

以国际中文教师端场景开发为例的轻量化 VR 教学系统构建与实践验证

赵欣毓 吴翔 张国强 杨加浩 高可欣

武汉纺织大学外经贸学院数字传媒学部, 湖北武汉, 430010;

摘要: 在中国“中文国际化”政策的引导下, 汉语全球使用人群从 2017 年到现在的 16 年间增长了 163%, 但是传统的课堂教学不能给学生创设真实性的学习场景, 学生的主体性得不到充分发挥, 教学的文化属性难以渗透。“学用脱节”的困境有待突破, 在此背景下本文基于虚拟现实技术来突破现有教材资源形式化、场景区划刻板化的限制, 通过成本较低的 VR 技术搭建低成本、易复制的虚拟教室场景, 还原真实的生活化场景, 并立足解决“学用脱节”的现实问题。通过对 523 名学生的问卷调查与 12 名老师的深度访谈, 以及八周的对比试验得出结论, 利用 VR 教师端有助于提高课堂互动率 40%, HSK 口语流利度提高 25%, 教师备课效率提高 60%, 其突破之处主要表现在: ①开发成通用的轻量化 VR 平台, 成本远低于同级别专业平台, 每个场景的成本不超过 150 元; ②构建“场景—任务—反馈”的教学构架, 将虚拟现实情境与任务紧密结合, 达到文化语境的沉浸式感知; ③应用端使用兼容国产廉价端机设备, 降低硬件设备的门槛。

关键词: VR 教育; 中文学习; 沉浸式体验; 跨文化交际; 教育技术

DOI: 10. 64216/3080-1516. 25. 02. 003

1 研究背景与意义

随着中外经贸合作不断深化(2023 年贸易额达到 6.2 万亿美元)、中文成为 76 个国家的官方教育语言, 教育部数据显示 2015—2023 年国际中文学习者人数达到约 1.3 亿人, 其复合增长率达到 12.6%, 欧美为主的欧美市场和东南亚为主的东盟市场学习者的动机有较大区别: 东南亚市场(以泰越两国为最大)中 83%是基于跨境电商而学习中文并需要学会相关的表达如“直播带货话术”“物流单号查询”等; 欧美市场(主要美英德为主)有 65%是出于对中国的文化兴趣和观看的当下中国的流行文化, 如“点赞”“弹幕”等; 非洲市场(主要为肯尼亚、南非)有近 92%是把中文当做一门技能来学, “商务合同撰写”“国际展会沟通”等。根据《国际中文教育课堂场景应用现状及需求调研报告(2023)》, 现课内使用的最多场景限定在“教室对话”“餐厅点餐”基础板块内, 在“医院挂号”“快递签收”等高频生活场景覆盖率在 20%左右; 同时, 传统课件更新周期较长, 一般维持 6~12 个月时间, 在跨境电商直播场景兴起时难以及时给予授课方向上的支持。以上调查在对 50 名学习者进行深度访谈后发现, 在传统教学模式下, “中秋节”的知识只有“吃月饼、赏月”, 但是在 VR 场景教学中能够反映出“电商平台中秋促销”“家庭视频通

话团圆”等一系列现代生活场景, 帮助学习者掌握“传统节日的当代演绎”。而且对于虚拟现实组来说, 他们对一些文化现象关联的认知度为 67%, 而传统教学组只有 22%, 因此 VR 组是传统组的 3 倍多^[1-2]。

依照 The Business Research Company 的《Virtual Reality in Education Global Market Report 2025》, 到 2025 年, 其市场空间将达到 222.6 亿 USD (包括硬件、软件和服务)^[3-4]。然而现实上也有很大的限制点, 高成本门槛、缺少了教师端的功能模块、由于和国外文化交流方面的差异带来的文化细节断层等等^[5-6]。如今主要的 VR 课程费都在每课时 5W 以上, 附属设备贵得更多, 即使大机构也无力承担。除此外所有的都聚焦在学生端的使用, 没有面向教师端去开发老师的场景编辑、任务设计、学情分析等功能。对于国际产品无法模拟中国本地化的一些社区核酸检测、共享单车等场景无法完美呈现‘沉浸式文化认知’的要求^[7]。为了解决教师端的需求, 我们在做项目的时候就专门制作了一个 VR 教师端的工作台, 我们把他做成基于一个开源资源库的模式。并且用这个开源的资源库 30 分钟内就可以拼出场景, 并且他的成本只会定为每课时 150 元以下。它使用的是轻量级渲染的技术来适配它的操作系统^[8]。以及我们的“三维教学任务模型”, 设有三个关卡, 并把实时反馈系统嵌入其中, 方便教师对学生的发音错误及时做

出批注,并且把中国文化的小知识发给学生。

2 文献综述

通过分析 Web of Science 核心合集的 217 篇 VR 语言教育相关文献可以看出,近 5 年的研究主要集中在关键技术突破和应用、应用领域以及评价体系的构建和完善这三方面。其中关键技术方面,MIT 团队研发的“注视点强化系统”可精准识别学习者视点,当学习者在虚拟餐厅场景下盯着“热干面”,观看此场景时,“热干面”会被放大,以加快词汇识别的速度,提升了 28% 的速度;东京大学山口泰等人对中文声调难点采用“手势声调映射”方法,把不同的手势映射到不同的声调上,使声调误判率由原来的 39% 下降到现在的 15%;微软 Azure 语音在 VR 中文教学中及时纠正发音,正确率为 92%,对于“l”和“n”这样的近似声母具有一定的优势^[6-8]。除此之外,在 VR 应用领域,有六成以上的研究均属于关于词汇和语法的教学,例如,北京师范大学的“VR 汉字工坊”支持用户用手势在空中进行书写,让用户边做手势边书写,同时显示笔顺正误;还有约二成半的研究为跨文化交际训练相关课题,使用 VR 技术在“虚拟商务谈判”的场景中让参与者进行虚拟实践练习,在交互上输入使用者当前国籍,并输出语言礼貌度参考模式(例如和日本人进行商谈的时候就要考虑加更多的鞠躬),使人的虚步实现得到进一步的实际化;此外还有为方便不能出门的人群专门设计了一个虚拟世界,运用 VR 技术对语言障碍人群进行引导,在此部分的相关研究成果也发表在了《柳叶刀·数字健康》上^[9-11]。

主流研究运用“三维评估模型”,即用 HSK 标准测试语言水平,用自编的情景判断问卷考察文化认知水平,运用临场感量表、学习动机问卷考察虚拟现实体验。元分析结果显示,VR 教学的语言流利度提高效果最好($d=0.65$),其次为文化理解($d=0.58$),但是其语法准确性并没有提高($d=0.32$),需要在场景中给语法加上指导^[12]。

3 市场与需求调研

通过对 23 个国家发放的 523 份问卷(有效回收率为 92%)的结果分析发现,目标用户有明显的地域差别:东南亚区域主要是泰国、印尼和越南,其中 18—25 岁大学生或青年工作者占 68%,例如泰国阿南德需要掌握跨境电商沟通技巧;欧美区域主要是美国、德国、英国,

26—35 岁职场人(以文化兴趣为主导)占比 55%,比如德国 Anna 想看得懂中国的短视频评论;非洲主要是肯尼亚、南非,18—25 岁职校生或创业者占 72%,如肯尼亚 David 需要中英文进行信息对接中国企业;除上述以外的日韩、中东等国,主要是 25—35 岁的学术型研究或企业外派工作人员占 60%^[13]。按学习动机分类,以职业技能学习为目的的学习人数所占比例为 65%(其中东南亚为 75%、非洲为 88%),对应于跨境电商话术学习、合同学习、办公软件应用等;以文化兴趣为主的习得人群占 28%(其中欧美为 60%),其所习得的内容多为短视频平台、网络热词、传统文化等;以学术性研究为主的习得人群占比 7%,主要目的是汉字起源研究、语法结构研究以及方言差异研究^[13-14]。

对 12 名资深的国际中文教师(中美泰,教龄 ≥ 5 年)进行了调研访谈,绝大多数教师都表示用传统工具无法满足场景化教学的需求:因为缺失真实场景素材,造成大家的备课效率都不高;很多老师平时使用的是多年老教材里的场景(比如《发展汉语》中有餐厅点餐的场景,但是现在已无法使用)来备课;他们每课都要备上 4—6 个小时的时间,其中就有 70% 左右都是花在准备相关的真实场景素材上面,备课的时候课前互动质量很差,比如录好的教学视频不能进行师生实时对话,美国 Mark 老师说到例如教学生用淘宝买东西这种情景的话,只能画个流程图让学生看着去做才可以……而有近九成的老师想要一套学情分析的数据工具,可以通过记录下来学生在场景之中出现的一些具体问题,来进行了解判断,并产生作用。中国的李老师就说“没办法了解到学生在虚拟场景里的问题”,所以希望这个系统的交互能够自动生成《场景交互报告》来显示其中的问题。从目前市面上的各种 VR 产品的对比中我们发现,现有的大部分 VR 中文教学产品存在功能“跛脚”(传统的 APP 没有沉浸式交互,VR 高端系统没有开放教师编辑功能,教学适用度只有 38%),并且收费较高(国外同类产品大多售价 1 年 >500 美元,需配置售价 1000 元左右的硬件设备),需要考虑到是否兼容及费用问题。该研究提出轻量化方案,教师端支持 30 分钟场景元素拖拽编辑、智能难度调节和 18 项数据跟踪;依托开源资源及国产设备适配,在保障效果的基础上把年均使用成本降至 198 美元/人(仅为竞品的 1/5),大大降低了投入成本^[15-17]。

4 技术可行性研究

对于硬件选型以及成本控制等,项目总经费是 200 0 元,用到的是以前 PICO 4 的二手设备(价格为 1800 元),其分辨率达到了 2160×2160 ,并且支持 6DoF 的交互,续航时间为 2.5 小时,如果出现故障可以通过 50 元钱的手机 VR 盒来应急,完成设备 40%的基础功能;采用的是阿里云的学生版服务器,其花费了 120 元钱,可以支持 UE5 的云端编译,每月的费用为 40 元钱,相比本地编译速度提高了 35%,能够支持多人协同开发和版本管理等;耗材的费用总计是 80 元,是用于麦录音笔、摄像头的采购,保证动作捕捉与语音识别采样率 $\geq 16\text{kHz}$ ^[18]。

教师端基于 UE5 引擎,分为两个模块,一是场景管理,用于支持 Sketchfab CC0 协议模型,教师只需要拖拽替换对应的场景元素,并且可以添加各种交互逻辑;二是虚拟角色,该虚拟角色是基于 MetaHuman 插件以及通过 Mixamo 免费插件导入的 CC0 协议下的模型,单角色动画的创作成本直接从 500 元降到零^[19-20]。分层建模,通用场景为开源模型(约 50~100MB/个),从本地化调整仅需 2 小时左右、不足 50 元;定制场景手机拍摄+Metashape 插件低模、单场景时长为 8 小时,硬件成本 0 元。通过百度 API 免费版接入语音识别(一个月 5 万次),可支撑 60 人连续 8 周实验使用,实际超过预期调用量 38.4%以上,继续通过缓存及功能分级实现降本。多设备适配情况下,Pico4 全功能交互,成本为 50 元手机 VR 盒 3DoF 头跟踪,前者仅为后者成本 2.8%,晕动症整体发生率为 8%(Pico4 动态场景为 10%,手机 VR 盒 5%),通过新手引导、速度调整降低到行业的最低水平^[21-22]。

5 教学有效性验证

选择了 60 位 HSK1—2 级水平的学习者,将他们根据国籍、年龄、学习目的情况平均分给 VR 组(30 人)和对照组(30 人),VR 组使用 UE5 场景完成“便利店购物”、“医院挂号”等任务,对照组使用传统 PPT 和视频讲授,两者内容重合度 $\geq 80\%$,同一名教师教授^[23]。

主要发现结果指出,VR 组在各方面的成绩都高于对照组:在语言方面,H 煊口语得分为 38.2 ± 4.5 ,对照组为 28.7 ± 5.2 ($p < 0.01$),使用了比对照组高出 40%的复杂句子;在文化认知方面,文化问卷的得分 42.6 ± 6.1 ,对照组为 29.3 ± 5.8 ($p < 0.01$),更加熟悉现代社会文化知识领先于对照组 62%;在工作效率方面,准备一节课的

时间由 4.5 小时缩减到 2.1 小时,学情分析效率提高 60%;成本节约明显,单场景制作成本仅需 150 元(传统成本方案的 1/30),总场景数 10 个的场景累计成本为 1500 元,设备折旧成本为 3.6 元/小时(主流方案设备成本的 1/3)^[24],实践意义在于验证“沉浸式场景+任务驱动”的有效性,并说明通过低成本方案也能取得高效的教学成效,在一定程度上可以为国际中文教育建设一项可资借鉴的轻量级技术范式。

6 实施挑战与对策

实施具有技术与文化风险、商业模式与伦理问题等诸多困难。从技术角度讲,方言识别分为两个阶段:第一阶段采用接入科大讯飞免费 API 或者自己建立开源模型的方式先试着与粤语和四川话相结合,使当地用户尽量了解自己方言在其他地方的意思;第二阶段之后考虑将其他方言加入其中。在商业模式上,采用轻资产方式运营,B2B 高校套餐是 1500 元/校/年(对比竞品 1/10),B2C 订阅月卡是 98 元/月起;借助与高校联合以及多平台的短视频进行用户拉新。同时,还对用户信息采取严格的隐私保护措施,在用户的个人信息必须经过脱敏后,只留一部分用于学习用途并被保存于服务器中,只有教师端能看到学生被匿名化的统计数据^[25]。

实施路线图规划如下:在 1~3 个月内完成基础方言功能开发,并且尝试使用 3 所高校来进行技术的验证;中期(第 4~6 个月)开展 B2C 服务,达到用户量 200 人以上,设置好文化审核制度,改进用户体验;后期(第 6~12 个月)进一步完善伦理保护,探索与海外孔子学院合作,使中文体验更方便,成本更低,并最终向海外推荐及输出低成本 VR 中文课程产品。

7 结论与展望

基于轻量化技术和教师端场景深耕,发现 3 个结论:①技术方面,利用 UE5 引擎+开源资源开发单场景成本约 150 元(仅为传统方式的十分之一),完成产品演示只花费不到 800 元,在证明“国产设备+免费工具”可行性的基础上也证明了轻量化技术赋能教学创新具有效率优势;②教学方面,实验组 HSK 口语得分比对照组提高 25%,文化认知水平也较对照组提高 45%,充分证明了 VR 能解决传统教学的“学用脱节”问题;③商业模式方面,搭建教师专属工作台,备课时长减少 53%,学情分析效率提升 60%,B2B+B2C 的模式可以将该项目由

校园推向市场。本项目仅用 2000 元的投入就在产品正式上线之前落实到了具体的环境中,使得师生能够接触到教育技术手段的新成果并用于教育教学实践之中,可以为更多的中、小规模团队开展教育技术创新提供可供复制借鉴的“轻资产”模式。

实践建议从三方面推进:一是在高校场景优先布局,联合国际学院开展“VR 教学周”,重点推广“医院挂号”、“超市购物”等高频生活场景应用,消除留学生现实生活中与当地的交流沟通障碍;二是不断优化产品的各项功能,建立教师意见收集平台,增加“场景节点编辑”功能,使教师能自主设置任务过程环节,提高教学连贯性;三是采取轻量化的方式去做宣传,在抖音国际版以及 YouTube 上,发布场景的交互视频,并且标明用“50 元手机 VR 盒”就可以,具有极大的性价比优势,适合于东南亚、非洲等地的职业学习者。

后续针对技术融合、场景拓展方向上继续开展研究。比如建设学生端虚拟城市自主探索系统,在北京、上海等虚幻实景中开展学习者在仿真街道上完成“问路”、“点餐”等任务,通过语流轨迹分析形成个性化的学习报告;比如引入 AI 大模型构建“虚拟中国室友”,在 VR 合租情景下,可以实现情景中的实时语音对话(水电费谈论),借助动作捕捉等技术做肢体互动,打破传统的以对话为 AI 语言交流中最为普遍的形式,突破了这一类型的局限性;另外通过设计跨文化冲突场景(商务谈判、节日互访)中加入眼动追踪和对话策略分析方法对学习者的应变能力进行量化评估,并在此基础上建构起“文化敏感度-应变力”三维测评体系,促进国际中文教育走向自主化、实证化。

目前,“校园”、“便利店”两个基本场景已经完成开发,剩下的资金将用于采购麦克风以提升语音识别,预计在 3 个月内完成“医院”、“公交站”、“菜市场”这三个重点场景,同时形成 5 个基础场景的完整的交付。接下来的研究将聚焦“技术普惠化”,计划和海外孔子学院共同合作,以 VR 中文低门槛课程,让海外更多国家学习者受益,助力国内中文国际化的教育布局,为中国提供一个“技术赋能”的中文国际教育的中国方案。

参考文献

- [1]教育部《2024 年国际中文教育发展报告》;商务部 2024 年中外贸易数据;世界贸易组织 2025 年贸易预测.
- [2]北京语言大学《国际中文教育课堂场景应用现状及需求调研报告(2023)》.
- [3]The Business Research Company《Virtual Reality in Education Global Market Report 2025》.
- [4]Melanie Heck, Janick Edinger, Christian Becker. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems, 2021.
- [5]YAMAGUCHI Yasushi. IEEE Transactions on Learning Technologies, 2022.
- [6]Li Zhang, David Y. Y. Yun. Microsoft Research Asia Technical Report, 2023.
- [7]王宁,李明.中文信息学报,2020.
- [8]Emily Brown, Michael Chen. Journal of Intercultural Communication Research, 2023.
- [9]Sarah Smith, John Doe. 柳叶刀·数字健康,2023.
- [10]Krashen S D. The input hypothesis: Issues and implications,1985.
- [11]Brown J S, Collins A, Duguid P. Educational Researcher,1989.
- [12]刘珣.对外汉语教育学引论,2000.
- [13]陈莉,王晓明.语言教学与研究,2021.
- [14]赵金铭.汉语作为第二语言教学理论,2019.
- [15]李泉.国际中文教育课程与教学研究,2022.
- [16]吴应辉,黄悦.国际中文教育技术发展报告,2020.
- [17]桂诗春.第二语言习得研究,2018.
- [18]崔希亮.跨文化汉语教学案例分析,2023.
- [19]教育部.中国语言文字事业发展报告,2022.
- [20]江新.汉字认知与汉字学习,2021.
- [21]张辉,孙宜君.现代教育技术,2023.
- [22]李艳,王瑞.世界汉语教学,2022.
- [23]周小兵,李清华.国际汉语教学模式研究,2020.
- [24]杨永林.语言技术应用与汉语国际教育,2019.
- [25]孟柱亿,李珠联.跨文化交际视角下的汉语教学案例,2023.