

# 第一章 虚拟现实概览

此系列为《Virtual Reality》的中文译版，涵盖了虚拟现实系统的基础知识，包括几何建模、转换、图形渲染、光学、人类视觉、听觉和前庭系统，跟踪系统，界面设计以及一些技术问题等。下面为原书 chapter 1 的中文译版。

## 1.1 虚拟现实是什么？

虚拟现实（VR）作为前沿科技，近年来发展十分迅速，这也导致设备更新换代的速率越来越快。如果我们着眼于具体的设备来研究虚拟现实，则可能很快就会被时代淘汰。因此，在这本书中，我们重点关注虚拟现实的基本理论。首先，我们必须搞清楚虚拟现实的概念，对虚拟现实进行定义要抓住它的根本，要保证定义的准确性不会随着技术的发展而轻易的改变。此外，对VR 做出的定义还应该具有普适性，不仅要适用于如今我们对虚拟现实的理解，也应包含我们对其未来形态的预测。

在开始介绍虚拟现实技术之前，我们先举两个具有代表性的例子：

1) 一个人通过挥动“翅膀”体验飞跃旧金山的感觉(图 1.1)

2) 老鼠通过在跑轮上跑动来探索投影在四周的虚拟迷宫(图 1.2)

在 1.2 节中我们还会介绍更多的样例，我们希望这些样例都能够包含在我们对虚拟现实的定义中。

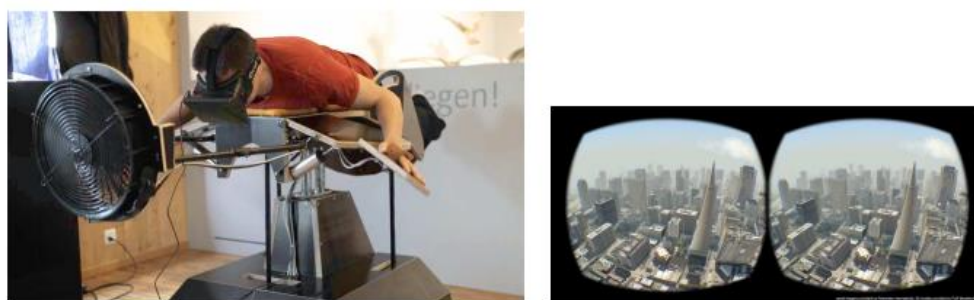


图 1.1 苏黎世艺术学院的“鸟类体验”装置，用户穿戴 VR 头盔，挥动“翅膀”体验飞跃旧金山的感觉，平台和风扇用来提供额外的感官刺激。右图展示了每只眼睛看到的画面。

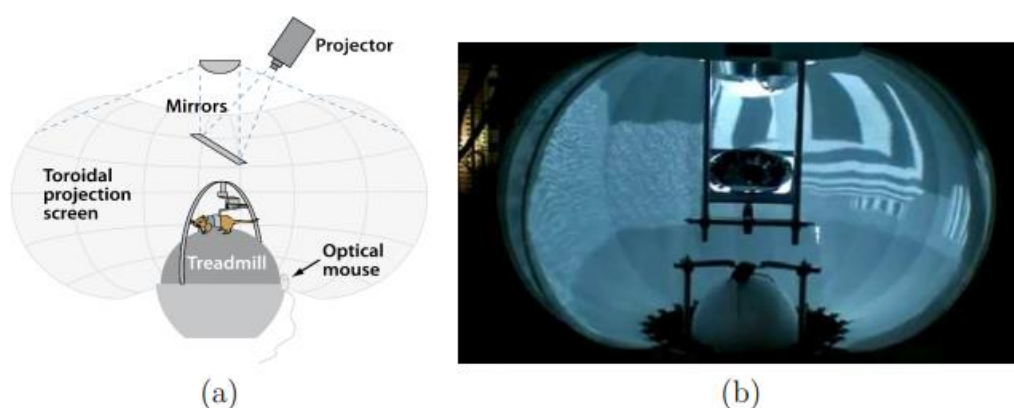


图 1.2 (a) 慕尼黑大学的神经生物学家制作的实验装置，将小鼠放在中心的跑轮上，随着小鼠的运动，四周墙壁上的画面不断变化，为其提供视觉刺激（图源于 Kay Thurley） (b)类似的实验装置图，由普林斯顿大学的研究人员制作

## 虚拟现实的定义

**虚拟现实定义：**利用人工设计的感官刺激，引诱生物体产生一种正在进行某种特定行为的感觉，而整个过程中该生物体几乎无法意识到人工干扰的存在。

定义中出现的四个关键词：

**特定行为：**生物体正在进行由人工设计的“动作”，比如说飞行、散步、探索、看电影、社交等等。

**生物体：**可能是你，他人，或是其他物种，像是果蝇，蟑螂，鱼，老鼠，猴子等等（科学家发现虚拟现实技术对上述生物体有影响）

**人工感官刺激：**通过工程学的技术，将生物体内的一种或多种感官劫持，然后使用人工的刺激来替代原有的感官输入

**意识：**整个过程中，生物几乎没有感受到人工干扰的存在，从而成功地被“骗入”虚拟世界。这会使得生物体认为自身正处于全新的世界当中。

## 作用范围

显然，图 1.1 和图 1.2 的例子，以及穿戴 VR 头盔体验虚拟场景等等均包含在我们的定义当中。那么，定义的作用范围到底有多大？用耳机听音乐算不算？在剧场看电影算不算？显然，电影和音频系统的技术已经可以给予人工的感官刺激。再进一步说，欣赏画展中的画作算是虚拟现实吗？它有着包含涂料、画布等技术。甚至看小说是不是也可以被认为是虚拟现实，它也包含了书写和印刷等技术。这些刺激虽然都是视觉上的，但是似乎不像电影场景和音频系统那样直接。其实，我们并不需要给虚拟现实定下一个明确的界限，上述的例子都值得我们讨论研究。而且，从历史角度来看的话，它们也是很好的切入点，我们会在 1.3 节详细介绍。

## 谁被骗了？

让我们回到虚拟现实的定义，“欺骗”生物体听上去似乎不切实际，但是，神经生命学的相关研究告诉我们，这是可行的。从原理上来看，当生物来到新的环境时，首先会构建一系列神经结构，这些结构由位置细胞组成，用来记忆周围的空间信息，如图 1.3(a)。当生物回到之前的环境，关于这个地方的所有位置细胞会被立即激活。虽然我们还没有完全弄懂原理，但格点细胞记忆位置信息的方式类似于笛卡尔坐标系，如图 1.3(b)。研究表明，进行 VR 体验时，生物体的神经系统也有可能构建一系列用来记忆环境的神经结构。换句话说，我们的大脑对虚拟的环境也会做出反应，生成相应的位置细胞。这表明虚拟现实或多或少地欺骗了我们的大脑。你可能会想，那么小说中描绘的环境又是否会让我们的的大脑做出反应从而生成位置细胞呢。

我们甚至还会想，我们是否一直在被欺骗，现实的环境是否只是一场梦境，是否还存在着更大的现实尚未被我们揭示。古今中外很多哲学家都讨论过这个问题，最早可以追溯到古希腊时期柏拉图《理想国》中的洞穴寓言。在寓言中，苏格拉底是这样描述的，一群整日束缚在洞穴之中的囚徒，他们面临着一堵空白的墙壁，他们身后有一堆火在燃烧，火和囚徒之间有一些人拿着器物走动，只有当人们经过时才会看到投影到墙壁上的阴影。他解释说，哲学家就像这些洞穴人，只有从洞穴中解脱出来才能看到现实的本质，而不能仅通过预测来认知现实。这个观念一直广为流传，也与宗教和灵性信仰有着千丝万缕的联系。1641 年，勒内·笛卡尔提出，我们的所见所想都是由一个全知全能的恶魔创作的幻象。1973 年，吉尔伯特介绍了“缸中的大脑”（图 1.4）的思想实验，来具象化“笛卡尔恶魔”的行为。这听起来很像是《黑客帝国》，没错，这部电影正是在这个思想实验的影响下创作出来的。电影中，一个名为“矩阵”的计算机人工智能系统控制着人们的身体，并将人们的大脑连接到一个模拟世界当中。主角 Neo 必须决定是否要回到真正的现实，或是喝下记忆消除药丸幸福地活

在梦境之中。

### 有关“现实”的术语

“虚拟现实”的概念最早由德国哲学家伊曼努尔·康德提出，尽管当时还未出现相关的技术。20 世纪 80 年代由 Jaron Lanier 进行进一步的推广。现如今，虽然虚拟现实已经广泛应用到各行各业，但是一些术语仍然有所争议。比如说*虚拟环境*，它早于*虚拟现实*被广泛使用，也更受到研究人员偏爱。它通常被认为与*虚拟现实*同义，然而，在本书中我们强调，这两个词还是有差异的，在*虚拟现实*中，我们对环境的感知不都是源于完全合成的世界，也有可能来自镜头捕捉到的“真实”世界，因此，这个环境并不一定是“虚拟”的。*增强现实* (AR) 指的是视觉刺激通过镜头或照相机直接进入眼睛，并且叠加一些额外的内容。*混合现实* (MR) 的概念有时被认定为一种技术组合，包含*虚拟现实*、*增强现实*、和现实世界，但近年来 VR/AR/MR 的含义也可以互相指代。*远程呈现*指的是让用户感觉自己在现实世界中的其他地方的系统。在此基础上，如果用户可以控制一些设备，比如无人机等，那么用*远程操控*描述更加合适。按照我们的说法，*虚拟环境*、*增强现实*、*混合现实*、*远程呈现*、*远程操控*都可以被视为*虚拟现实*完美的应用实例。必须要指出，虚拟现实的关键在于通过工程学方法改变用户对现实的感知，而不是用户认为环境是真实的还是虚拟的，因此，我们还可以用*感知工程*来描述。

从字面上来看，“虚拟”和“现实”似乎是矛盾的，这看起来像是一个哲学问题，【3】中也将其修正为*虚拟*。尽管如此，我们仍要继续使用*虚拟现实*这个词语，同时，我们要注意区分：*真实世界*指的是包含用户的物理世界，*虚拟世界*指的是在虚拟现实体验中所感知到的世界。

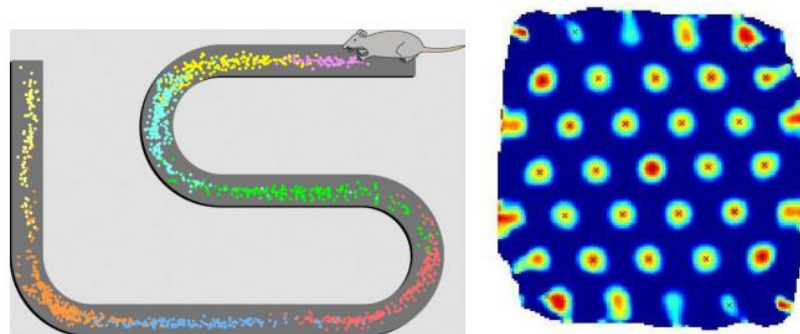


图 1.3 (a) 动物体将神经元分化为位置细胞，当其回到有记忆的地方时会被激活。图中描绘了大鼠沿着特定路线进行往返运动时，脑中 8 个位置细胞的空间分布。(图源于 Stuart Layton)。(b) 在空间中均匀分布的格点细胞，用于对位置坐标进行编码记忆 (图源于 Torkel Hafting)。

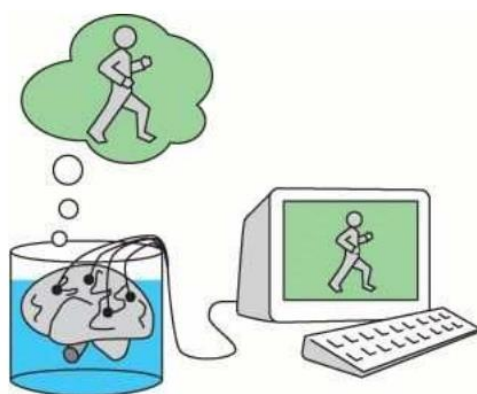


图 1.4: 关于虚拟现实的思想实验: 缸中的大脑, Gilbert Harman 于 1973 年提出 (图源于 Alexander Wivel)。

## 交互性

虚拟现实系统还有另一个特性: 交互性。感官刺激是否依赖于机体采取的行动? 如果不是, 那么 VR 系统是一个开环系统; 否则, 它是闭环的。在闭环 VR 系统中, 机体对刺激有一定程度的反馈, 表现在刺激可能会随着身体的运动而不断变化, 包括眼睛、头部、手部或腿部的运动; 也可能由于语音命令、心率、体温和皮肤电导性 (是否出汗?) 等原因而有所变化。

## 第一人称 VS 第三人称

如果你正在读这本书, 我猜测你很可能想要开发 VR 系统。那么, 你一定要注意关于第一人称/第三人称的事情。研究人员设计生物实验时, 分工是清晰的: 实验体是体验者, 是第一人称; 研究人员是观察者, 是第三人称。他们通过设计 VR 系统, 进行实验来验证自己的科学假设。比如说, 抑制大鼠脑中的部分神经元, 观察它的导航能力是否收到影响? 如何被影响? 与研究人员不同, 当工程师或开发人员设计 VR 系统时, 他们通常将自己代入到系统中, 以便于对系统进行评估和改进, 也就是说, 在“实验体”和“研究人员”两种角色间不断进行切换。读完整本书你会发现, 这种做法非常的糟糕。这些设计者想要取得效果的同时避免重复的实验, 但是, 他们本身知道系统会给予人一种怎样的体验, 会有什么样的效果, 这种意识会导致他们的感受与首次接触的人的感受相差很大。更糟糕的是, 设计者会逐渐地在生理或心理上适应系统的某些缺陷, 导致这些缺陷不能被及时发现并解决。日常生活中也有类似的经历, 比如说, 我们很难预测其他人会对自己的文章做出怎么的反应; 还有, 相比他人的文章, 我们更难以正确地审视自己的文章。在虚拟现实的研究中, 这些影响会变得更加严重, 你必须时刻注意这一点。当你那些始终信任的感官被劫持时要小心, 你将会进入到一个完全未知的领域。

## 比现实更“真实”?

VR 的体验到底应该有多么真实? 毫无疑问, 让它无限接近我们的现实世界是很诱人的想法。在 10.1 节中, 我们会详细介绍 *通用模拟原理*, 是指现实世界中的任何交互机制都可以在虚拟现实中进行模拟。我们的大脑对这些机制最为熟悉, 因此很容易进行模拟。这种理念在游戏行业中非常火热, 例如, 在第一人称射击游戏 (FPS) 的开发中, 随着显卡性能的不断提升, 游戏画面也被渲染的非常漂亮。但这并不意味着卡通式的环境一无是处, 有的时候, 简单的环境也许是更可取的选择。纵观历史, 有很多这样的例子, 我们会在 1.3 节中详



细介绍。

作为设计整个 VR 系统的人，你需要考虑你想要用户获取什么样的信息，会做出怎样的反应，以及期望的效果是怎样的。经过精心的设计，你甚至可以使得整个体验比现实世界还好。用户可以体验什么？上数学课？观看文艺表演？开发软件？设计房屋？异地情侣之间的交流？玩游戏？冥想放松？到地球甚至宇宙中去旅行？对于每一种情况，不同的需求也会有产生不同的体验，要结合实际需求考虑。例如，在 VR 环境中开发软件。我们现如今开发软件的方法是在电脑里面敲代码，但要知道，在 20 世纪 50 年代这是难以想象的。通过 VR，我们可以完全模拟如今的开发环境，让程序员认为自己坐在办公桌前，面对电脑屏幕敲代码。然而，这有些“大材小用”，我们可以在 VR 中创建几乎任何东西，无需拘泥于现有的思维定式。我们可以设计一个全新的环境，不是在办公桌前，而是让屏幕漂浮在海滩上、倒映在森林中。此外，我们还可以设想，是否可以让调试器把整个程序的运行轨迹显示出来？怎样显示效果会更好？

## 合成 VS 捕捉

设计 VR 系统，我们首先要构建一个虚拟世界。构建的方法主要分为两种，第一种，我们可以构建一个合成世界，完全由几何原型和物理仿真制作，这在视频游戏中很常见，我们可以认为，这是几十年前 VR 体验的主要方式；另一种，我们可以用镜头捕捉真实场景，再用成像技术将其还原出来。在一个多世纪前，人们已经做到了将摄像机拍摄到的画面投射到屏幕上观看。如今，我们捕捉全景图像和视频，通过 VR 系统可以从任何角度观看，这也是摄影技术的延伸。然而很多情况下，将真实场景投影到相机传感器上会丢失很多图像信息，当用户的头部位置和观察区域发生改变时场景会怎样变化？我们应该捕捉更多的信息来尽可能的还原真实场景。通过深度传感器和 SLAM（即时定位与地图构建）技术，我们可以获取到空间定位的信息，完成地图的构建或场景的生成。然而，想要还原出完全精确的真实场景非常困难，除非整个空间是专门为此设计的。（例如，动作捕捉工作室）

与用户进行互动的时候，跟踪他们的动作并及时作出反应是 VR 系统的关键，这也是我们需要捕捉的信息。穿戴 VR 头盔时候他们的面部表情是怎样的？我们是否应该看懂他们的手势？他们的情绪是怎样的？他们是否将注意力集中系统中？必要的话，我们可以在虚拟世界中构造一个化身，以合成的方式模拟注意力或情绪状态，此外通过运动跟踪和其他属性来增强这个化身，使其更接近用户，以此来进行人机交互并且在一定程度上保护了用户的隐私。然而，这由引出了另一个著名的哲学问题——恐怖谷理论，指的是当机器人与人类相像程度超过一定程度，由于机器人与人类在外表、动作上都相当相似，所以人类亦会对机器人产生正面的情感。直至到了一个特定程度，人类的反应便会突然变得极之反感。哪怕机器人与人类有一点点的差别，都会显得非常显眼刺目，让整个机器人显得非常僵硬恐怖。目前，用户还没办法很好地把自己在 VR 系统中的所见所感分享给其他人，当然，他们可能也不想这么做。

## 健康和安全

随着目标的不同，我们对 VR 系统可能会有各种各样的要求，但唯一不变的就是用户的健康和安全。与广播或电视等传统媒体不同，VR 有可能会影响到感官和大脑，导致疾病疲劳。几十年前，已经有学者在研究模拟晕动症，本书中，我们将 VR 应用中产生的不良症状统称为 VR 疾病。尽管有时候 VR 疾病是由于低端的软硬件导致的，但大多数情况是由于开发人员没有重视到 VR 的副作用，这也正是为什么这本书的大部分内容都是有关人类生理学和感知心理学的。为了开发出令人舒适的 VR 体验，这些都是必须要了解的。从原理上讲，疲劳是由于给予感官的异常刺激难以被大脑整合理解导致的。有时候，接受到与预期不一致

的刺激，或是其他感官给予的刺激，甚至会导致头晕恶心。

导致疲劳的另一个因素是 VR 互动需要大量的肌肉运动，比方说，你想要移动沙盒游戏中的物体，那么你必须不停地摆动你的胳膊，这会导致疲劳和大猩猩臂效应，即人很难在手臂往外展开的姿势下长时间操作。鼠标的发明正是源于此，通过小而轻微的操作来控制屏幕上光标的移动。经过一段时间的适应期，我们的大脑就会将这两种动作结合起来，使整个过程变得自然，并且大大的减少了疲劳。

## 1.2 现代 VR 体验

当下的 VR 时代是由智能手机行业的显示，传感和计算技术进步而带来的。从 Palmer Luckey 2012 年设计的 Oculus Rift 到简单地在智能手机上构建此功能[10,16,19]，随着 VR 头戴设备的大量生产并发布世界各地，这个世界已经迅速发生了变化。这种趋势在许多方面与家用电脑和网络浏览器革命相似；随着越来越多的人接触到这项技术，他们也正利用 VR 来创造更多可能。

这节内容能让你快速了解当今人们正利用 VR 做什么事，并且提供一个找寻在互联网方面相似的切入点。在此，我们只能用文字和图片来描述体验，它还是会与你自己亲身体验所带来的惊喜相差甚远。书类的纸质媒体对于充分传达虚拟现实技术来说还是极其不足的。也许这就像 1890 年代在报纸上描述一个电影院是什么样的！如果可能的话，我们强烈建议你亲身去尝试体验一些 VR 技术，从而形成一种对其直观的印象并且激发你的想象力去实现更好的功能。

### 视频游戏

人们几十年来一直梦想着进入视频游戏的世界。在 1982 年以前，这种观念已经因为被迪士尼电影 Tron 而流行起来。图 1.5 展示了一些 VR 游戏的体验。大多数玩家现在都追求通过一个虚拟实景探索更大，更现实的世界。图 1.5(a)为 Valve's Portal 2，它是在 HTC Vive 头戴式设备上开发的一个冒险类的解谜游戏。图 1.5(b)为一个全方位的跑步机外设装置，当人们的脚在地面上滑动时，它可以提供给使用者一种走路的感觉。这两个例子都提供了用户一种当事人的视角。相反，图 1.5(c)展示了 Lucky's Tale，它相反地提供了一个舒适的第三方视角当用户看起来像是要漂浮在她控制的角色之上。图 1.5(d)与其他图片都不同，它被明确地设计用来利用 VR 的力量。

### 沉浸式影院

好莱坞电影持续提供升级版的现实体验。为什么不让观看者感觉到他们就是场景中的一部分呢？图 1.6 展示了一个沉浸式的小故事。电影导演已经进入到一个关于电影的迷人的新时代。20 世纪学到的技巧需要被重新研究，因为它们建立在电影摄影师控制摄影机视角的假设上。在 VR 中，观看者们可以选择任何的视角，甚至可以在场景中走动。为什么他们能够这样做呢？你怎么确保他们不会错过故事中的部分场景？这个故事应该是固定的吗，还是应该对观看者的行为有所调整适应？观看者在影片中应该是当事人吗，还是一个对别的人物不可见的旁观者？怎么能让一群好友同时体验 VR 影片？什么时候模拟相比于真实场景的捕捉会更真实合适？

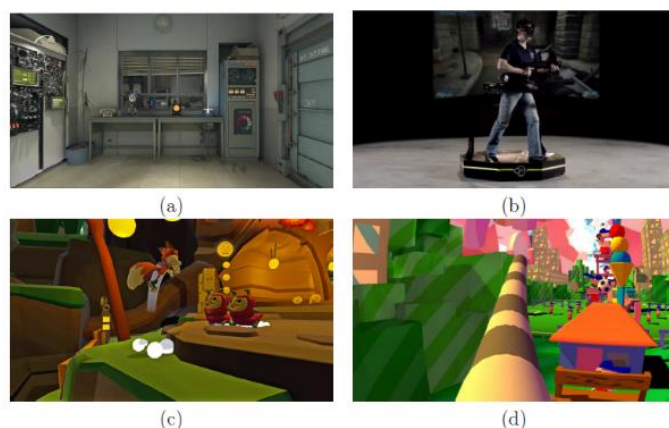


图 1.5: (a) Valve's Portal 2 为 HTC Vive 头戴式设备设计的演示是一款在虚拟世界中解谜的体验游戏。  
 (b) 支持第一人称穿越射击游戏的 Virtuix Omni 跑步机。(c) 当玩家在角色上方浮动时, Oculus Rift 上支持的 Lucky's Tale 保持第三人称视角。(d) 在 DePaul 大学开发的 Dumpy 游戏中, 玩家似乎有一个大象鼻。游戏的目的是通过摆动的象鼻将东西击倒, 从而享受这种不寻常的体验。

解决这些现存的以及之后会出现的无数问题需要花费很多年的时间。在这期间, VR 也可以用来作为一种现存电影的“包装”。图 1.7 展示了 VR 影院的应用, 它允许用户在虚拟的影院中任意选择座位。无论在用户硬盘上的是何种标准的电影或是视频文件, 都可以被传输到影院的屏幕。他们可以是 2D 也可以是 3D 的。背后的投影仪发出闪烁的灯光并且音频会被调整以模仿真实剧场的音效。它提供了一个快速的途径来转化所有原本创作用来在屏幕上观看的内容, 并且将它们变为 VR。许多简单的扩展可以在不修改影片内容的前提下实现。举个例子, 在一个关于僵尸的电影中, 一些虚拟的僵尸能进入到剧场并开始追逐你。在一个关于龙卷风的电影中, 也许剧场会被撕裂。你还可以体验社交。想象跟来自世界各地的朋友在虚拟的剧场中坐在一起, 共同度过一个“电影之夜”。你们甚至可以在剧院里体会到捣乱的快感, 而不会被赶出去。

## 远程呈现

营造一种我们仿佛在其他地方的第一步是捕捉远程环境的全景(图 1.8)。谷歌的街道场景和地球应用就是依赖于从全世界各地几百万个地方捕捉的全景图像。查询街景服务器的简单 VR 应用程序会使用户直接感觉他仿佛站在这些位置, 同时能轻松地在附近的位置之间转换(图 1.9)。全景视频的拍摄更加引人注目。图 1.10 展示了一个沉浸式摇滚音乐会的框架。更好的方式是提供现场全景视频界面, 人们可以通过该界面参加体育赛事和音乐会。实时界面使交互变为可能。人们可以把视频会议发展至一个让遥远的参会者仿佛在场出席会议的新阶段。通过在机器人上装备全景相机, 用户甚至可以在遥远的地方自由移动。现在的 VR 技术让我们可以虚拟地去到许多遥远的地方并且能进行大多数以往只能通过亲身到场才能进行的交互活动。这会提供给大家远程办公的机会。这最终可能有助于扭转 19 世纪工业革命引发的城市化趋势, 当人们更均匀地分布在世界各地市实现去城市化。



图 1.6: Oculus Story Studio 制作了艾美奖获奖作品 Henry, 这是一个简短的沉浸式故事, 讲述的是一只不受喜爱的刺猬, 希望能结交一位新朋友, 就是屏幕前的观众。



图 1.7: 由 Joo-Hyung Ahn 为 Oculus Rift 制作的 VR 电影。 你可以选择座位并观看喜欢的任何电影。



(a)



(b)

图 1.8: 实现远程呈现的重要组件是捕获全景图: (a) 顶部有摄像头和深度传感器的汽车, Google 使用它来制作街景视图。(b) Bublcam 是一种便宜便携的方式来捕捉和流式传输全景视频。



(a)



(b)



图 1.9：简单的 VR 体验，通过 VR 头戴式设备呈现 Google 街景图像：（a）巴黎的熟悉场景。（b）在设备内部创建左眼和右眼视图，同时还考虑到用户的观看方向。



图 1.10：Jaunt 拍摄了 Paul McCartney 音乐现场的全景视频。Let Die 提供了让用户感到他们仿佛与摇滚明星一同在舞台上的 VR 体验。



图 1.11：机器人头戴设备的例子：（a）宾夕法尼亚大学的 DORA 机器人模仿用户的头部运动，允许他在保持立体视图的同时在遥远的世界中环视（全景是单色的）。（b）Plexidrone，一款低成本飞行机器人，专为流式全景视频而设计。

## 虚拟社会

当远程呈现技术让我们感觉到仿佛处于物理世界的另一部分时，虚拟现实还允许我们构建出整个社会，提醒我们这个物理世界的存在，但实际上只是包含与真人相关的虚拟形象的合成世界。图 1.12 显示了人们通过头戴设备在幻想世界中互动的第二人生场景；此类体验最初设计为在屏幕上观看，但现在可以通过 VR 体验。由于各种原因，一群人可以在这些空间中共度时光，包括共同的特殊爱好，教育目标或者只是想逃离普通的生活。

## 同理心

VR 提供的第一人称视角是一个让人们对他人的状况有同理心的有力工具。世界持续在为使不同种族，宗教，年龄，性别，性取向，社会地位和教育被接受和获得平等而斗争，而进步的最大障碍是大多数人无法理解拥有不同身份的感觉。图 1.13 展示了一个由联合国赞助的 VR 项目旨在设身处地体会那些陷入 2015 年叙利亚危机的人们的感受。我们之中的一些人可能会对别人的不幸感到同情，但是真正设身处地体会到他们的挣扎将会是一种更强烈

的感受。图 1.14 是一个允许男人和女人互换身体的 VR 系统。通过虚拟的社会，我们可以探索更多的可能。如果你比其他任何人都矮十公分将会怎样？如果你以不同的性别教授你的课程将会怎样？如果你是警察种族歧视的受害者呢？使用虚拟现实技术，我们可以想象许多如果你不在“合适的”群体中，是无法体会到的“生命游戏”。



图 1.12：虚拟社会通过在虚拟世界中相互作用的交互设备发展，这些虚拟世界由一台通用服务器维护。第二世界的快照如图所示。



图 1.13：在西德拉云上，电影制片人克里斯牛奶为叙利亚难民的痛苦提供了第一人称视角。

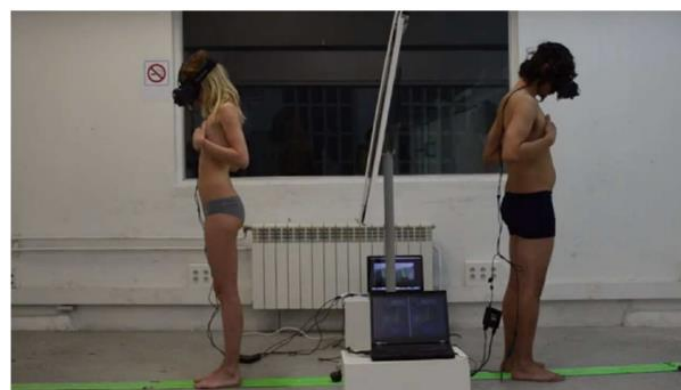


图 1.14：巴塞罗那的学生提供了一个可以与其他性别交换身体的体验。每个人都戴着一台 VR 头戴式设备，前端安装有摄像头。因此，每个人都从另一个人的近似观点看世界。他们被要求以协调一致的动作移动双手，以便他们看到他们新身体的适当移动。

## 教育

除了教会同理心，第一视角也可以对教育的许多方面进行革新。在工程学、数学、和科学方面，虚拟现实提供了把难以解释的概念或数据中几何关系可视化的机会。此外，虚拟现实技术自然适合于实践培训，因为在现实的虚拟环境中开发的技能可能自然地转移到真实的环境中。由于有时真实环境的成本高昂或会产生健康风险，所以开发此技术的动机尤其强烈。VR 中最早也是最常见的训练例子之一是飞行模拟（图 1.15）。其他例子包括消防，核电站安全，搜救，军事行动和医疗程序。



图 1.15: 美国空军使用的飞行模拟器（Javier Garcia 拍摄）。用户坐在物理驾驶舱中，同时被显示环境的显示器包围。



图 1.16: 参观亚述国王 Ashurnasirpal II 的 Nimrud 宫殿，这是一个由 Learning Sites Inc.和伊利诺斯大学开发的 VR 体验。

除了这些 VR 的常见用途之外，也许对于 VR 教育来说最大的机会在于人文学科，包括历史学、人类学和外语的学习。考虑一下读一本英国维多利亚时代的书籍和有机会漫游于由历史学家精心拟建的 19 世纪伦敦街头之间的区别。我们甚至可以去到一个由废墟重建出的古老城市（图 1.16）。迷人的可能性体现在我们既可以通过虚拟现实界面观看实体博物馆，也可以直接在虚拟博物馆中扫描和展示文物。这些例子都属于数字遗产。

## 虚拟原型

在现实世界，我们建立原型来理解一个提出的设计模型具体是怎样工作的。由于 3D 打印跟相关技术的出现，这项工作比之前容易了许多。同时，虚拟原型设计使设计师能够在包



含原型的虚拟世界中居住（图 1.17）。他们可以快速地进行交互并修改。它们也能够带客户进入他们构造的虚拟世界以便交流想法。想象你想要重新改造你的厨房。你可以在 VR 中构建一个模型并且向承包人解释它看起来应该是怎么样的。VR 的虚拟原型设计在商业中也有重要的用途，包括房地产、建筑业、以及飞机，航天器，汽车，家具，服装和医疗器械的设计。

## 健康医疗

虽然健康和安全是颇具挑战性的 VR 话题，这些技术还是能够帮助我们改善健康状况。分布式医疗的发展趋势越来越明显，医生们培训人们在遥远的世界各地执行常规医疗程序。医生们能通过远程呈现提供指导建议，并且运用 VR 技术来培训。在 VR 的另一用途中，医生们可以让自己沉浸在利用医疗扫描数据生成的 3D 器官模型中（图 1.18）。通过在手术前简短地了解病人的身体状况，能让他们更好地规划和准备医疗步骤。这也能更好地向病人或其家人解释医疗方面的一些选择以便他们做出更加知情可靠的决定。当下一个另外的应用是 VR 可以直接提供理疗以帮助病人，包括通过重复地暴露克服恐惧症和压力失调，改善或维持认知技能以对抗老化，并且改善运动技能以克服平衡，肌肉或者神经系统紊乱失调等问题。虚拟现实系统有一天还可以通过使老年人虚拟旅行，进行有趣的物理治疗，并通过与亲朋好友通过界面进行互动，使他们在远程活动中感受到自己的存在和参与，来克服孤独感从而延长寿命。

## 增强和混合现实

在一些应用中，用户可以通过叠加一些附加图形来观看真实世界，以增强其外观；如图 1.19。这被称为增强现实或者混合现实（我们在此书中都将它们归为 VR 的一部分）。通过在现实世界中放置文本、图标以及其他图形，用户可以利用互联网的力量来辅助许多操作，例如导航社交互动和机械维护。迄今为止的许多应用都旨在帮助企业更有效地开展业务。想象一下，在一个工厂的环境，工人可以在需要组装的零件上看到标识，或者他们可以直接看到机器的内部来确定潜在的需要替换的零件。







图 1.17: 建筑是虚拟原型具有无价价值的最佳例子。这个名为 Ty Hedfan 的演示由设计师 Olivier Demangel 创建。真正的厨房在上面, 虚拟厨房在下面。

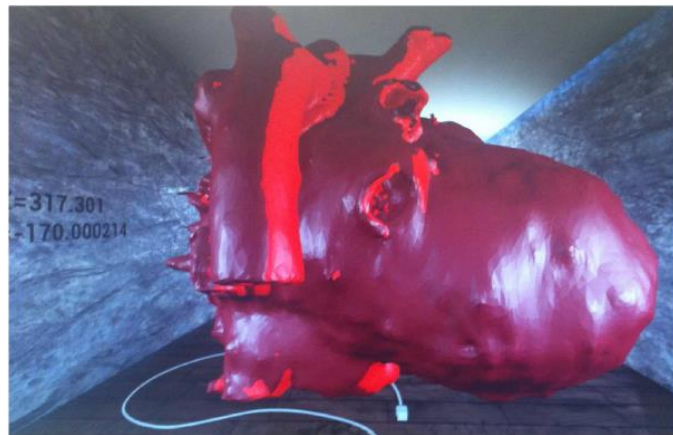


图 1.18: 基于真人心脏图像的心脏可视化系统。这是由 Jump Trading 模拟和教育中心以及伊利诺伊大学开发的。

这些技术都十分依赖于先进的计算机视觉技术。在确定如何绘制仿佛像自然嵌入的虚拟对象之前, 必须先识别对象, 重构形状, 并识别真实世界中的光源。由于视觉算法在不可预见的环境中经常出现错误, 所以实现高度的可靠性就成为了一项挑战。必须先估计真实世界的光照条件来确定如何绘制虚拟物体, 以及它们可能投射到环境的实际部分以及其他虚拟物体上的任何阴影。此外, 在某些应用实例中, 真实和虚拟的物体都需要被完美地进行排列, 这给追踪和计算机视觉系统都会带来很大的负担。

在虚拟现实中还存在着这几种可能。一个固定屏幕可以显示通过 3D 眼镜增强的图像。数字投影机可以通过将光照射到物体上来增强环境, 为它们提供新的颜色和纹理, 或者在真实世界中放置文本。作为智能手机或平板电脑一部分的手持式屏幕可以用作增强或混合世界的窗口。这是流行的任天堂 *Pokemon Go* 游戏的基础; 图 1.20。与本书更相关的案例涉及将显示器安装在头上。在这种情况下, 存在两种主要的方法。在透视显示器中, 用户只需要通过透明材料就能看到大部分真实世界, 而虚拟对象出现在显示器上以更改部分视图。近期采用先进透视显示技术的头戴式设备包括 Google Glass, Microsoft HoloLens 和 Magic Leap。实现高分辨率, 宽视野和阻挡入射光的能力仍然是经济实惠的消费级设备面临的重大挑战; 然而, 它可能会在几年内得到很好的解决。另一种方法是直通显示器, 它将图像从外置摄像头发送到头戴式设备内部的标准屏幕。直通显示器克服了当前的透视显示问题, 但是却存在延迟,

光学失真，颜色失真和有限的动态范围等问题。



图 1.19: 微软 HoloLens 使用先进的透视显示技术将图形图像叠加到普通的物理世界中，如透过眼镜看到的那样。



图 1.20: 任天堂的 Pokémon Go 是一款网络游戏，它允许用户想象一个叠加在现实世界中的虚拟世界。他们只能通过他们的智能手机屏幕才能看到口袋妖怪角色。



图 1.21: (a) Epic Games 通过虚拟客厅创造了一个狂野的过山车。(b) 断头台模拟器由 Andre Berlemont, Morten Brunbjerg 和 Erkki Trummal 制作。参与者在刀片掉落时被朋友击中了脖子，他们可以在头部滚动时获得适当的视角。

## 新的人类体验

最后，VR 的重点可能只是提供一种新的人类体验。通过远程呈现，人们可以通过机器人或其他人的眼睛体验不一样的世界。然而，我们还可以进一步给予人们在现实世界中不可能（或者致命）的体验。大多数情况下，艺术家们是作出这一努力的人。如同鸟一般的人类飞行体验（图 1.1）就是一个很好的例子。图 1.21 显示了另外两个。如果我们改变规模会

怎么样？想象一下，身高只有 2 毫米，看着正对着脸上的蚂蚁。相比身高 50 米，高耸于城市而人们尖叫着跑过你。如果我们模拟系统中药物的作用会怎样？如果你能成为你最喜欢的动物呢？如果你成了一块食物呢？艺术家的创作可能性似乎是无止境的。我们只是受限于身体机能，令人兴奋的冒险其实就在眼前！

### 1.3 VR 的历史发展

#### VR 究竟从何而来？

VR 的历史可以从一个很早的时间开始说起，这个时间虽然早于普遍意义上理解的 VR 开端的时间，但其中包含了许多对 VR 来说至关重要的方面。很久以前，我们的祖先看着墙壁的画，想象着一个三维的故事世界，图 1.22 显示了一些这样的例子。如图 1.22（a）所示为距离 3 万年前洞穴绘画，图 1.22（b）展示了欧洲中世纪的一幅画，与洞穴绘画类似，但它之中有更多的绘画场景细节，因此留给观赏者的想象力较少。图 1.22（c）显示了一幅后期绘画，其中透视图被画家精心诠释，塑造出一座美丽的宫殿景观，它甚至不需要人们的想象就可以将其视为 3D 的场景。到了 19 世纪，许多艺术家已经厌倦了这种现实主义，并开始了有争议的印象派运动，图 1.22（d）展示了其中的一个例子，这些绘画留给观看者的想象更多，就像早期的洞穴绘画一样。

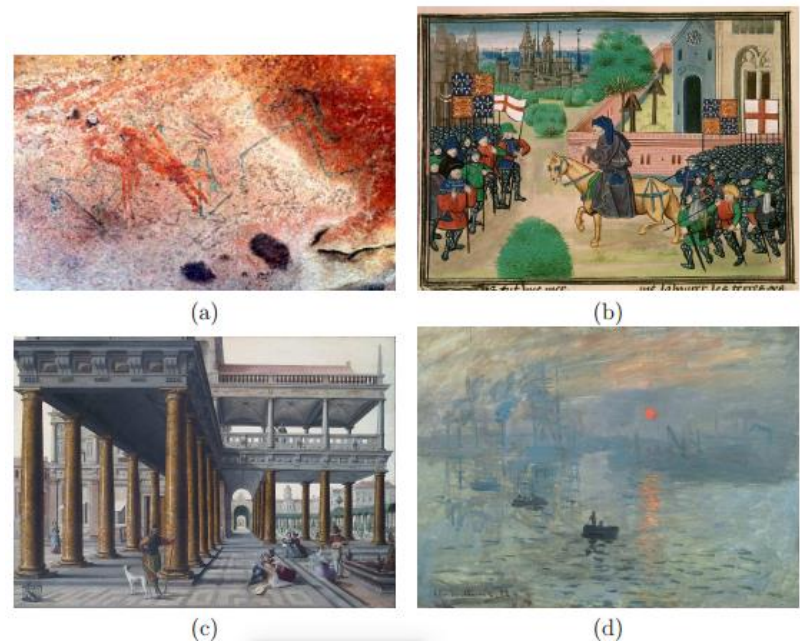


图 1.22:（a）一幅来自印度 Bhimbetka 岩石庇护所的 30,000 年的绘画作品（印度考古调查照片）。（b）约 1470 年左右的一幅英国画，描绘了约翰·鲍尔鼓励泰勒泰勒反叛者（未知艺术家）。（c）1596 年 Hans Vredeman de Vries 的一幅画。（d）1874 年 Claude Monet 的印象派绘画作品。

#### 运动的画面

频闪观运动现象是我们今天称之为电影或动态图像的基础。通过一系列图片快速翻转会产生运动错觉，即使速度低至每秒两张。若每秒超过 10 张照片，动作甚至看起来是连续的，而不是被视为单独的照片。最早的例子之一就是 1878 年爱德华·慕布里奇（Eadward Muybridge）创作的赛马电影，见图 1.23。



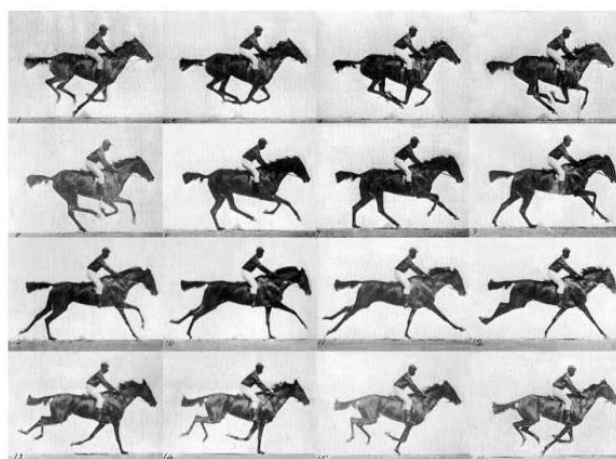


图 1.23: Eadward Muybridge 拍摄的这部 1878 年的动画电影是通过沿轨道均匀放置 24 个摄像机并在马匹经过时通过绊索触发它们而创建的。这部动画是在电影放映机的前身 zoopraxiscope 上播放的，但机械上类似于电唱机。

电影技术迅速得到了改善。到 1896 年，电影院里已经有一些 3D 电影的雏形，比如在一部影片中，一列火车进入火车站的场景会让观众们相信火车即将撞上它们（图 1.24 (a)），但是那时候还没有声音。随着观众期望的增加，特殊效果所产生的现实主义程度也随之增加。1902 年，观众可以在电影上观看到“月球之旅”，虽然画质较为拙劣（如图 1.24 (b) 所示），但到了 2013 年，要保持观众相信的必要条件是高度的现实还原主义（如图 1.24 (c) 和 1.24 (d) 所示）。

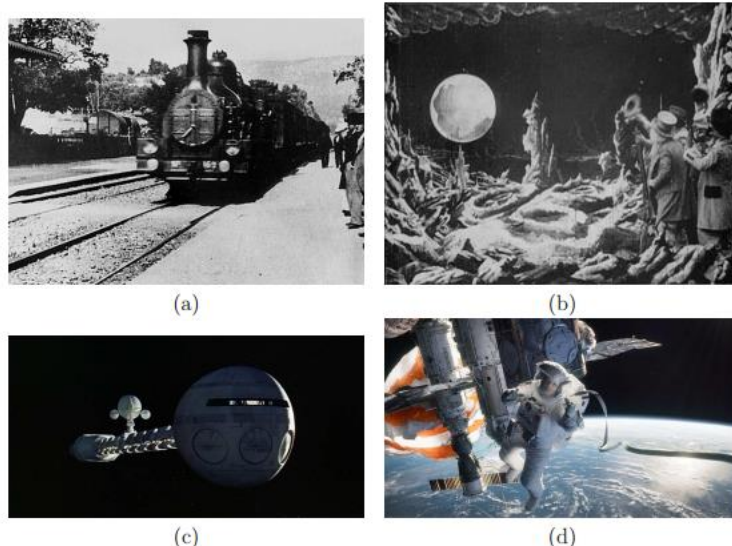


图 1.24: 一系列特殊效果: (a) 1896 年在拉西奥塔火车站列车。(b) 1902 年的月球之旅。(c) 1968 年的 movie2001。(d) 重力，2013 年

与此同时，电影观众也愿意接受现实主义程度较低的作品。就如同绘画一样，它们之间相同的原理就是留给观众的想象力更多。动画（也称为动漫或漫画）的流行就印证了这个道理（图 1.25）。从 1908 年 Fantasmagorie 的简单线条图起，动画开始萌芽（图 1.25 (a)）；随着 1928 年米老鼠的推出，动画出现更多细节部分的元素（图 1.25 (b)）；到 2003 年，动画电影取得了更高度度的现实主义（图 1.25 (c)）。不过，简单的动画也受到广泛的欢迎（如图 1.25 (d)）。



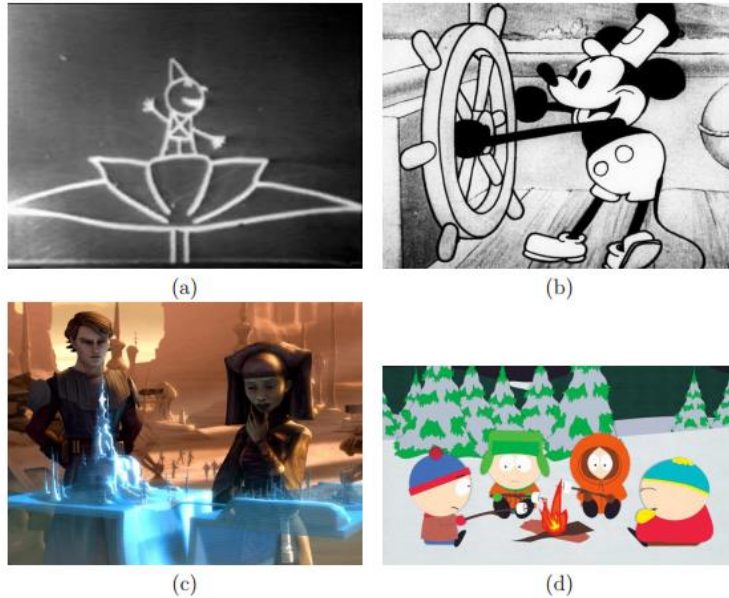


图 1.25：漫画的进展：(a) Emile Cohl, *Fantasmagorie*, 1908 年。(b) 1928 年 *Steamboat Willie* 的米老鼠。(c) 克隆战争系列, 2003 年。(d) 南方公园, 1997 年

### 成本与便携性

成本和便携性是让人进一步接受较低程度的现实主义的主要原因。如图 1.26 所示，家人愿意聚集在电视机前观看家中的免费广播节目，尽管他们知道可以在剧院观看高分辨率以及全景的 3D 电影。但对于我们目前的预期，这种显示在小小电视机屏幕上的画面依旧是观看体验上的缺陷。因此迫切亟待的要求是：1) 必须去某个地方观看它，2) 能够在家中观看它，3) 能够随身携带它。无论是图片，电影，电话，电脑还是视频游戏，以上提到的三种需求都是发展的趋势。因此我们期望 VR 系统也是如此的发展趋势。与此同时，我们期望从便携式设备获得的质量也需要达到相同的高质量水平。



图 1.26：尽管有大屏幕的电影院可用，但家庭也很愿意聚集在电视机周围，这些电视机产生的观看质量在目前的标准下是无法忍受的，如 1958 年的这张照片所示。

### 视频游戏

运动图像从某种程度上会产生一种被动的旁观者体验，相比之下，拥有交互性质的视频游戏更接近主观体验。视频游戏是闭环 VR 的重要一步，而动态图像则是开环的。如图 1.27

所示，我们看到了从简单到增进现实，然后回归简单的趋势。最早的视频游戏如 Pong and Donkey Kong，给人留下了很大的想象空间。像 Doom 这样的带入性质的射击游戏为玩家提供了主人公视角，并对接下来的十年中的游戏发展造成了深远影响，这种游戏的重点旨在提高图像质量和逼真度。刺客信条展示了一个现代的、现实的视频游戏的典型场景。与此同时，关注简单这个性质，也出现了很多广受欢迎的游戏。愤怒的小鸟看起来像 20 世纪 80 年代的游戏，Minecraft 允许用户创建和居住由方块组成的世界。减少现实主义往往导致更简单的任务要求。在 2015 年，先进的 FPS 游戏可能需要强大的 PC 和图形卡，而简单的游戏则可以在基本的智能手机上运行。因此，游戏做的越逼真不一定越好。

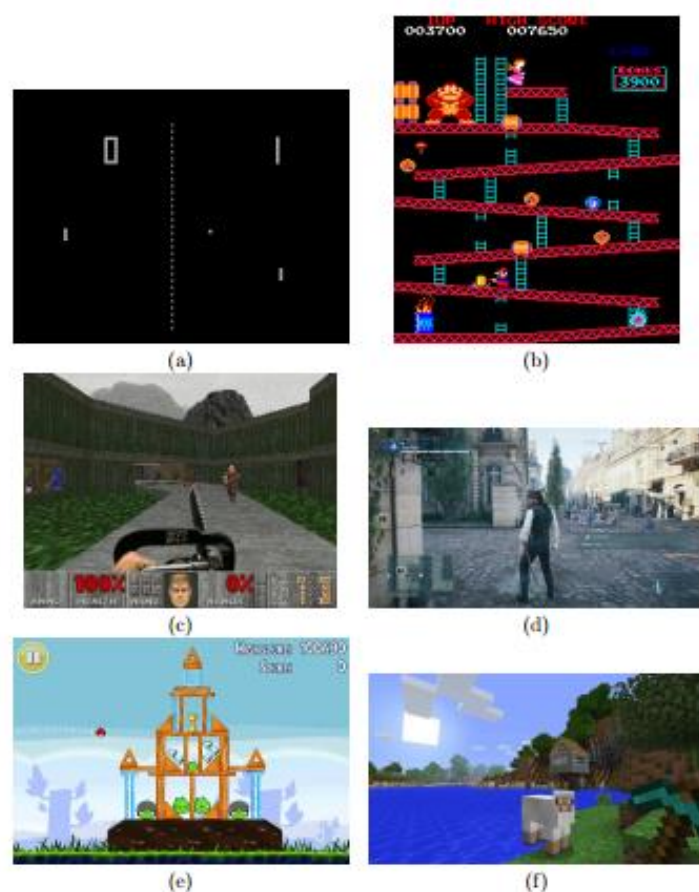


图 1.27：视频游戏的进展：(a) Atari 的 Pong，1972 年。(b) 任天堂的 Donkey Kong，1981 年。(c) id Software 的 Doom，1993 年。(d) Ubisoft 的刺客信条团结，2014 年。(e) Rovio Entertainment 愤怒的小鸟，2009 年。(f) Markus “Notch” Persson’s Minecraft，2011

### 观看视野的突破

到目前为止，我们所提到的观看概念仍然是停留在一个矩形的平面上。若是想要突破这个概念，有两个重要的方面需要考虑：1) 向每只眼睛呈现单独的图片以引起“3D”效应。2) 增加视野，使用户不会被边界影响而分散注意力。我们的大脑从眼睛推断物体的距离的一种方式立体视觉，通过观察和匹配世界上左右眼可见的特征来获取信息。它们在视网膜上的图像之间的差异产生关于距离的信息。1838 年，Charles Wheatstone 在一个称为立体镜的系统中进行了立体视觉的 3D 效果（图 1.28 (a)）。到 20 世纪 30 年代，便携式立体镜成为当今被称为 View-Master 的成功商业产品（图 1.28 (b)）。沿着这个想法进一步产生了 Sensorama，它增加了视觉，声音，振动，甚至是气味的不同感官体验（图 1.28 (c)）。这些设计的一个不幸的局限是要求视点相对于图片是固定的。如果设备太大，那么用户的头部也会变为固定。

自 20 世纪 50 年代以来，电影院已经有 3D 电影的趋势。当观看者佩戴特殊眼镜，使用偏振滤光片为每只眼睛选择不同的图像时，立体视觉就会实现。这种流行的 3D 电影，在今天的剧院中都以同样的方式观看。

另一种增加沉浸感和深度感的方式是增加视野范围。20 世纪 50 年代的 Cinerama 系统提供了一个弯曲的宽视场，类似于今天提供的弯曲的大型 LED（发光二极管）显示器（图 1.28（d））。沿着这些路线，我们可以放在我们周围放置屏幕。沿着这个思路一个重要的虚拟现实系统系列诞生了，称为 CAVE。它于 1992 年在伊利诺伊大学推出[49]（图 1.29（a））。用户进入一个房间，视频投影到几面墙壁上。CAVE 系统还通过使用偏振光和特殊眼镜向每只眼睛呈现不同图像来提供立体观看，通常还会执行头部追踪以允许视点相关的视频出现在墙上。

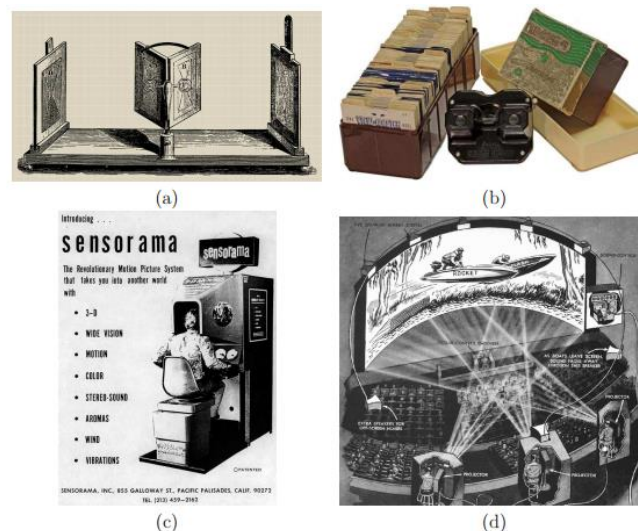


图 1.28: (a) 1838 年由 Charles Wheatstone 开发的第一台立体镜，使用镜子为每只眼睛呈现不同的图像；不久之后镜片被透镜替代。(b) View-Master 是一种大规模生产的立体镜，自 20 世纪 30 年代以来一直可用。(c) 在 1957 年，莫顿海利格的 Sensorama 增加了电影，声音，震动，甚至是气味。(d) 在与立体观看竞争中，全景电影提供了更大的视野。更大的电影屏幕导致 3D 电影的流行在 20 世纪 50 年代衰落

## VR 头戴式设备

VR 头戴式设备再一次显示了便携性的趋势。1968 年，VR 的一个重要里程碑是引入了伊凡萨瑟兰的达摩克利斯之剑，它利用了现代显示器和计算机的原理（图 1.29（b））[315,316]，构建了被广泛认为是第一个 VR 头戴式设备。当用户转动头部时，屏幕上呈现的图像会进行调整以进行补偿，以便虚拟物体看起来像在空间中固定。这让我们首先看到了一个重要概念：平稳感。为了使物体在移动时看起来是静止的，产生刺激的装置必须改变其输出来补偿运动。这就要求传感器和跟踪系统成为 VR 系统的一部分。20 世纪 80 年代，Jaron Lanier 的 VPL 公司开始出现商用 VR 头戴式设备，从而推广了护目镜和手套的交互设备，如图 1.29（c）。在 20 世纪 90 年代，基于 VR 的视频游戏出现在市场（图 1.29（d））和家庭（图 1.29（e））中，但在那时候，使用体验并不令人十分满意，没有吸引大众的兴趣。目前的 VR 技术由于智能手机行业的发展，以及高分辨率屏幕和传感器的广泛应用，已经产生了可以提供轻量级，低成本，高视野的头戴式设备，如 Oculus Rift（图 1.29（f））。这极大地提高了 VR 体验的质量，同时显著降低了开发者和业余爱好者的入门门槛。这也引起了近来对 VR 技术和应用的兴趣。



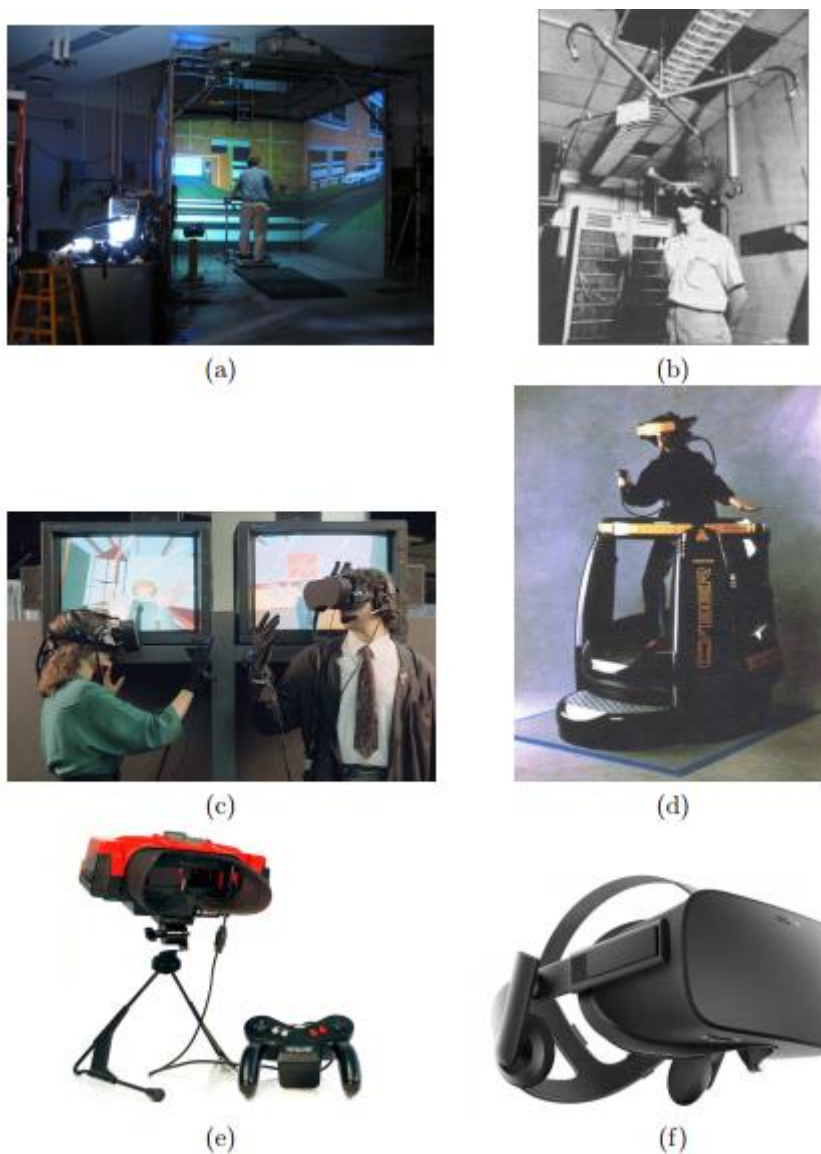


图 1.29: (a) CAVE VR, 1992 年。(b) 达摩克利斯之剑, 1968 年。(c) VPL 眼镜, 1980 年代。(d) 90 年代的虚拟游戏。(e) 任天堂虚拟男孩, 1995 年。(f) Oculus Rift, 2016 年。

### 群体聚集体验

群体聚集体验是一个容易忽视的重要方面，即人与人之间或社会之间的相互作用。我们使用诸如现场演出，教室或演讲厅等方式为人们提供交流或娱乐。我们写作和阅读小说，互相讲故事。在写作之前，熟练的讲故事者会将经验传播给其他人。几个世纪以来我们可以通过写信给对方进行沟通。而眼下更新的技术使我们能够立即进行直接交互。音频部分已通过电话传输了一个多世纪，现在视频部分也通过互联网上的视频会议传输。与此同时，简单的短信已成为我们互动的宝贵部分。自 20 世纪 70 年代以来，通过因特网进行交互的在线用户群体一直在增长。在游戏的背景下，早期的多用户地下城（MUD）已经发展成为今天的大型多人在线游戏（MMORPG）。在教育方面，伊利诺伊大学的 PLATO 系统是第一个计算机辅助教学系统，包括留言板，即时消息，屏幕共享，聊天室和表情符号。这是许多基于社区的在线学习系统的前身，如 Khan 学院和 Coursera。今天最大的在线社交互动是通过 Facebook 进行的，这些应用程序可以通过文字直接沟通，并共享图片，视频和链接。





图 1.30: Second Life 于 2003 年推出，通过化身进行社交并基本上建立一个虚拟世界的生活方式。作者在 2014 年开放社区大会上发表主题演讲。开发人员构建开源软件工具，用于构建和托管这些虚拟人物社区，以及身后的真人。

回到虚拟现实部分，我们可以在虚拟环境中创建自己的虚拟形象并“共存”，就像 Second Life 和 Opensimulator（如图 1.30）一样。一部大众科幻小说描绘了一个令人激动的反乌托邦式的未来世界，在这个世界里面每个人都喜欢通过虚拟现实进行互动。但未来到底会是怎样的趋势，这还仍有待观察。

随着目前科学技术的发展，在制作 VR 体验的时候发现其实不是描绘的越真实就越好，给用户留足想象力的空间也是很重要的环节。尽管随着技术演变的变化，人类的设计美感仍然是一样的，需要了解我们感官、大脑和身体如何工作，这对于 VR 今后的需求发展至关重要。

## 进一步阅读

本书的每一章都会对前面正文中未提及的其他相关文献进行总结。目前，VR 领域已经具有具有指导意义的书籍。消费级 VR 革新之前的几本关键教科书是 W. R. Sherman 和 A. B. Craig 编写，2002 出版的《理解虚拟现实》[288] 和 D. A. Bowman 等人编写，2005 出版的《3D 用户界面》[31]。基于当前技术的书籍包括 [135, 183]。有关增强现实概念的调查，请参阅 [348]。有关增强现实的最新报道超出了本书的范围，请参阅 [280]。

目前也已经有大量关于 VR 的研究文献。不幸的是，目前的消费级 VR 行业似乎已经忘却了 VR 研究的长期历史，导致学术与商业之间存在相当大的认识差距。当今提出的许多问题和工业中引入的方法在几十年前就得到了很好的解决。早期的大部分工作放在今天仍然有用，因此值得仔细研究。一个很好的参考读物是 2015 年出版《虚拟环境手册》[108]，其中包含了数十篇最近的调查文章和数千篇相关的参考文献。最近的资料可以在发表 VR 相关论文的网站、论坛、出版物中找到。阅读这些最新的会议、出版物可能很有用：IEEE 虚拟现实（IEEE VR），IEEE 混合和增强现实国际会议（ISMAR），ACM SIGGRAPH 会议，ACM 应用感知研讨会，ACM SIGCHI 会议，IEEE 3D 用户界面研讨会，《Journal of Vision》，《Presence: Teleoperators and Virtual Environments》。