

MoveVR: Enabling Multiform Force Feedback in Virtual Reality using Household Cleaning Robot 假设检验分析

1120200822 郑子帆

1 文章概要

1.1 基本问题/任务

现今人机交互技术的发展中，VR 设备所提供的触觉反馈相比声音和视觉上的反馈弱很多，无法做到比较逼真、真实的效果。另外，一些能提供良好触感体验的设备很多都尚处于研发阶段，还未能从实验室中走进主流市场。该论文团队研发了 MoveVR 产品，即通过对扫地机器人、和一些家用设备、工具进行改装组合得到一个能够提供四种力反馈（张力、反作用力、对抗阻力、冲击力）的 VR 设备。

1.2 解决思路

该团队通过设计研发 MoveVR 产品，使得用户在家中进行 VR 游戏时能够获得到较为逼真的触感，提高游戏的可玩度和用户游玩的体验感。

硬件方面，MoveVR 由一个家用扫地机器人、机器人控制器 (RMC)、带追踪设备的头戴 VR 显示器 (HMD) 和一个图形工作站组装构成。其中，用于交互的对象 (文中称作代理) 有机器人携带代理、用户持有代理、二者连接共享代理三种类型。绳子可作为共享代理模拟张力反馈，纸盒可作为机器人携带代理模拟对抗阻力，木棒可作为用户持有代理模拟材料的反作用力等等。

软件方面，在 RMC 上运行实现一个控制脚本，可控制图形工作站在不同场景下通过机器人的控制单元发出控制命令，以控制扫地机器人的移动。其移动的方向、速度大小等都是根据游戏场景中所出现的力反馈类型、大小等经过某些特定算法进行转化而得到的。

2 实验分析

2.1 文中已有实验的分析

原论文中一共有 2 个实验，分别为力感知实验和用户体验调查。

2.1.1 力感知实验

该实验的目标为探索 MoveVR 提供多种层次可分辨的力强度、方向的能力强弱。实验中可出现的力反馈等级共 20 个，其中张力强度 4 个等级 (L0-L3)，张力方向 5 个 (L, FL, F, FR, R)，模拟材料刚性强度 5 个不同材料强度 (L0-L4)，阻力强度 3 个等级 (L1-L3)，冲击力强度 3 个等级 (L1-L3)。实验的依赖变量 (Dependent Variable) 共 5 个，分别为张力强度分辨准确率、张

力方向分辨准确率、模拟材料强度分辨准确率、阻力强度分辨准确率、冲击力强度分辨准确率。其计算方法为 $\text{分辨准确率} = \frac{\text{参与者正确分辨的次数}}{\text{该力反馈类型实验次数}} \times 100\%$ 。通过分析，我认为，该实验的**控制变量 (Control Variable)** 有同一 MoveVR 设备（电量等等尽可能一致）、同一测试地点（保证地板材质相同）等；**随机变量 (Random Variable)** 有之前对于 VR 设备的使用经验、参与者的身高体型等。**测试条件 (Test Condition)** 只有 1 个，即使用 MoveVR 设备

实验一共有 12 位参与者（平均年龄 23.4 岁，年龄标准差 3.29，4 女性 +8 男性）。实验过程共分为练习和测试两轮。在练习中，每位参与者会分别亲身体验到 20 种力反馈且会得到对应的力反馈等级。在测试轮中，每个参与者会依次感受到这 20 种力反馈各一次，且这 20 种力反馈出现的先后顺序随机，参与者将根据当前的感受体验给出当前力反馈的形式及等级，由工作人员记录下来。一名参与者将会进行 4 次测试。参与者分辨力反馈的准确性结果如下图（横轴为真实的各力反馈等级、纵轴为参与者分辨的等级）。从实验结果可以看出，参与者对于 5 个力表达形式的分辨准确率分别高达 97.40%，97.08%，96.25%，96.88%，99.31%。

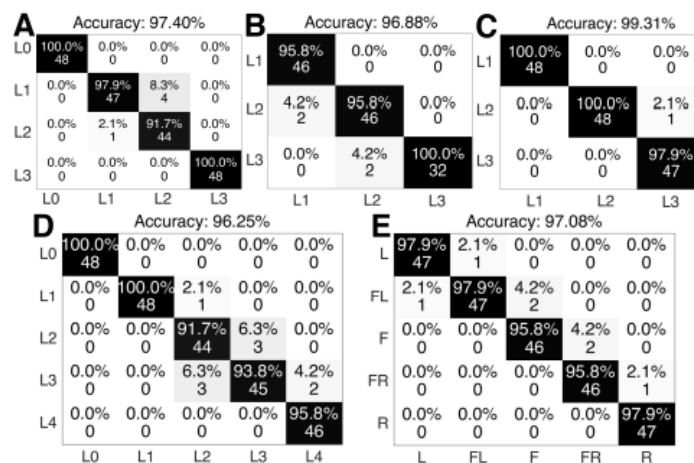


图 1: 力感知实验结果

该实验与课堂上所讲大部分实验并不是同一类型（测试条件只有一个，没有对比），故此实验没有也不需要假设检验分析。此实验有诸多优点和可取之处：

1. 实验过程中各种力反馈出现的顺序随机，避免了参与者的定式学习思维对结果产生重大影响。
2. 实验流程完整，在正式测试前会有详细的讲解和练习，保证实验的正常进行；同时实验保证每个参与者的参与时间在 30min 之内，并有每人 15 美元的补偿，基本可以保证参与者的耐心认真参与以及在体力充沛、思维活跃的状态下参与，减少实验的不确定性偏差。
3. 每位参与者进行 4 轮实验，多次的测试可以减少偶然因素造成的偏差。

当然，我认为该实验也有一些待改进的地方，如：

1. 实验缺乏**独立变量 (Independent Variable)**，该实验看起来更像是按照流程做了一件事情并算出某些数，而缺少对于多个可能潜在影响因素的分析。
2. 实验中只有 12 人参加，与课中所讲 15-30 人为宜有一定差距，可以增加一些参与者；同时年龄方差只有 3.29，说明参加者都在 23-24 岁左右，这样无法体现出其他年龄段对于 MoveVR 力反馈分辨的能力，代表性不够强。所以在参与者抽样选取上可以进行一些调整。
3. 缺少对比与对照。由于实验团队认为 MoveVR 在提供触感方面要比市面上某些流行的 VR

设备好，所以我认为有必要加一组对比实验。

2.1.2 力感知实验的改进

首先，我们可以对于参与者的抽样进行改进。一方面，我们可以增设人的年龄为一个新的独立变量，将 20-30、30-40、40-50、50+ 岁的人群分开，并在各自年龄组中简单随机抽样找出 15 位左右的参与者进行实验。此时我们可以对于不同年龄人群对于力感知分辨的能力进行分析讨论。由于人的年龄属于自然属性 (naturally occurring attributes) 且为定序变量 (Ordinal Variable)，故我们只能进行被试间 (Between-subjects) 实验，根据课上内容，假设检验分析法可用非参数检验 (Non-parametric Test) 中的 Kruskal-Wallis 法，空假设 (null hypothesis) 可为“不同年龄对于力感知分辨的能力无明显差异”。最终可根据实验计算得到的 P 值进行总结分析。

另外，我们可以再在使用 MoveVR 的基础上增加一个新的测试条件：让测试者们用一台市面上成型量产的可提供上述多种力反馈的 VR 设备进行和原实验类似的测试实验。这样实验共有 2 个独立变量：人的年龄和 VR 设备。其中，人的年龄这一独立变量有 4 个测试条件（即 4 个年龄段），VR 设备有 2 个测试条件（MoveVR 和一个其他 VR 设备）。只单独考虑 VR 设备对于力感知分辨的影响（不考虑两个独立变量的相互影响），控制人年龄不变，我们可以利用 Mann-Whitney U 法进行假设检验分析，空假设为“MoveVR 和市面其他设备提供多层次力反馈的能力无明显差别”，若修正的 P 值 $< .05$ 那么可以认为 MoveVR 可以提供更好的多层次和类别的力反馈。

2.1.3 用户体验调查

该实验的目的为评估 MoveVR 能够提供真实和舒适的 VR 体验的能力。实验流程设计包括一个完整的 VR 游戏场景：玩家遛狗回家，在家门口移开挡路的盒子，同时捡起一根木棍击退一个突然袭击的陌生人。实验中共有 24 个参与者（平均年龄 24.2 岁，年龄标准差 7.6，15 男 +9 女），其中共有 8 位曾经无 VR 经验，其余人员曾有过 1-3 次 VR 体验。在实验开始时会有演示实验，在正式测试阶段每位参与者有 2 次实验机会，实验后参与者将填写一个有关于不同测试条件下不同场景的游戏体验问卷 (Likert Scale)，包含游戏真实性 (Q1)，舒适感 (Q2)，易接受程度 (Q3) 和意见建议 (Q4)。

该实验为被试内 (Within-subject) 实验。实验的独立变量为不同 VR 设备，共有 4 个测试条件：(1) 空手条件，无触感体验。(2) 振动触感体验，由 HTC Vive 控制器提供。(3) 被动触觉，在实验房间中的对应位置放与虚拟场景中一样的实物。(4) MoveVR 组。依赖变量为 Q1-Q4 的综合评分，即 VR 游戏体验感。控制变量有实验场景（包括实验室环境布置、游戏中虚拟场景的流程布置）等。随机变量有参与者对于 VR 设备的使用经验等。

实验结果如下两图所示。其中左图分别为 4 个场景（遛狗、推箱子、打人、被打）中四个测试条件下的 Q1、Q2 的平均得分。右图为四个测试条件下总的 Q1、Q2 的平均得分。

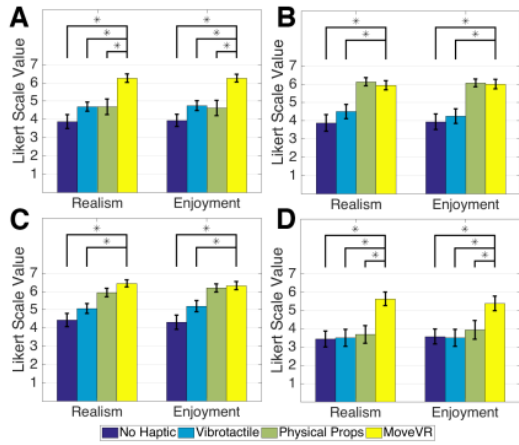


图 2: Q1、Q2 分场景得分

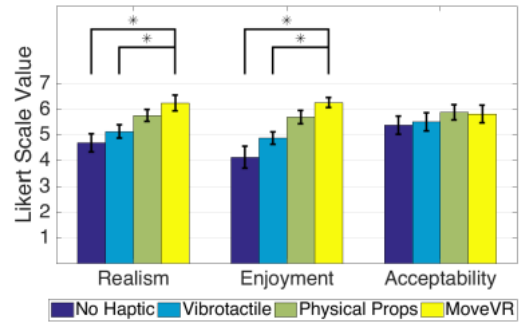


图 3: Q1、Q2 总得分

该实验团队采用了单因素、被试内的 ANOVA 和事后检验 (Post-hoc test) 进行了假设检验分析。测试条件 $n = 4$, 实验人数 $m = 24$, 故 effect 为 3, residual 为 $m - n = 20$, 故自由度分别为 3, 60。MoveVR 的 Q1 均分 $6.24/7$ ($F_{(3,60)} = 5.37, p < .05$); Q2 均分 $6.25/7$ ($F_{(3,60)} = 10.14, p < .005$), 故可得结论 MoveVR 组的游戏真实性和舒适度在统计学意义上和空白组、震动触感组有显著的差异。但是相比实物组, $F_{(3,60)} = 0.51, p = .68$, 说明二者无显著性差异。若分别针对 4 中游戏场景进行分析, 也可得到 MoveVR 组相比前两组真实性和舒适度有显著差异。综上所述, 相比其他的 VR 环境或设备, MoveVR 设备可以给用户真实的、良好的 VR 体验。

该实验的各组成部分、流程设计以及假设检验分析都很完整, 同时结果也达到了预期值。当然, 我认为该实验还可以有一些补充, 另外也有一些疑问。

1. 该实验和上一个实验存在一个共同的问题, 没有考虑到不同年龄人群的影响, 尤其 Q3 (易接受程度) 可能会有很大的影响。故可以将年龄增加成为独立变量。后续实验分析中可以增加人的年龄和不同 VR 设备的交互影响 (Two-way ANOVA)。
2. 实验中使用了课堂中所讲的针对单因素、被试间实验的方法进行了自由度的求解, 但该实验确实被试内实验, 所以在求自由度的方法选择上是否存在错误和偏差呢?

2.2 可增添的实验分析

2.2.1 MoveVR 组装便捷性调查

由于该实验团队设计的 MoveVR 的一大特色优点为其构成材料均可以来自于家中的常见物品 (除了控制器以及脚本程序)。显然, 如果 MoveVR 能够成功推广普及, 那么用户在家中可以轻松方便地组装和拆卸一定是一个必备条件。故可补充一个实验, 有关于 MoveVR 的组装便捷性调查。

该实验可以设计为被试内实验, 邀请 20 位左右的参与者。独立变量有 VR 设备 (MoveVR 和一个市面量产的可提供一定触感的 VR 设备), 依赖变量为 VR 设备的易接受度 (包括组装使用便捷性、游戏体验等综合考量), 利用李克特式量表 (1-7, 4 为 Neutral, 1 为 Strongly disrecommend, 7 为 strongly recommend) 对参与者进行调查。

实验分析可以利用单因素、被试内的 ANOVA 方法计算 P-Value, 再据此结果结合评分均值判断两种 VR 设备在易接受使用评分方面是否有显著的差异。若有, 则说明 MoveVR 的组装流

程与方式受到了大众的认可，为实现普及化提供了基础。