中国科学: 信息科学 2016 年 第 46 卷 第 12 期: 1779–1784
SCIENTIA SINICA Informationis
《中国科学》杂志社
SCIENCE CHINA PRESS
虚拟现实纵览专刊 · 观点与争鸣

虚拟现实技术的现实挑战

张文波*, 曹雷, 熊君君

北京三星通信技术研究有限公司, 北京 100095
* 通信作者. E-mail: wenbo.zhang@samsung.com

收稿日期: 2016–08–02; 接受日期: 2016–10–27; 网络出版日期: 2016–12–09

摘要 随着计算机技术的发展, 虚拟现实 (virtual reality, 简称 VR) 技术自 20 世纪 50 年代以来, 尤其近年来得到了长足的进步和较为广泛的应用, 正在从科学与专业领域向日常生活扩散, 2016 年更被广泛认为是 VR 元年. 本文从基本概念出发, 以工业界及消费者视角, 分析和总结了当前消费级虚拟现实技术的现状和挑战, 并对虚拟现实技术的未来给出有益的思考.

关键词 虚拟现实 人机交互 挑战

1 引言

虚拟现实 (virtual reality, 简称 VR) 是一种通过计算机模拟产生的, 具有身临其境感觉并可交互的虚拟场景的技术, 该场景通过提供视觉、听觉、嗅觉、触觉等感官的模拟, 提供给使用者沉浸式的体验和自然的交互方法. 虚拟现实技术融合了计算机、多媒体、传感器等新兴技术, 并随着这些技术的发展而具有越来越广泛的应用.

2 虚拟现实技术简短回顾

虚拟现实概念的提出, 可以追溯到 1935 年, Stanley G. Weinbaum 的小说中所设想的一种通过视觉、触觉、嗅觉的模拟来提供沉浸式体验的虚拟现实设备 [1]. 然而, 公认最早的虚拟现实设备 Senorama 直到 1957 年才被 Heilig 发明出来 [2,3]. 该机器通过模拟立体视觉、立体声、振动和风来实现身临其境的感觉. Ivan Sutherland [4] 于 1965 年对虚拟现实技术做了系统性描述, 并发明了包含 HMD、位置跟踪和实时虚拟空间绘制的 SKetchPad, 正式开启了虚拟现实的现代研究, 包括 20 世纪 60 年代 Thomas Furness 为美国空军开发的虚拟飞行系统; 1977 年 Dan Sandin 等开发的第一个数据手套 Sayre Glove; 1985 年, 由于太空探索的需求, NASA 开发出 VIEW 系统 [5]. 这些设备或系统表明虚拟现实技术在专业领域和研究机构获得了长足的进步. 在商业领域, 1984 年诞生的 VPL Research 公司开始发售虚拟现实软硬件, 包括数据手套、HMD 和绘制引擎 ISAAC; 1994 年, 任天堂推出的 VrBoy 游戏机掀起了虚拟现实的大众化浪潮, 然而该产品最终以失败告终.

引用格式: 张文波, 曹雷, 熊君君. 虚拟现实技术的现实挑战. 中国科学: 信息科学, 2016, 46: 1779–1784, doi: 10.1360/N112016-00162

© 2016《中国科学》杂志社

虚拟现实技术的再次兴起, 得益于计算机硬件和相关软件的快速发展. 近年来, 基于摩尔定律的计算能力提高以及现实技术的发展, 相继出现了 Google Cardboard[1)], GearVR[2)], Oculus Rift[3)], HTC Vive[4)]等基于不同应用场景的虚拟现实设备, 在产品体积、价格和用户体验上都有了巨大的飞跃, 2016 年也被广泛认为是虚拟现实产业 “元年”.

回顾虚拟现实技术的历史, 我们会发现, 尽管虚拟现实是新兴技术, 但并不是近年才出现的技术. 在半个多世纪的发展中, 虚拟现实技术既有长足的进步, 也有大众化产品的失败历史. 随着软硬件技术的爆发式发展, 有越来越多的消费级产品开始涌现在市场上. 然而, 虚拟现实技术是否真的能够应用于大众化产品, 是否真的能够满足人们对想象和身临其境体验的追求, 本文从工业界角度给出了自己的分析.

3 虚拟现实技术的基本特征

虚拟现实技术通常包含 3 个主要特征 [6].

沉浸感 (Immersion). 虚拟现实技术的目的, 是在计算机的帮助下, 制造出具有真实感觉的虚拟环境, 从而让使用者具有沉浸在现实环境的感觉, 甚至产生认知体验.

交互 (Interaction). 在计算机生成的虚拟环境中, 使用者可以借助各种传感器来与虚拟环境进行交互, 从而产生与现实环境交互相同的互动体验.

想象力 (Imagination). 虚拟场景, 除了实现对现实的高度还原, 人的大脑所产生的任何想象, 都可以借助虚拟现实技术呈现到虚拟场景中.

以上的 3 个主要特征互相依赖, 共同构建起虚拟现实技术的基本概念, 即构建根据想象所产生的, 经由计算机和传感器技术实现的, 可以自然交互的虚拟环境, 给予使用者身临其境的沉浸式体验.

基于如上理解, 可以分析人类对虚拟现实技术的终极需求和当前虚拟现实技术现实情况之间的距离, 并由此展望未来.

4 虚拟现实的现实挑战

虚拟现实技术在计算机技术、多媒体技术和多种传感器技术的帮助下, 提供虚拟环境给使用者以沉浸式的现实体验, 主要通过人的感觉, 如视觉、听觉、触觉等将虚拟环境的信息送达使用者的大脑, 从而产生感官体验. 因此, 虚拟现实技术在多大程度上能够模拟真实的感觉, 就成为决定该技术成败的关键. 本文从如下几个较有代表性的方面出发, 从工业界和使用者角度逐项进行分析, 并在最后给予总结.

4.1 视觉虚拟

随着显示技术的提升, 显示设备可以做到轻、薄的同时, 显示效果有极大提高. 然而, 要模拟真实的视觉体验, 虚拟显示技术尚有许多不足.

1) https://www.google.com/get/cardboard.
2) http://www.samsung.com/global/galaxy/gear-vr.
3) https://www.oculus.com.
4) https://www.vive.com.

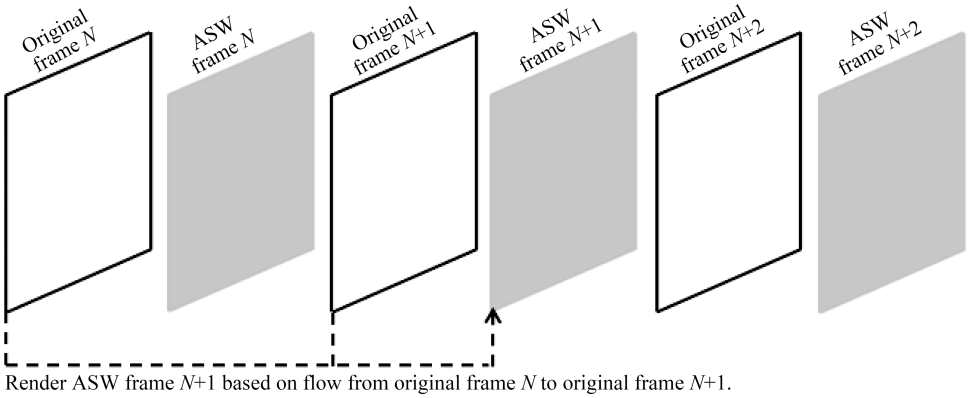

图 1 异步空间扭曲

Figure 1 Asynchronous space warp

解析度. 当前主流头盔式虚拟现实显示设备正在从 4 K 向 8 K 迈进. 那么人眼需要多大的解析度呢? 研究发现 [7, 8], 人眼的角度分辨率约为 0.35~0.39 分, 假设为 0.4 分, 则对 360° 全景所需分辨率为 360×60/0.4 = 54000, 即理论上人眼分辨率为 54 K×54 K 像素. 尽管市场上已经出现昂贵的 20 K 设备, 并且可以对解析度作出妥协, 如 AMD 公司认为 16 K×16 K 已经足够, 但要想达到人眼对真实世界扫描的解析度, 并出现在大众化产品上, 硬件显示设备还有漫长的道路.

刷新率/帧率. 是对显示设备更新内容速度的综合评价. 当前的主流产品刷新率约为 90 Hz, 移动 VR 设备如 Samsung Gear VR 为 60 Hz. 虽然人眼由于视觉残留特性, 30 fps 的帧率即可满足基本需求, 如电影, 但由于 VR 设备的交互性, 使用者头部的动作需要以极高的频率进行采样, 并及时对新的位置进行图像更新. 目前, 业界广泛认为帧率应达到 144 Hz 才能保证真实感.

视场角. 人眼的双目总视场在垂直方向上大约为 130°~135°, 在水平方向上大约为 200°~220°, 其中重叠部分约 120° [9], 然而, 目前主流 VR 头显的水平视场角在 90°~110° 之间. 如果能够在虚拟环境中模拟现实场景的 200°~220° 水平总视场角, 则需要在头显中缩短人眼到透镜的距离, 或使用较大的透镜, 其带来的问题或是因成像无法聚焦于视网膜导致看不清, 或是因放大倍数太大造成局部颗粒感严重, 甚至由于镜头尺寸太大而损伤眼睛. 由于人类视觉的真实感, 不仅仅来源于视场的双目重叠部分, 如何能在有限的重量体积限制下, 低成本地生成 200°~220° 的水平视场来模拟真实环境就成为较大挑战.

抖动. 现实场景中抖动仅出现在特殊情况下, 但 VR 中包含不同原因的抖动. 在硬件能力限制下, 如何使 VR 视频平滑呈现, 催生了异步时间扭曲 (ATW) 和异步空间扭曲 (ASW) 技术, 异步时间扭曲技术通过惯性传感器 (IMU) 的高频采样来预测用户视角的变化, 在需求的绘制任务无法及时完成时, 将当前已绘制好的最后一帧经过补偿呈现给用户, 从而保持画面的连续性, 该技术目前已经较为常见. 然而, 该技术仅补偿视角的旋转而不包括平移, 当绘制场景有近距离物体, 并且用户视角存在平移时, 近处将出现位置抖动 (positional judder). 异步空间扭曲通过传感器采样用户位置和视角变化, 并根据最近两帧图像来预测图像中物体的移动, 并根据这种移动来生成 ASW 帧, 从而解决了位置抖动问题. 如图 1 中, 通过计算原始帧中相邻两帧的光流, 预测出下一 ASW 帧中物体的绘制位置, 从而解决了由于视角移动所造成的视差带来的重影问题, 同时帧率提高了一倍.

其他图像质量相关. 现有主流显示设备的颜色重现与自然场景依然有差距, 如在色域范围上, 仍然是人眼可见色域的子集; 而颜色的精细程度, 与自然场景下的色彩精细程度有较大距离. 另一方面,

高端显示设备的动态范围虽然已经可以媲美人眼 10000 : 1 的动态范围, 但如何使大众化的成本支出与高动态范围、高色彩还原能力和低噪声水平等技术要求趋于平衡, 并使这种平衡与人眼的特性保持一致, 从而给人以 “现实” 的感觉, 既是一个技术问题, 也是一个感知心理问题.

4.2 眼动

头戴式显示设备虽然可以单独传送图像到眼睛, 但是有如下几个问题.

深度感知. 真实场景中, 眼睛的聚焦点和双眼视线汇聚点是重合的, 因此, 在深度的感知上是统一的. 对于当前的头显来说, 由于采用 LCD 眼镜方式, 眼镜的聚焦点和双眼视线汇聚点不在同一个深度上, 聚焦点仅仅是屏幕的表面, 而双目视线汇聚点代表虚拟物体的距离. 其结果是导致人眼的 “肌肉深度”, 即控制眼睛聚焦的肌肉群的动作所对应的聚焦点距离, 和 “视觉深度”, 即双目视线汇聚点距离之间的冲突, 也成为眩晕感或者不适的来源. 光场技术可能是解决这一问题最根本的方法, 但距离实用还比较遥远.

视线跟踪. 将人眼作为一个光学系统来看, 当眼睛转动时, 其光心和瞳距都在变化, 对于自然场景, 此类变化已经被人类所适应. 而对于头显来说, 此类变化可能造成图像呈现质量的下降或三维感觉的减轻. 由于人眼移动的角速度最高可达每秒 900°, 视线跟踪在虚拟显示中应用的难点, 在于准确地、快速地进行视线判断和响应.

注视点跟踪. 在自然场景中, 人类的视觉会仅关注注视点附近的细节, 而忽略/模糊注视点之外的部分. 因此, 在虚拟显示中, 当注视点和其周围都很清楚的时候, 使用者潜意识里会觉察到与真实场景的不同, 从而在长期使用后产生不适感. 此外, 在深度方向上, 人类在观察某一个深度的时候, 同一方向的其他深度的物体也会由于景深的有限而变模糊. 在上述两种情况中, 如何快速有效地判断出注视点并作出相应的绘制改变, 还需要虚拟显示技术在硬件和软件上有长足的发展.

4.3 移动反馈

人类在与环境的互动中真实感的重要来源, 是在与环境的交互中可以得到基于历史经验学习而来的反馈. 从基本的身体动作出发, 可以简单归纳为如下几个方面.

头部姿态变化. 包括旋转和平移. 当前对头部姿态的跟踪较为成熟, 如三星 Gear VR 基于陀螺仪获得头部旋转信息, 其信息更新间隔可以达到 2 ms; HTV Vive 使用外置摄像头对头部的旋转和平移进行跟踪. 但在进一步基于此类位置信息进行图像重绘的过程中, 其反应速度还依赖于帧率和刷新率.

身体部分的移动. 如手脚肘等部位位置在虚拟环境中的体现, 亦即真实的身体部分, 如何反应在虚拟的环境中, 并且与虚拟环境的交互能够达到与真实世界交互无异. 以手为例, 市场上已经有多种设备, 如 Oculus Touch, HTC Vive 手柄, 微软 handpose, 三星 Rink 等等. 借助多种传感器的信息融合, 基于手的虚拟现实交互方法已经逐渐成熟, 但对于身体的其他方面, 目前还较少研究, 比如脚、膝盖、手肘等身体部位与虚拟场景的交互, 目前还不成熟.

重力感/失重感. 人类在真实环境中移动、跳跃时, 会产生重力感或失重感. 然而, 当我们在虚拟环境中移动、跳跃、下坠时, 大脑会根据物理状态的不同而 “预期” 不同的重力感觉和身体状态. 而现实中, 我们的身体基本是停留在原地不动的. 这种 “预期” 状态与现实状态的不同, 也是虚拟现实系统设计者所必须考虑的问题.

综上所述, 虚拟现实技术尽管已经获得长足进步, 亦能在很多方面满足消费级产品的需求, 但依然距离给人以真实感有相当的距离. 当前产品容易造成使用者的眩晕或不适, 其最主要的来源, 是由

于虚拟世界中的 “感觉” 和现实身体的感受的冲突. 这种冲突, 既体现在视觉的质量、人眼的属性上, 也体现在身体的其他感官在虚拟与现实环境中感受相背离上. 从技术上讲, 其中的一些问题, 如滞后、图像质量等可以通过技术的发展来解决, 但另一些问题, 如身体感觉在虚拟和现实中的冲突等, 还没有很好的解决方法, 产品级应用就更加遥远了.

5 虚拟现实的未来和挑战

随着硬件和软件技术的快速发展, 虚拟现实技术开始以消费级产品的形态服务于普通消费者. 然而, 当前的虚拟现实设备较容易导致实用者的不适, 究其根本原因, 在于人与虚拟环境的交互体验中, 出现了与现实状态相违背的现象, 因此, 人的大脑开始潜意识地认为某些信息难以接受, 从而出现不适. 因此, 未来的虚拟现实如何发展以及可能出现的问题, 我们思考如下 3 个方面.

感官体验. 随着技术的进步, 满足视觉、听觉甚至触觉的感官真实感是可以解决的, 比如显示分辨率、视场、三维声场、力反馈等等, 随着软硬件技术的发展, 更加轻薄、更加精细的元器件可能超越人类感觉极限值, 从而满足人类需求.

认知体验. 虚拟现实技术是人类与计算机系统的一种交互方式, 是对人与现实场景交互的一种模拟. 人类对现实场景具有认知能力, 即对客观世界得来的感官刺激所产生的信息进行分析、归纳等加工活动, 从而形成特定的知识或心理活动, 并指导个人行为的能力. 相似地, 人类对虚拟环境也具有认知体验, 如果虚拟环境不能产生与现实场景相似的认知体验, 要么对使用者原有认知产生冲突, 要么使用者形成与现实不同的认知, 其后果可能是灾难性的. 如何处理好人对虚拟场景的认知体验, 将是未来大众化产品必须要完美处理的课题.

社会伦理. 遥远的未来, 从技术上我们也许可以完美地制造出可以媲美真实场景的虚拟, 完成真正的虚拟现实技术. 而从使用者角度出发, 可能会出现混淆虚拟和现实的可能, 即把虚拟世界中的认知逻辑, 应用于现实世界中. 如何解决这个问题, 让人类不会迷失在虚拟场景中, 不论是采用技术的还是非技术的解决方法, 都是有待探索并必须解决的关键点.

参考文献

1 Weinbaum S G. Pygmalion’s Spectacles. Wonder Stories, http://freeread.com.au/@RGLibrary/SGWeinbaum/PygmalionsSpectacles.html. 1935
2 Heilig M L. Stereoscopic-television Apparatus for Individual Use. Patent No.2955156 Oct. 1960
3 Heilig M L. Sensorama Simulator. Patent No.3050870 Aug. 1962
4 Ivan E S. The ultimate display. In: Proceedings of the IFIP Congress, Sutherland, 1965
5 Fisher S S, McGreevy M, Humphries J, et al. The virtual environment display system. In: Proceedings of the Workshop on Interactive 3D Graphics, Chapel Hill, 1987. 77–87
6 Burdea G, Coiffet P. Virtual Reality Technology. 2nd ed. New York: Wiley-IEEE Press, 2003. 464
7 Curcio C A, Sloan K R, Kalina R E, et al. Human photoreceptor topography. J Comp Neurol, 1990, 292: 497–523
8 Clark R N. Visual Astronomy of the Deep Sky. Cambridge: Cambridge University Press and Sky Publishing, 1990. 355
9 Dagnelie G. Visual Prosthetics: Physiology, Bioengineering, Rehabilitation. Berlin: Springer, 2011. 453

The challenging facts of virtual reality technology

Wenbo ZHANG*, Lei CAO & Junjun XIONG

Beijing Samsung Telecom R&D Center, Beijing 100095*, China*
*E-mail: wenbo.zhang@samsung.com

Abstract The Virtual Reality (VR) technology has been widely used in many areas such as gaming, scientific and military application from 1950s. With the computer technology development and cost reduction, the VR technology is also progressing from professional application to consumer level. The year 2016 is under the consensus of the year of VR. In this paper, the current VR technologies are analyzed from the industry and consumer point of view based on the fundamental concepts. The technical status of current consumer level VR is reviewed with respect to the challenging facts.

Keywords virtual reality, HCI, challenge

Wenbo ZHANG was born in 1979. He received his Ph.D. degree from Nanyang Technological University, Singapore in 2011. His working experience including Hewlett-Packard, Huawei and currently he is principal engineer in Beijing Samsung Telecom R&D Center working on computer vision.

Lei CAO was born in 1985. He received the Ph.D degree in pattern recognition and intelligent system from Beihang University, Beijing, in 2016. Currently, he works at Beijing Samsung Telecom R&D Center. His research interests include computer vision and image processing.

Junjun XIONG was born in 1977. He is graduated from Institute of Computing Technology, Chinese Academy of Sciences at 2005 as Ph.D. He then joined Beijing Samsung Telecom R&D Center. Currently he is a principal researcher on computer vision.