
任务 6	47.2	12.0
总任务	194.7	86.2

2.2 眼动数据

由表 3、图 4 可知传统二维显示交互系统的各项眼动数据均低于 MR 显示交互系统,这表明 MR 系统能够更有效地吸引和维持操作人员的注意力。将 2 种显示交互系统的平均注视时间、瞳孔直径、眨眼频率、注视点个数、首次注视时间分别进行配对 T 检验,结果如表 3 所示, T 检验结果均小于 0.05,说明 2 种显示交互系统的 eye 数据存在显著差异,眼动数据的差异进一步支持了 MR 显示交互系统在提升任务绩效方面的优势。

平均注视时间的差异表明,MR 系统使操作人员能够更快地锁定关键信息,减少了对无关信息的

注视时间。瞳孔直径的差异可能反映了 MR 系统在提供视觉刺激方面的优势,因为更大的瞳孔直径通常与更高的视觉兴趣和注意力集中程度相关。眨眼频率的差异则可能表明,在 MR 环境中,操作人员的眼睛更加放松,减少了因长时间注视二维屏幕而产生的疲劳感。

表 3 传统二维显示交互系统与 MR 显示交互系统眼动特征的平均数据

Tab. 3 Average eye-tracking data for traditional 2D display interactive systems versus MR display interactive systems

眼动特征	传统二维显示交互系统	MR 显示交互系统	配对 T 检验
平均注视时间/ms	398.42	404.22	$t(9)=-2.505, p<0.05$
瞳孔直径/mm	3.32	3.5	$t(9)=-5.122, p<0.05$
眨眼频率/(次·min ⁻¹)	9.88	10.73	$t(9)=-9.613, p<0.05$
注视点个数	43	50	$t(9)=-5.422, p<0.05$
首次注视时间/ms	100	130	$t(9)=-18.01, p<0.05$

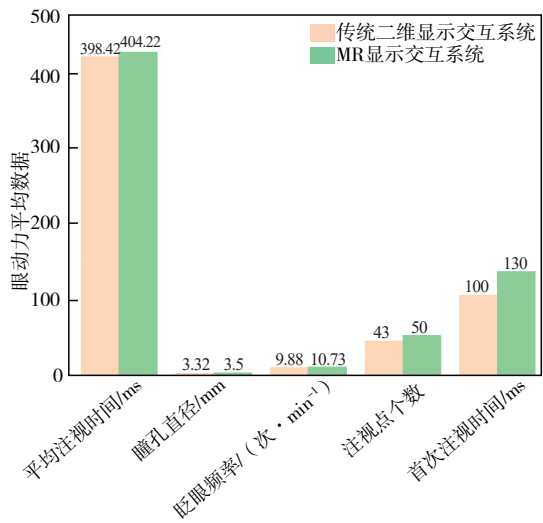


图 4 基于表 3 的传统二维与 MR 显示交互系统眼动特征平均数据柱状图

Fig. 4 Barchart of average eye movement characteristics data for traditional 2D and MR display interaction systems based on table 3

2.3 外围生理数据

表 4、图 5 中收缩压、舒张压和血样饱和度 3 项数据在 2 个不同的显示交互系统的实验中数据值相差不显著,但三者均比实验前高。这反映了实验对操作人员的生理状态产生了一定影响。值得注意的是,使用 MR 显示交互系统的实验中心率上升更高,这可能与 MR 系统提供的沉浸式体验导致的更高心理负荷有关。然而,这种心率上升可能并不完

全是不利的,因为它也可能反映了操作人员对 MR 环境的更高投入度和参与度。

表 4 传统二维显示交互系统与 MR 显示交互系统的外围生理特征数据

Tab. 4 Peripheral physiological characteristics data of traditional two-dimensional display interaction system and MR display interaction system

外围生理特征	实验前	传统二维显示交互系统	MR 显示交互系统
心率/(次·min ⁻¹)	70	80	86
收缩压/mmHg	140	143	144
舒张压/mmHg	62	65	65
血氧饱和度/%	97.2	98.3	98.2

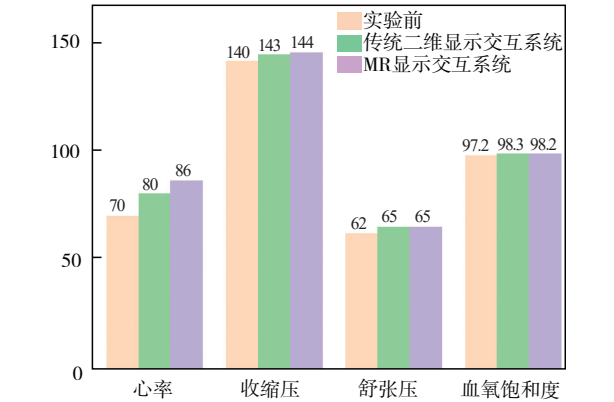


图 5 基于表 4 的传统二维与 MR 显示交互系统外围生理特征数据柱状图

Fig. 5 Barchart of ceripheral chysiological characteristics cata for traditional 2D and MR tisplay interaction systems based on table 4

2.4 脑电数据

本实验选取如图 6 所示的中央区至顶区 9 个电极点(C3、CZ、C4、CP3、CPZ、CP4、P3、PZ、P4)作为 P300 成分的分析电极。统计被试者的 P300 脑电数据得出表 5、图 7 所示。对不同显示交互系统下的波幅和潜伏期进行平均值计算,分别为 1.84 和 4.26、353.5 和 399。将二者的 P300 波幅和潜伏期分别进行配对 *T* 检验,显著性结果分别为 $t(9)=-8.322, p<0.05, t(9)=-5.226, p<0.05, T$ 检验结果均小于 0.05,说明 2 种显示交互系统的脑电数据的波幅和潜伏期均具有显著差异。

P300 波幅和潜伏期的差异表明,MR 显示交互系统对操作人员的认知加工过程产生了显著影响。具体而言,MR 系统下的 P300 波幅更大,潜伏期更短,这通常与更快的刺激识别和更高的认知资源投入相关。这种差异可能源于 MR 系统提供的丰富环境信息和交互手段,使得操作人员能够更快地识别和响应任务相关刺激。

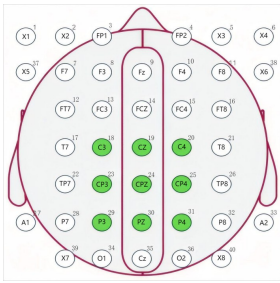


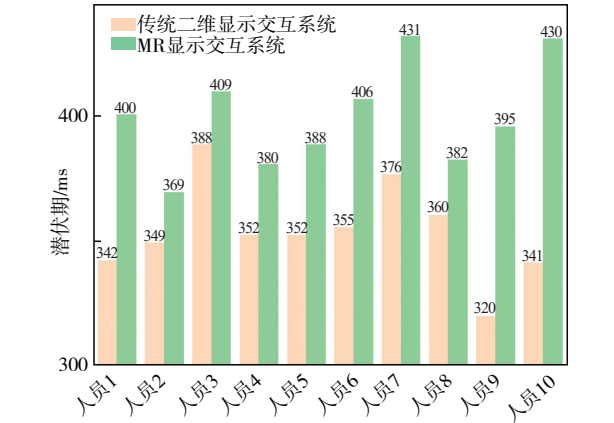
图 6 电极位置排布

Fig. 6 Electrode position arrangement

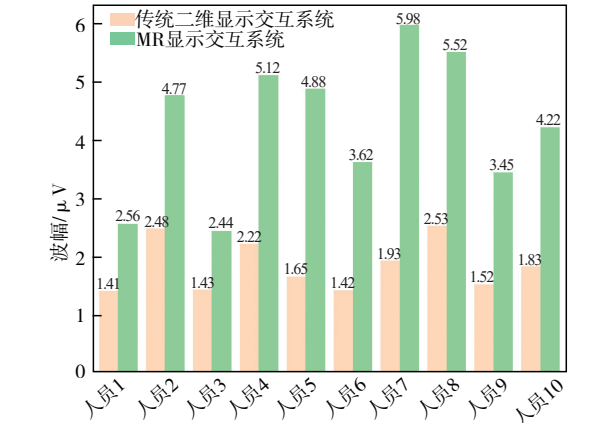
表 5 P300 波幅与潜伏期

Tab. 5 P300 amplitude and latency

组号	波幅/ μV		潜伏期/ms	
	传统二维显示交互系统	MR 显示交互系统	传统二维显示交互系统	MR 显示交互系统
1	1.41	2.56	342	400
2	2.48	4.77	349	369
3	1.43	2.44	388	409
4	2.22	5.12	352	380
5	1.65	4.88	352	388
6	1.42	3.62	355	406
7	1.93	5.98	376	431
8	2.53	5.52	360	382
9	1.52	3.45	320	395
10	1.83	4.22	341	430



(a)潜伏期



(b)波幅

图 7 基于表 5 的传统二维与 MR 显示交互系统 P300 波幅与潜伏期数据柱状图

Fig. 7 Bar chart of P300 amplitude and latency data for traditional 2D and MR display interaction systems based on table 5

2.5 主观数据评价

主观数据评价的结果进一步支持了 MR 显示交互系统在提升操作人员认知和体验方面的优势。如表 6、图 8 所示,传统二维显示交互系统和混合现实显示交互系统的主观认知负荷平均值分别为 40.68 和 47.68,对其进行配对 T 检验: $t(9) = -8.566, p < 0.001$,T 检验结果小于 0.05,说明 2 个显示交互系统具有显著差异。MR 显示交互系统的得分高于传统二维显示交互系统,这表明尽管 MR 系统可能带来更高的心理负荷,但这种负荷是积极的、有益的,因为它与更高的任务绩效和更好的用户体验相关。

表 6 主观认知负荷

Tab. 6 Subjective cognitive load

组号	传统二维显示交互系统	MR 显示交互系统
1	42.10	51.30
2	39.10	48.10
3	37.60	46.50
4	43.50	49.20
5	40.50	45.90
6	41.00	43.50
7	38.20	44.20
8	42.00	46.30
9	42.60	54.60
10	40.20	47.20
平均数	40.68	47.68

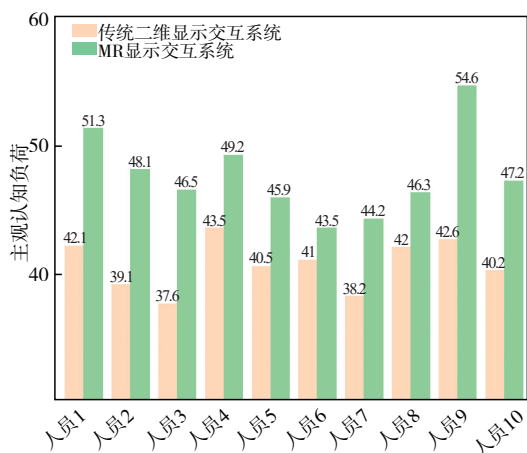


图 8 基于表 6 的传统二维与 MR 显示交互系统主观认知负荷数据柱状图

Fig. 8 Bar chart of subjective cognitive load data for traditional 2D and MR display interaction systems based on table 6

$$\mathbf{X} = \begin{bmatrix} 194.7 & 398.42 & 3.32 & 9.88 & 43 & 100 & 80 & 1.84 & 353.5 & 40.68 \\ 86.2 & 404.22 & 3.5 & 10.73 & 50 & 130 & 86 & 4.26 & 3.99 & 47.68 \end{bmatrix}$$

数据归一化处理: $\mathbf{X}' = (x'_{ij})_{2 \times 10}$

$$\mathbf{X}' = \begin{bmatrix} 0.693 & 1 & 0.496 & 4 & 0.486 & 8 & 0.479 & 4 & 0.462 & 4 & 0.434 & 8 & 0.481 & 9 & 0.301 & 6 & 0.469 & 8 & 0.460 & 4 \\ 0.306 & 9 & 0.503 & 6 & 0.513 & 2 & 0.520 & 6 & 0.537 & 6 & 0.565 & 2 & 0.518 & 1 & 0.698 & 4 & 0.530 & 2 & 0.539 & 6 \end{bmatrix}$$

对归一化后的数据进行求解加权计算得:

$$y_{ij} = x'_{ij} w_i \quad (6)$$

2.6 数据归一化处理

对所测得的主客观数据进行统计处理,结果如表 7 所示。

表 7 主客观数据

Tab. 7 Subjective and objective data

数据指标	传统二维显示交互系统	MR 显示交互系统
任务绩效/s	194.7	86.2
平均注视时间/ms	398.42	404.22
瞳孔直径/mm	3.32	3.5
眨眼频率/(次·min ⁻¹)	9.88	10.73
注视点个数	43	50
首次注视时间/ms	100	130
心率/(次·min ⁻¹)	80	86
脑电 P300 波幅/ μV	1.84	4.26
脑电 P300 潜伏期/ms	353.5	399
主观评价	40.68	47.68

建立主客观数据矩阵 $\mathbf{X} = (x_{ij})_{m \times n}$, 表示第 i 个方案下的第 j 个指标的数据:

$$\mathbf{X} = (x_{ij})_{m \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \quad (3)$$

对数据进行归一化处理:

$$x' = x'_{ij} = \begin{cases} \frac{X_{ij}}{\sum_{i=1}^m X_{ij}} \in [1, n], \\ \text{the larger the value, the better} \\ \frac{1 - X_{ij}}{\sum_{i=1}^m (1 - X_{ij})} \in [1, n], \\ \text{the smaller the value, the better} \end{cases} \quad (4)$$

得归一化矩阵 $\mathbf{X}' = (x'_{ij})_{m \times n}$ 。

$$\mathbf{X}' = (x'_{ij})_{m \times n} = \begin{bmatrix} x'_{11} & \cdots & x'_{1n} \\ \vdots & & \vdots \\ x'_{m1} & \cdots & x'_{mn} \end{bmatrix} \quad (5)$$

式中: $0 < x'_{ij} < 1, \sum_{i=1}^m x'_{ij} = 1$ 。

将数据带矩阵得 $\mathbf{X} = (x_{ij})_{2 \times 10}$ 。

各数据指标的权重是通过邀请 12 名人机工效、人机界面评估和混合现实技术方面的专家进行指标

权重评价,运用层次分析法经过统计、分析和计算得到的,各数据指标权重如表 8 所示。

表 8 各数据指标权重
Tab. 8 Weight of each data index

数据指标	任务绩效	平均注视时间	瞳孔直径	眨眼频率	注视点个数	首次注视时间	心率	脑电 P300 波幅	脑电 P300 潜伏期	主观评价
权重/%	14.61	11.34	12.26	11.36	3.56	2.96	8.56	10.32	10.32	14.71

各数据指标加权量化结果如下:

$$Y_{ij} = \begin{bmatrix} 0.101\ 3 & 0.056\ 3 & 0.059\ 7 & 0.054\ 5 & 0.016\ 5 & 0.012\ 9 & 0.041\ 3 & 0.031\ 1 & 0.048\ 5 & 0.067\ 7 \\ 0.044\ 8 & 0.057\ 1 & 0.062\ 9 & 0.059\ 1 & 0.019\ 1 & 0.016\ 7 & 0.044\ 3 & 0.072\ 1 & 0.054\ 7 & 0.079\ 4 \end{bmatrix}$$

对 2 个不同显示交互系统实验的每一项数据指标的加权得分进行求和,计算每个系统的加权总分:

$$T_i = 200 \times \sum_{j=1}^n y_{ij} \tag{7}$$

经计算,传统二维显示交互系统的总得分为 97.92,MR 显示交互系统的总得分为 102.08。表明在完成相同任务的情况下,MR 显示交互界面表现优于传统二维显示交互界面。综合绩效、认知负荷、视觉疲劳和主观评价方面,在完成相同操作任务时使用 MR 显示交互系统是更优的选择。

3 讨论

在 MR 环境中,MR 显示交互系统通过提供更加真实、直观的三维环境,降低了操作人员的认知负荷,使他们能够更快地理解任务要求并作出相应操作。同时,MR 显示交互系统的交互方式更为自然,减少了传统二维界面中常见的视角切换和缩放操作,从而提高了任务完成速度。因此,对 MR 系统导致的认知负荷变化进行细致分析,对于评估其实际影响及可接受性至关重要。

在 HoloLens2 代构建 MR 环境下进行实验时,操作人员的双眼需要不断进行调节,导致眼睛具有更高的视觉疲劳,同时,MR 显示交互系统提供的丰富视觉信息和更高的沉浸感导致操作人员需要更多的视觉注意来处理 and 解读信息。瞳孔直径的增加反映了操作人员在高认知负荷下的生理反应。MR 环境中视觉刺激的增强和注意力集中度的提高导致了眨眼频率的上升。注视点个数和首次注视时间的增加则表明 MR 显示交互系统在引导操作人员视线方面更为有效,使得他们能够更快地锁定关键信息点。因此,在完成相同任务的情况下,MR 显示交互系统的认知负荷和视觉疲劳高于传统二维显示交互系统。

在 2 种系统操作实验过程中,实验者都表现出一定的压力和紧张情绪。虽然收缩压、舒张压和血氧饱和度在 2 种系统间无显著差异,但心率数据显

著表明,在 MR 系统实验中,操作人员的心率上升更为显著,这反映了 MR 系统带来的更高认知负荷和紧张感。尽管 MR 显示交互系统提高了任务完成效率,但其丰富的视觉和交互体验也导致操作人员需要更多的心理和生理资源来应对。这种生理反应在一定程度上验证了 MR 显示交互系统在提升任务性能的同时,也要求操作人员使用过程中具有更高的认知负荷。

脑电数据的分析进一步揭示了 2 种显示交互系统在认知加工过程上的差异。具体而言,MR 显示交互系统引发的 P300 脑电成分的平均波幅是传统显示交互系统的 2.31 倍,而其平均潜伏期也延长了 1.13 倍。这些数据表明,MR 显示交互系统下的认知加工过程更为复杂且深入,需要更多的时间和认知资源来完成。相比之下,传统的二维显示交互系统在认知负荷方面显著低于 MR 显示交互系统。

主观评价量表的结果显示,MR 显示交互系统的认知负荷得分显著高于传统二维系统 ($p < 0.001$)。然而,这一结果并不直接表明 MR 系统带来了更高的认知负荷,而是可能反映了操作人员对 MR 系统更高性能、更直观操作和更丰富体验的认可。主观评价中的高认知负荷得分可能源于操作人员对 MR 系统复杂性和新颖性的感知。尽管 MR 系统提高了任务完成效率和操作体验,但其丰富的信息和交互方式也可能让操作人员感到更加具有挑战性。因此,在评价过程中,他们可能倾向于给予更高的认知负荷评分。然而,这种主观感受并不完全等同于实际的认知负荷水平

综上可以得出结论,MR 显示交互系统在提高任务完成效率、优化视觉注意分配、降低交互成本以及增强操作的直观性和准确性方面相较于传统的二维显示交互系统展现出明显优势。尽管有研究表明,使用 MR 显示交互系统可能会导致认知负荷有一定程度的增加,但这种增加是基于系统提供更丰富、更直观的信息和交互方式所带来的必然结果。在实际应用中,操作人员通常能够适应并接受这种程度的认知负荷增加,因为它往往伴随着任务执行

效率和操作体验的提升。因此,在需要完成复杂操作任务的场景下,MR显示交互系统通常被视为一个更为优越的选择。

为了进一步优化用户体验,提出简化界面布局,减少不必要的视觉元素和交互步骤,以降低操作人员的认知负担;将复杂任务分解为更小的子任务,并提供清晰的指导和反馈,以帮助操作人员更好地理解 and 执行任务以降低认知负荷。

4 结语

本文以地空导弹指控系统典型任务为实验背景,采用主观与客观评价法相结合的多维度综合评价法,对传统二维显示交互系统与混合现实显示交互系统的指控界面进行了对比实验。通过对2类不同显示交互系统实验环节中各项关键指标数据的深入分析,探讨了用户在不同系统下认知负荷的动态变化。实验结果表明,混合现实显示交互系统在提升用户体验及操作效率、信息呈现丰富度及用户满意度等性能方面优于传统二维显示交互系统。然而,其伴随的认知负荷也相对较高。这一结果验证了多指标综合评价法在评估混合现实指控界面人机工效方面的有效性,为进一步优化用户交互界面设计提供了重要的参考依据。

尽管本研究已取得初步成果,但仍存在若干局限性与不足之处。首先,实验所采用的研究样本具有局限性,这可能对研究结论的普遍适用性构成潜在影响。为增强研究的可靠性和推广价值,未来的研究将增强样本容量和多样性以提高研究结论的普适性,包括增加样本容量、拓宽年龄范围、丰富专业背景,以覆盖更广泛的用户群体。其次,本研究聚焦于部分主客观评价指标,未能全面覆盖所有相关维度,特别是情感反应和学习曲线等关键指标。情感反应作为用户体验的重要组成部分,直接影响用户的使用体验和长期忠诚度;而学习曲线则反映了用户从初次接触到熟练掌握界面所需的时间和努力,是评估界面易用性和学习友好性的重要依据。因此,后续研究应进一步拓宽评价指标体系,将情感反应和学习曲线等关键指标纳入其中,以实现混合现实指控界面人机工效更为全面而深入的评估,为界面设计的优化提供更为丰富的数据支持和理论依据。

参考文献

[1] 邢祥莹. MR技术的发展与现状及其在教育中的应用[J]. 中国现代教育装备, 2024(6): 31-33.

XING X Y. Development and Present Situation of MR Technology and Its Application in Education [J]. China Modern Educational Equipment, 2024 (6): 31-33. (in Chinese)

[2] 孙靖杰, 郭仲平, 王广会. 混合现实技术在军事领域的应用探析[J]. 仪表技术, 2024(4): 60-62.

SUN J J, GUO Z P, WANG G H. Analysis on the Application of Hybrid Reality Technology in the Military Field[J]. Instrumentation Technology, 2024(4): 60-62. (in Chinese)

[3] SOUCHET A D, PHILIPPE S, LÉVÊQUE A, et al. Short and Long-Term Learning of Job Interview with a Serious Game in Virtual Reality: Influence of Eye-strain, Stereoscopy, and Apparatus[J]. Virtual Reality, 2022, 26(2): 583-600.

[4] 李龙跃, 贾忠慧, 皮雳, 等. 美军杀伤网的概念内涵、发展现状与趋势[J]. 航空兵器, 2024, 31(5): 11-18.

LI L Y, JIA Z H, PI L, et al. Conceptual Connotation, Development Status and Trend of US Army Kill Network[J]. Aviation Weapons, 2024, 31(5): 11-18. (in Chinese)

[5] 肖彬, 李琳琳, 刘文. 面向认知的指挥控制系统界面设计方法[J]. 指挥与控制学报, 2024, 10(2): 244-249.

XIAO B, LI L L, LIU W. Cognitive Oriented Command and Control System Interfaces Design[J]. Journal of Command and Control, 2024, 10(2): 244-249. (in Chinese)

[6] 陈睿欣. 特种车辆人机界面布局优化与评估方法研究[D]. 大连: 大连理工大学, 2019.

CHEN R X. Research on Optimization and Evaluation Method of Man-Machine Interface Layout for Special Vehicles[D]. Dalian: Dalian University of Technology, 2019. (in Chinese)

[7] 陈玉, 王志远, 耿雷, 等. 基于灰色聚类分析的煤矿机械人机界面安全综合评价[J]. 煤炭技术, 2023, 42(7): 144-148.

CHEN Y, WANG Z Y, GENG L, et al. Comprehensive Evaluation of Safety of Coal Mine Machinery Human-Machine Interface Based on Grey Cluster Analysis[J]. Coal Technology, 2023, 42(7): 144-148. (in Chinese)

[8] WANG H B, XUE C Q, LIU Q. The Eye Movement Experiment and the Usability Evaluation of the Fighter Cockpit Digital Interface[C]//2010 2nd International Conference on Information Engineering and Computer Science. Wuhan: IEEE, 2010: 1-4.

[9] 孙林辉, 韩贝贝, 张伟. 基于眼动实验的英语学习类手机APP界面设计评价[J]. 人类工效学, 2021, 27(2): 1-8.

SUN L H, HAN B B, ZHANG W. Evaluation of Mobile APP Interface Design for English Learning Based on Eye Movement Experiment[J]. Chinese Journal of

- Ergonomics, 2021, 27(2): 1-8. (in Chinese)
- [10] 王大颜, 瞿珏. 基于脑电指标的指控系统人机界面评价研究[J]. 兵器装备工程学报, 2021, 42(11): 196-203.
- WANG D Y, QU J. Research on Human-Computer Interface Evaluation of Allegation System Based on EEG Indicators[J]. Journal of Ordnance Equipment Engineering, 2021, 42(11): 196-203. (in Chinese)
- [11] 王玮, 路易霖, 肖利伟, 等. 基于 MR 混合现实技术的界面布局方法及系统研究[J]. 工程机械, 2024, 55(1): 103-112, 9-10.
- WANG W, LU Y L, XIAO L W, et al. Research on Interface Layout Method and System Based on Mixed Reality Technology[J]. Construction Machinery and Equipment, 2024, 55(1): 103-112, 9-10. (in Chinese)
- [12] 刘翠莲. 基于 MR(混合现实)工业机器人标准化教学模式与效果评价研究[J]. 中国设备工程, 2023(22): 30-32.
- LIU C L. Research on Standardized Teaching Mode and Effect Evaluation of Industrial Robots Based on MR (Mixed Reality) [J]. China Plant Engineering, 2023(22): 30-32. (in Chinese)
- [13] HART S G, STAVELAND L E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research[M]// Human Mental Workload. Amsterdam: Elsevier, 1988: 139-183.
- [14] 孙建华, 蒋婷, 王春慧, 等. 航天器软件人机界面工效评价指标与评价方法研究[J]. 载人航天, 2020, 26(2): 208-213.
- SUN J H, JIANG T, WANG C H, et al. Research on Ergonomic Evaluation Indicator and Methodology for Interactive Interface of Spacecraft Software[J]. Manned Spaceflight, 2020, 26(2): 208-213. (in Chinese)
- [15] 郑弋源. 民用飞机驾驶舱人机交互机理与评价方法研究[D]. 上海: 上海交通大学, 2015.
- ZHENG Y Y. Research on Human-Computer Interaction Mechanism and Evaluation Method in Civil Aircraft Cockpit[D]. Shanghai: Shanghai Jiao Tong University, 2015. (in Chinese)
- [16] 刘家义, 王刚, 夏智权, 等. 基于深度强化学习的防空反导智能任务分配[J]. 火力与指挥控制, 2024, 49(1): 43-48, 55.
- LIU J Y, WANG G, XIA Z Q, et al. Intelligent Task Assignment Research for Air Defense and Anti-Missiles Based on Deep Reinforcement Learning[J]. Fire Control & Command Control, 2024, 49(1): 43-48, 55. (in Chinese)
- [17] 张锦, 王刚, 郭相科, 等. 防空反导作战杀伤链/网构建与发展研究分析[J/OL]. 现代防御技术. (2024-08-26) [2024-09-05]. <https://link.cnki.net/urlid/11.3019.TJ.20240806.1006.002>
- ZHANG J, WANG G, GUO X K, et al. Research and Analysis on the Construction and Development of Kill Chain/Network in Air Defense and Antimissile Operations[J/OL]. Modern Defense Technology. (2024-08-26) [2024-09-05]. <https://link.cnki.net/urlid/11.3019.TJ.20240806.1006.002>. (in Chinese)
- [18] 方峰, 李清伟, 贺嘉璠, 等. 自适应作战场景的反导系统火力分配策略[J]. 指挥信息系统与技术, 2023, 14(3): 31-38.
- FANG F, LI Q W, HE J F, et al. Adaptive Weapon-Target Assignment Strategy for Missile Defense System in Different Scenarios[J]. Command Information System and Technology, 2023, 14(3): 31-38. (in Chinese)
- [19] 杨传福. 层次分析法和模糊加权的装配式工程管理绩效评估[J]. 九江学院学报(自然科学版), 2022, 37(4): 31-34.
- YANG C F. Performance Evaluation of Assembled Engineering Management Based on Analytic Hierarchy Process and Fuzzy Weighting[J]. Journal of Jiujiang University (Natural Science Edition), 2022, 37(4): 31-34. (in Chinese)

(编辑: 刘勇)