

第十章 交互

在 VR 中，用户应该如何与虚拟世界交互呢？他们应该如何移动？他们如何抓住和放置物体？他们应该如何与对方互动？他们如何使用互联网？下文则为上述问题提供了许多见解。

通用仿真原理

VR 可以模拟来自真实世界的任何交互机制。

例如，用户可以通过转动旋钮并拉动来打开门。此外，用户还可以坐在模拟驾驶舱内操作虚拟飞机（如图 1.15 所示）。人们甚至可以戴上 VR 显示器，从而获得与梦境相媲美的体验！

尽管有通用的模拟原则，但回忆第 1.1 节，VR 的目标不一定要完全契合现实，而是追求更好的交互感。因此，本章将介绍在物理世界中没有的交互机制。

第 10.1 节介绍了一般的运动学习和控制概念。最重要的概念是重映射，其中现实世界中的运动可以被映射到虚拟世界中的另一种运动。这一概念为许多强大的交互机制的实现奠定了基础。这些机制的任务是开发易于学习，易于使用，快速执行并提供舒适用户体验的交互流程。第 10.2 节讨论了用户如何在虚拟世界中移动自己，同时保持现实世界中的固定状态。第 10.3 节介绍了用户与虚拟世界中其他对象进行交互的方式。第 10.4 节讨论了社交互动机制，它允许用户直接交流。10.5 节简要介绍了一些额外的交互机制，如文本编辑，3D 结构设计和 Web 浏览。

10.1 运动程式和重映射

动作程序

在生活中，我们会学习动作技能以完成很多特定任务，例如写文字，绑鞋带，扔球和骑自行车。这些通常被称为动作程序，通过反复的试验学习得到，随着练习量的增加，精确性和易用性逐渐提高[196]。最终，我们甚至可以无意识地产生动作。例如，大多数人可以在不注视方向盘，制动器和加速器的情况下驾驶汽车。

同样，我们大多数人都学会了如何使用计算机的接口设备，例如键盘，鼠标和游戏控制器。有些设备更容易学习。例如，学会使用鼠标不需要很长时间，但想利用键盘快速输入可能需要数年才能掌握。是什么原因让一个技能比另一个更难呢？这并不容易预测，比如由 Destin Sandlin 利用逆向转向操作设计的逆向大脑自行车，其以相反的转向控制前轮[21]。Sandlin 花了六个月的时间学习如何骑行这种自行车，最后他却无法骑一辆普通的自行车。因此，他花费了大量精力学习新的骑行方式，却忘记了如何骑一辆普通的自行车。

设计注意事项

在虚拟现实交互机制的发展中，主要考虑因素是：

1. 达到所需速度，精度和运动范围的有效性。

2. 学习新的动作程序的难度; 理想情况下, 不应期望用户花费数月时间掌握新机制。

3. 在认知负荷方面的易用性; 换句话说, 就是在一些实践之后, 交互机制应该很少或根本不需要用户集中注意力。

4. 在长时间使用期间需要得到放松; 除非是进行一些体育锻炼, 否则用户不应该出现肌肉疲劳。

为了设计和评估新的交互机制, 首先了解获得动作技能和程序的生理学和心理学。第 5 章和第 6 章介绍了视觉感知, 这是将感官输入转化为感知体验的过程。我们现在考虑在物理世界中以身体运动的形式产生输出的相应部分。在这种情况下, 大脑将运动信号发送给肌肉, 使其移动, 同时利用感知过程引入感官反馈。

运动的神经生理学

首先考虑参与学习, 控制和执行自愿运动的神经组成。如图 10.1 (a) 所示, 大脑皮层的某些部分为人体运动服务。运动皮层是控制运动神经信号的主要源, 运动前区和辅助运动区也会参与运动的准备和规划。如图 10.1 (b) 所示, 还有许多部分参与运动并通过神经信号进行通信。最有趣的部分是小脑, 意思是位于头骨后部的“小脑袋”。它是一个特殊的处理单元, 主要用于运动, 但也涉及注意力和语言等功能。广泛认为小脑损伤会影响精细运动控制和学习新的运动项目。据估计, 它包含大约 1010 亿个神经元[7], 远远超过整个大脑皮层, 其中包含大约 200 亿。虽然小脑所占体积要小得多, 但通过更小, 更密集细胞也能形成一个大的集合。除了协调动作外, 它也是动作技能的储存中心。

小脑在 VR 中最相关的一个作用是学习感知模拟器的关系, 这些关系被编码成一个动作程序。所有的身体运动都涉及某种感官反馈。最常见的例子是手眼协调; 然而, 即使闭着眼睛移动手臂, 本体也会以动作信号的参考副本的形式提供信息。因而, 在动作控制信号与感知信号和之间建立紧密的联系对许多任务至关重要。这在工程系统中也广为人知, 其中传感器反馈和运动控制结合在机器人和飞机稳定控制中有所应用; 与此相关的主题被称为控制系统。众所周知, 闭环系统是优选的, 其中传感器信息在执行期间提供反馈, 与开环相反, 开环指定运动信号作为时间的函数。

学习某个动作模式所需要的时间是最重要的因素之一。在这一点上, 不同人的差异很大。一个关键的概念是神经可塑性, 这是大脑重组神经结构并形成适应新刺激途径的能力。幼儿具有高水平的神经可塑性, 随着时间的推移, 突触修剪过程会大大降低神经可塑性 这导致健康的成年人每个神经元的突触数量比两岁的孩子多一半[99]。结果便是成年人获得新技能的难度更大, 例如学习新语言或学习使用复杂的界面。除了随着年龄减少神经可塑性之外, 它在同龄人中的差异也很大。

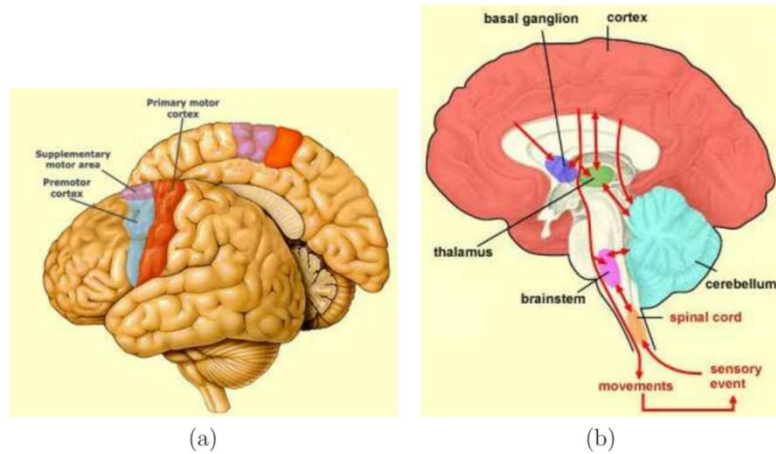


图 10.1: (a) 部分大脑皮层致力于运动。(b) 许多其他部分与皮质相互作用以产生和执行运动，包括丘脑，脊髓，基底神经节，脑干和小脑。（由 McGill University 的 The Brain from Top to Bottom 提供的数据）



图 10.2: (a) Atari 2600 桨控制器。(b) Atari Breakout 游戏，其底线部分是一个虚拟桨，允许球弹跳到顶部并消除接触后的块。

学习动作模式

现在考虑学习一个计算机界面呈现的动作程序。一个简单而经典的例子是 1976 年由 Atari 开发的视频游戏 Breakout。玩家转动一个旋钮，如图 10.2 所示。这会导致屏幕底部的线段水平移动。相应的设备包含一个电位计，通过校准可以准确地估计旋钮的方向。玩家观察位于屏幕底部的线段并快速调整旋钮方向。因此，学习过程涉及从视觉感知中获取信息，并通过转动旋钮并确定感知运动关系来获取本体信号。熟练的玩家便可以快速转动旋钮，以便更快地移动线段，而不用在现实世界中来回移动一个小托盘。因此，我们已经有了一个例子，其在虚拟世界的版本比现实中的表现更好。

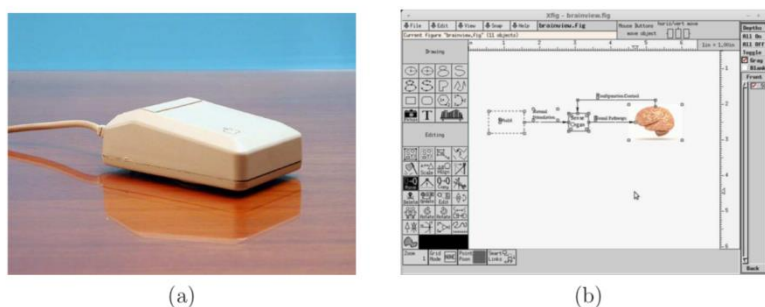


图 10.3: (a) Apple Macintosh 鼠标。(b) 当鼠标在桌面上移动时，屏幕上的虚拟手指会相应地移动，但它旋转了 90 度并需要移动较长的距离。

在 **Breakout** 示例中，旋钮方向和线段位置之间有一个一维映射学习过程。实际上，这里还有许多替代控制方案；但是，它们并不能让用户感到满意。如果你发现一个模拟器存在可以完善和突破的地方，那它很可能涉及到通过按键达到移动效果的问题。在这种情况下，按下按键的时间量对应于位移量。相应物体的速度则由程序设定。使用现代硬件的合理替代方案需要手指在触摸屏上来回移动，而相应的控制物体直接出现在触摸屏上。由于额外的自由度，手指不会受到太大的约束，而手指的快速来回运动可能会导致肌肉疲劳，尤其是在屏幕较大的情况下。在手划完成操作的情况下，运动由手指完成，手指灵活性高，而前臂移动少得多。该映射提供了身体运动和虚拟物体放置之间的关联，实现了高精度，快速放置和长期舒适性的交互。

图 10.3 显示了一个比较熟悉的例子，它是电脑鼠标。当鼠标在桌子上被推动时，编码器确定位置，该位置被转换为屏幕上的指针位置。该运动感知的映射看起来比 **Breakout** 例子要复杂一些。儿童往往能立即学会如何使用鼠标，而老年人则需要一些练习。鼠标的 2D 位置被映射到屏幕上的 2D 位置，具有两个基本的失真：1.与桌子相比，屏幕旋转了 90 度（水平到垂直运动）。2.运动被缩放，以小的物理运动产生了更大的屏幕运动。最初的施乐奥拓鼠标的优点在[39]中从人类技能学习和费茨定律[78,192]两方面进行了科学讨论，这从数学上将任务的难度指向了达到目标所需的时间。

最后假设通过按一个键，字母“h”立即以熟悉的字体出现在屏幕上。我们的视觉感知系统将“h”视为与纸面版本等效。因此，输入关键字会导致“h”的感知。这是一个相当舒适，快速和强大的操作。

VR 的动作模式

迄今为止所给出的例子已经与 VR 很相似。感知体验由通过硬件设备感知的身体移动来控制。使用通用模拟原理，可以将这些以及更多的交互机制引入到 VR 系统中。物理交互部分大体上是相同的，或者可以通过另一种控制器进行模拟。

使用第 9 章的跟踪方法，可以可靠地估计身体各部位的位置和方向并将其引入 VR。对于头部追踪的情况，保持视点的准确性和零有效延迟至关重要；否则，VR 体验质量会显著

降低。这是至关重要的，因为必须保持稳定的感觉，才能维持可信度和舒适度。感官的运动必须与跟踪系统相匹配。

重映射

对于其他身体部位的运动，这种完美匹配并不重要。我们的神经系统可以在舒适性方面产生更好的关联，就像鼠标和键盘在现实世界中的运行方式一样。因此，我们想要进行重映射，这涉及到学习一种感知运动映射，这种映射在虚拟世界中产生不同于现实世界所期望的结果。上面的键盘示例是重映射最常见的示例之一。按下一个按键取代了在纸上划动铅笔以产生一个字母的过程。

VR 的重映射是很自然的。例如，可以不用伸手抓住一个虚拟的门把手，而是按下一个按钮来打开门。对于更简单的情况，请考虑按照 HTC Vive 系统允许的方式，通过空间控制器追踪用户姿态。设置缩放参数可以使得真实世界中一厘米的手位移对应于虚拟世界中的两厘米位移。这与鼠标的缩放参数相似。第 10.2 节介绍了通过控制器实现从现实世界中的自然行走到虚拟世界中等价状态的重映射。10.3 节介绍了对象交互方法，它们通过重映射再次实现。可以预计在未来几年还会看到许多 VR 重映射方法。

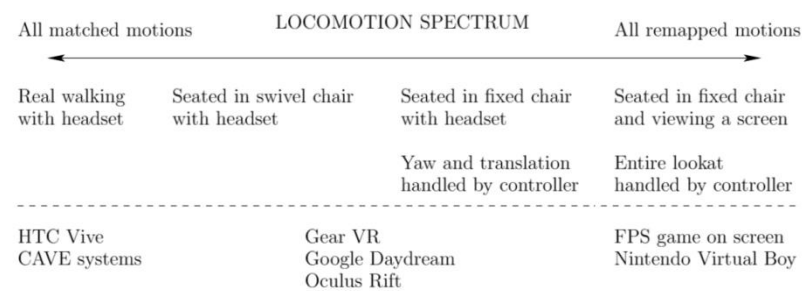


图 10.4：从左向右移动时，实际运动和虚拟运动之间的视点失配量增加。

10.2 运动

假设虚拟世界覆盖的面积远大于被追踪的真实世界的面积。换句话说，匹配区域相对于虚拟世界而言较小。在这种情况下，需要某种形式的互动机制来将用户移到虚拟世界中，同时用户仍然保持在现实世界的跟踪区域内。以这种方式移动用户的交互机制称为运动控制 (Locomotion)。就好像用户正在虚拟世界中驾驶虚拟车辆一样。

图 10.4 显示了一系列常见的运动场景。在左侧，用户在戴着 HMD 的同时四处走动。此情况下，除非虚拟世界大于开放空间，否则不需要运动控制。这种情况不涉及实际和虚拟运动之间的不匹配。

两种中间情况对应于佩戴 HMD 的坐在座位上的用户。在这些情况下，就需要使用交互机制来改变虚拟世界中匹配区域的位置。如果用户坐在旋转椅上，那么他可以通过旋转椅子来改变朝向（偏航方向）。这可以被认为是在虚拟世界中定位用户的躯干。如果用户坐在固定的椅子上，则通常使用控制器来改变虚拟躯干方向，这会导致更多的不匹配。极限情况显

示在图 10.4 的右侧，其中甚至没有头部跟踪。如果用户正面对屏幕，如在屏幕上的第一人称射击游戏的情况那样，则使用游戏控制器来改变用户在虚拟世界中的位置和方向。这是不匹配程度最大的情况，因为视点中的所有更改都是由控制器生成的。

重定向行走

如果用户在非常大的空间（例如每边至少 30 米的正方形区域）进行运动并被跟踪，重定向行走则有可能使她认为自己正在以直线行走数公里，而她实际上正在绕圈行走。这种技术被称为重定向行走[261]。在没有视觉线索的情况下长距离沿直线行走对于人类（和机器人）实际上是不可能的，因为在现实世界中不可能实现完美的对称。

想象一下虚拟城市包含长而直的街道的虚拟现实体验。当用户在街上走时，视点的偏航方向可以逐渐变化。这代表了真实和虚拟世界之间的一小部分不匹配，并且导致用户沿着圆弧行走。但这种技术的主要问题是用户有自由意志，并且可能使得用户走向现实世界中匹配区域的边缘，即使用户不能直接感受到这一点。在这种情况下，可能会出现警告，建议用户必须旋转以重置偏航方向。

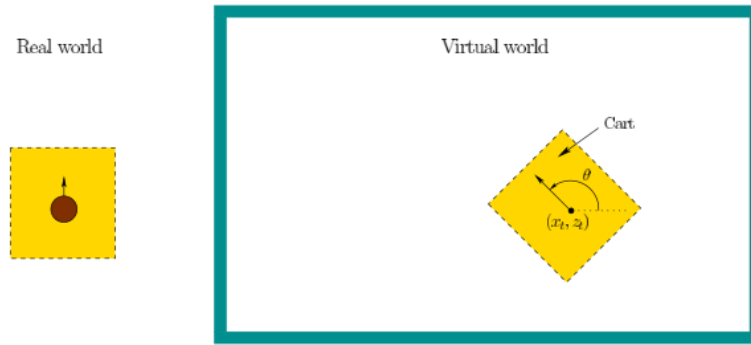


图 10.5：沿着水平地形的运动可以被模拟为在虚拟世界中的车辆。这里显示自上而下的视图。黄色区域是匹配的区域（回忆图 2.15），其跟踪用户视点。通过使用控制器来改变 x_t , z_t 和 θ 的值。

运动执行

现在考虑图 10.4 中坐着并戴着 HMD 的中间情况。运动可以通过控制器移动视点简单地实现。将匹配区域视为在虚拟环境中移动的可控车体是有帮助的；见图 10.5。首先考虑地面是水平面的简单情况。令 T_{track} 表示在物理世界中眼睛跟踪位置和方向的齐次变化矩阵。

9.3 节中描述的方法可用于为当前时间提供 T_{track} 。

车体的位置和方向由控制器决定。齐次矩阵：

$$T_{\text{cart}} = \begin{bmatrix} \cos \theta & 0 & \sin \theta & x_t \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & z_t \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10.1)$$

对位置 (x_t, z_t) 和方向 θ 进行编码（作为偏航旋转角，从 (3.18) 中借用）。 (10.1)

中的高度设置为 $y_t=0$ ，这样它就不会改变跟踪系统或其他系统所确定的高度（回顾 9.2 节，如果用户坐在真实世界中，但是在虚拟世界中是站着的，高度可以是人为设置的）。

通过连接矩阵 T_{track} 和 T_{cart} 获得眼部变化状态：

$$T_{\text{eye}} = (T_{\text{track}}T_{\text{cart}})^{-1} = T_{\text{cart}}^{-1}T_{\text{track}}^{-1} \quad (10.2)$$

回顾第 3.4 节，眼部变换是放置几何模型所需变换的逆变换。因此，(10.2) 对应于由于车体而产生的视角变化，其次是跟踪的头部视角。

为了将视点移动到固定方向 θ ， x_t 和 z_t 分量通过以下方程获得：

$$\begin{aligned} \dot{x}_t &= s \cos \theta \\ \dot{z}_t &= s \sin \theta. \end{aligned} \quad (10.3)$$

在时间步长 Δt 上积分 (10.3)，位置更新为

$$\begin{aligned} x_t[k+1] &= x_t[k] + \dot{x}_t \Delta t \\ z_t[k+1] &= z_t[k] + \dot{z}_t \Delta t. \end{aligned} \quad (10.4)$$

式子 (10.3) 中的变量 s 是前进速度。人类的平均步行速度约为每秒 1.4 米。通过按下设置 $s = 1.4$ 的按钮或按键，可以向前移动虚拟车体。另一个按钮可用于分配 $s = -1.4$ ，这会导致向后移动。如果没有按键或按钮被按下，则 $s=0$ ，使得车体保持静止。另一种控制方案是使用两个按钮来增加或减少速度，直到达到某个最大限制。在这种情况下，无需按住某一个键就可以保持运动。

除了前进/后退运动之外，按键还可用于提供侧向运动。这在视频游戏中被称为扫射。如果可能的话应该避免这种运动，因为它会导致不必要的侧向牵引。

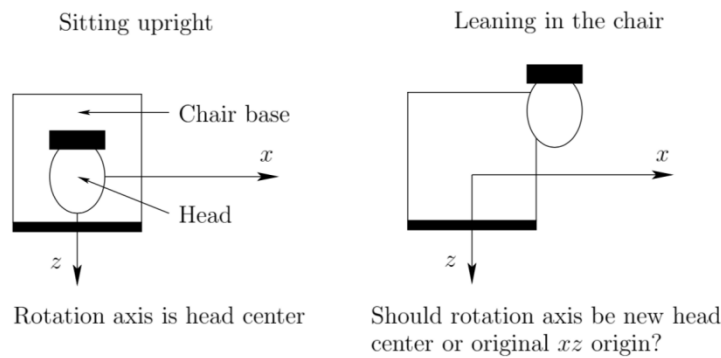


图 10.6：对于右侧直坐在椅子上的用户，偏航旋转轴线位于头部的中心。对于左侧，用户靠在椅子上。这时，旋转轴应该保持固定，还是与用户一起移动？

改变方向的问题

现在考虑方向 θ 。为了向不同的方向移动， θ 需要再分配。可以根据用户的头部偏航方

向进行分配。当用户坐在转椅上向前看时，这种操作比较方便且令人舒适。通过转动转椅，可以设定方向。（但是，这可能会引起有线耳机的问题，因为耳机线可能会缠绕用户。）

在固定的椅子上，控制 θ 可能没那么简单，因为舒适的头部偏航范围仅限于每个方向 60 度（回忆图 5.21）。在这种情况下，可以使用按钮在顺时针或逆时针方向上以小增量改变 θ 。不幸的是，根据恒定的角速度改变 θ 会导致偏航，这对很多人来说是不舒服的。一些用户喜欢点击一个按钮，每次立即偏航约 10 度。如果增量太小，则再次取向；如果增量太大，则用户会对其取向感到困惑。

另一个问题是如何定位旋转中心，如图 10.6 所示。当用户在真实世界中从椅子中心移开时会发生什么？旋转中心的位置应该是原来的头部中心还是新的头部中心？如果选择它作为原始中心，则当 θ 改变时，用户将感知到大的平移。如果产生新的中心，那么变化量就会减少，但可能与现实不相符合。

对于另一个变化情况，可以使用 8.3.2 节中的车状运动模型（8.30），以便转换视点才能旋转。换句话说，虚拟形象会有最小的转弯半径。更一般来说，可以通过控制任何虚拟车辆模型来改变视点。第 1 章的图 1.1 显示了“车辆”是一只鸟的例子。

减少相对运动错觉的策略

运动的主要问题是产生相对运动错觉并导致 VR 晕眩。在第 8.4 节，我们讲述了六种不同的运动，每个自由度都对应着一种运动。此外，影响敏感度的因素很多。减少这些因素的权重应该会减少对身体的负面影响，并希望减轻 VR 晕眩。

减少基于相对运动错觉所产生晕眩的几种策略是：

- 1.如果光流的视场减小，则视野减弱。一个常见的例子是制造阻挡大部分光流的驾驶舱或汽车内部。

- 2.如果视点太靠近地面，则运动特征的速度和加速度矢量的幅度更高。这就是为什么你可能会觉得，如果你乘坐的是一辆小型汽车，比起在卡车或小型货车上以相同的速度骑行时，要快得多。

- 3.令人惊讶的是，短时间内较大的不匹配可能优于长时间较小的不匹配；见图 10.7。

- 4.具有高空间频率将为人类视觉系统追踪提供更多特征。因此，如果通过的环境更加平滑，细节更少，则相对运动错觉将减少。考虑走楼梯的情况。如果这些步骤清晰可见，以至于它们表现为移动的水平条纹，那么用户可能会因强烈的垂直信号而迅速感到不舒服。

- 5.减少对比度，例如在加速时让世界看起来朦胧或模糊，可能会有所帮助。

- 6.提供其他感官线索，如吹风或移动音频源可能会提供更强的运动证据。包括隆隆或振动形式的前庭刺激也可能有助于降低前庭信号的信心。即使使用头部倾斜来诱发虚拟世界运动的变化也可能有所帮助，因为它会引起前庭信号分散注意力。

- 7.如果世界应该移动，而不是用户，那么通过提示或特别指示来清楚说明可以提供帮助。

8.提供具体的任务，例如对飞虫发射激光，可能会对前庭冲突提供足够的注意力。如果用户完全专注于动作，那么她可能会更快地晕眩。

9.反复练习可能会降低运动的负面影响。经常在大屏幕前玩 FPS 游戏的人似乎已经降低了对VR 中的感光度的敏感度。要求用户在晕眩减少之前进行练习对希望推出新产品的公司来说可能不是明智的策略。想象一下，尝试一些新的食物，让你在吃了前 20 次后感到恶心，但随后逐渐变得更加可以接受。谁会继续尝试呢？

最后的建议是尽可能避免运动！尝试设计不严重依赖于它的体验。

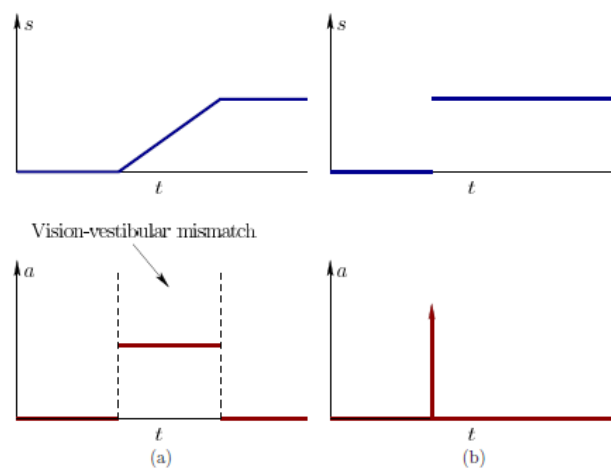


图 10.7: (a) 在一个时间间隔内应用恒定的加速度，以使停止的化身达到速度限制。上图显示了随着时间的推移速度。下图显示加速度。非零加速的时间间隔对应于与前庭感觉的不匹配。(b) 在这种情况下，施加加速脉冲，导致立即达到所需的速度限制。在这种情况下，不匹配发生在有效长度为零的时间间隔内。实际上，感知的速度在一对连续的帧中变化。令人惊讶的是，情况 (b) 比 (a) 更舒适。看起来大脑更喜欢在很短的时间间隔内出现异常值，因为在较长的时间间隔内（如 5 秒）应该会有较小的持续不匹配。

非平面运动

现在考虑更复杂的运动案例。如果用户在地形上行走，则可以简单地增加或减少 y 分量以反映高度的变化。这似乎是现实的，但请记住，它会增加真实世界与虚拟世界之间的不匹配数量，因为垂直切入与正向切入相结合。

在移动通过 3D 媒体的情况下，8.4 节的所有六种形式都可以使用。常见设置包括虚拟航天器，飞机或潜水员，他们可以轻松地生成偏航，俯仰和滚动视野。例如，想象一下飞行虚拟航天器，通过以滚动的方式运动，当周围的星星以环形模式旋转时，可能会引起滚动的相对运动错觉。如果开发人员必须以这种方式制作飞船，那么应该遵循先前关于减少切割强度的建议。此外，应该对人体进行仔细的实验，以确定哪种形式的视力在特定的应用中更差，详见第 12 章。为了避免奇点，对于所有 3 个旋转运动自由度都是可能的系统，虚拟车辆变换最好用四元数来维护（回顾在 3.3 节的问题）。

添加移动视点的特殊效果将导致更多的难度。例如，使角色上下跳动将导致垂直视角。

在走路时摇摆的头部动作也是一个影响视觉的情形，因为它导致不匹配度的增加。。设想一个更糟糕的情况，如果你是一个体操的虚拟角色，在多次翻转过程中，世界的观点可能变得无法忍受。

专门的硬件

现在已经开发了多种硬件来支持运动。最古老的例子之一是为飞机飞行模拟创建一个完整的驾驶舱（图 1.15）。图 10.8（a）显示了一个全方位跑步机，可以在任何方向和任何距离上行走。诸如固定式自行车之类的练习机已经连接到 VR 系统，以便用户可以踩踏和操纵引导自己穿越大型虚拟世界，如图 10.8（b）所示。图 1.1 显示了一个像鸟一样虚拟飞行的机械平台。



图 10.8: (a) 美国陆军用于 CAVE 系统的全方位跑步机，用于训练。(b) 由 Paul Dyan 开发的与 VR 设备相连的家庭自行车自行车系统。

瞬间移动

迄今为止涉及的运动方法主要集中于再现现实世界中熟悉的体验，这些体验提供了通用模拟原理的实例。然而，在虚拟现实，我们也可能采取物理上不合理的方式。最常见的是远距传物，就像电视连续剧“星际迷航”中的运输车一样。用户立即被运送到另一个地点。

如何确定所需的位置？一种简单的机制是虚拟激光指示器（或 3D 鼠标），通过用户持有与现实世界中的激光指示器形状相似的控制器的完成，甚至智能手机都可以使用。用户旋转控制器来移动虚拟世界中的激光点。这需要执行光线投射操作（第 7.1 节）沿着对应于激光的光线寻找最近的可见三角形。

要选择用户喜欢站立的位置，她可以简单地指向虚拟激光并按下一个键即刻传送。为了更容易指向地板，实际上可以是一个跟随重力的抛物线弧，类似于一股水流，见图 10.9。不可见的地方可以使用弹出式地图进行选择，甚至可以执行基于文本的搜索（可以使用语音命令代替打字）。一种叫做微缩世界的方法可以向用户展示一个虚拟的小型环境版本[312]，这实际上是一张 3D 地图。

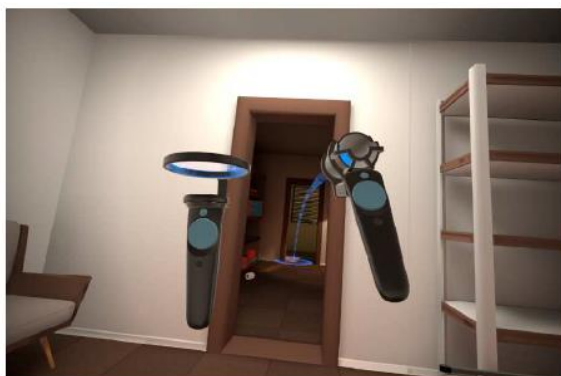


图 10.9: 一个虚拟的“激光指示器”，它遵循抛物线弧线，因此可以很容易地将传送目的地指定为地板上的一个点。（图片来自 HTC Vive 平台上的 Budget Cuts 游戏。）

路径寻找

学习空间表示并将其用于导航的认知问题称为路径寻找。这是一个比低层运动机制更高层次的过程，但这两者密切相关。运动系统在现实世界中并不熟悉的一个问题是，用户可能无法了解周围世界的空间布局。如果你能够在不同的地方传送，你的大脑是否仍然会形成真实世界环境的地方记忆细胞？我们广泛观察到这种现象，发现人们大多仅学习使用 GPS 或出租车服务导航城市道路，而不是自己寻找路线。

远距传送机制减少了运动，因此减少了 VR 晕眩，然而，这可能是以减少对环境空间布局的学习为代价的。进行远距传物时，不要改变视点的偏航方向；否则，用户可能变得更加迷失方向。他可能不了解他现在在虚拟世界中相对于以前的位置定位和定位的位置。

请注意，通用仿真原理可以再次用于从现实世界借用任何有效的导航工具。如果虚拟建筑和城市以现实世界中常见的方式布置，那么他们应该更容易导航。标志和地标甚至可以放置在虚拟世界中以帮助导航。在现实世界中，标志通常会告诉我们出口的位置，街道的名称或区域的边界。诸如高楼，风车或塔楼等地标提供了可以长距离导航的视觉线索。许多这些想法在[31]的第 7 章中讨论。

10.3 操作

我们与真实世界中的物体进行交互的原因很多。你可以通过在碗和嘴之间移动勺子来吃一碗汤，你可能会拿起一块石头并尽可能地扔掉，你可以穿上一条裤子……这些例子属于操纵的主题。在现实世界中，操纵涉及复杂的感觉运动关系，通过进化和经验，我们可以在各种设置下操纵对象。物体的变化包括尺寸，重量，摩擦，柔韧性，温度，易碎性等方面的差异，但不知何故，我们的身体都可以处理。让机器人以人类的方式进行操纵一直是一条漫长而令人沮丧的道路，只有有限的成功[203]。

由于现实世界中的操作复杂性，它是应用 10.1 节中的重映射概念在 VR 中尽可能简化操作的理想人选。虚拟世界不必遵循复杂的操纵物理学。相反，优选的是进行诸如选择，抓取，操纵，携带和放置物体的操作尽可能快和简单。此外，应避免广泛到达或其他形式的肌

肉紧张，除非 VR 体验旨在提供锻炼。

避免“大猩猩手臂”

在公众中最常见的误解之一是汤姆克鲁斯在电影“少数派报告”中使用的接口是可取的；见图 10.10。事实上，它很快导致了“大猩猩手臂”这个众所周知的问题，使用者很快感觉到手臂长时间疲劳。你可以在自己面前直接握住手臂而不会疲劳吗？



图 10.10：汤姆克鲁斯在 2002 年电影“少数派报告”中在全息显示中移动窗口。这对好莱坞来说是一个很好看的互动机制，但在现实中令人非常疲劳。用户会很快体验大猩猩的武器。

选择

在虚拟世界中选择一个物体的最简单方法之一就是使用虚拟激光指示器，该指示器在 10.2 节中有描述。多种变化可能有助于改进选择过程。例如，用户可能会拿着照亮潜在选择的虚拟手电筒。手电筒的视野可以调节[83]。可以放置虚拟镜像，以便可以在角落进行选择。[31]的第 5 章提供了许多其他建议。

通过使用指针，用户只需照亮感兴趣的对象并按下一个按钮。如果目标是检索对象，那么它可以立即放置在用户的虚拟手中或库存中。如果目标是以标准的、重复的方式操作对象，那么按下按钮可能会导致执行虚拟电机程序。例如，这可以用来转动门把手，从而打开一扇门。在这样的使用中，开发者可能希望设置激光指示器深度的限制，以使用户必须站得足够近才能进行交互。例如，把门把手从房间转过来似乎是不合适的！

如果物体很难看清楚，那么选择过程可能会很复杂。它可能在用户头后，这可能需要不舒服的转弯。该物体可能很小或很远，以至于它只能在屏幕上占据几个像素，因此难以精确选择它。如果感兴趣的物体周围存在大量杂物，特别是如果其他可选物体在附近，问题会变得更加严重。最后，该对象可能部分或完全被遮挡。

操纵

如果用户长时间携带物体，则不需要挤压或离合控制器，这会产生不必要的疲劳。在某些情况下，用户可能需要仔细检查物品，同时拥有它。例如，他可能想要将其移动到手中以确定其 3D 结构。可以将对象方向设置为完全遵循用户所持有的控制器的 3D 方向。用户甚至可以握住由外部相机跟踪的真实物体，但在虚拟世界中具有不同的外观。这使得熟悉的力量反馈给用户，这个概念在 13.1 节中重新讨论。请注意，一个对象甚至可以直接在虚拟世

界的原始位置进行操作，而不会将其靠近用户的虚拟身体[30]。在这种情况下，虚拟手被带到物体上，而物理手保持在原位。还可以模拟比正常情况更长的手臂[254]，以便更长距离地检索和放置物体。

放置

现在考虑解开物体并将其放入世界。对于用户来说，简单的情况是按下一个按钮，让对象简单地落在正确的位置。这是通过吸引域来实现的，该吸引域是在目标姿态(位置 and 方向)的邻域中定义的吸引力潜在函数，见图 10.11。潜在函数的最小值在目标上。物体释放后，物体通过移动落入目标姿态，以使电位降至最小。在许多 2D 绘图程序中可以看到这种行为，以便线段的端点方便地相遇。一个方便的对象放置的例子是马库斯佩尔松（Notch）在 2011 年的 Minecraft 沙盒游戏中，其中的积木很容易落到位。孩子们用这种方式建立了数以百万计的虚拟世界。

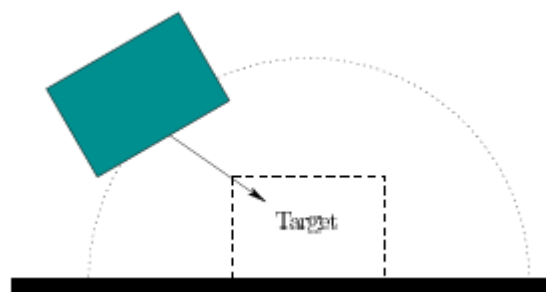


图 10.11：为了让使用者的生活更轻松，可以在一个物体周围定义一个吸引域，以便当物体进入时，落下的物体被直接吸引到目标位置。

或者，用户可能需要微妙地放置物体。也许应用程序涉及尽可能高的堆叠和平衡对象。在这种情况下，精度要求会非常高，给控制器跟踪系统和用户带来负担。

重映射

现在考虑重映射的功能，如 10.1 节所述。最简单的情况是使用按钮来选择，抓取和放置对象。除了按钮之外，用户可以生成连续的运动并由系统进行跟踪。示例包括转动旋钮，移动滑块，在手指上移动手指以及通过空间移动自由浮动身体。回想一下，重映射最重要的方面之一是易学易用性。减少重新映射的自由度数通常会简化学习过程。为了避免“大猩猩的手臂”和相关问题，可以在跟踪设备上施加一个比例因子，以便控制器中的少量位置变化对应于虚拟世界中的大运动。这个问题可以再次使用 Fitts 定律来研究，就像电脑鼠标一样。请注意，这可能会对虚拟世界的精度产生不利影响。在某些设置中，定向缩放也可能是可取的。在这种情况下，三维角速度 (ω_x , ω_y , ω_z) 可以按比例缩放，以在虚拟世界诱发更多的旋转，而不是真实世界。

目前系统设计

操纵互动机制的发展仍然是 VR 面临的最大的挑战之一。当前一代消费类虚拟现实头戴式

设备或者利用现有的游戏控制器，比如 2016 年将 Xbox 360 控制器与 Oculus Rift 捆绑在一起，或者引入假设大型手部动作的系统，如 HTC Vive 控制器，如图 10.12。使用户在空间中移动手的控制器与图 10.10 中显示的少数报告交互机制似乎并不相同。其他人正在开发不涉及硬件的手势系统，就像 9.4 节中图 9.25 所示的 Leap Motion 系统一样。这些可能是 20 世纪 90 年代流行的“护目镜和手套”愿景的更新版本（回顾 1.3 节图 1.29 (c)）。在未来几年，操作方法和技术的快速发展可以预期，并且越来越重视用户的舒适度和易用性。



图 10.12: (a) 2016 年随 HTC Vive 一起推出的一对手持控制器；该设备包括侧按钮，触发器和用于拇指的触摸板。(b) 尝试控制器的用户（原型版本）。

10.4 社交互动

沟通和社交互动是很大的主题，远远超出了本书的范围。此外，虚拟现实或社交虚拟现实中的社交互动仍处于起步阶段，并在进行实质性的实验和对范式的反思。尽管如此，将人类连接在一起是 VR 技术最大的潜力之一。尽管在我们周围展示我们自己与世界之间的显示可能是孤立的，但我们也可以通过成功的交互机制将它们拉近在一起。本节重点介绍了几个有关社交互动的有趣问题，而不是提供完整的评论。

超越 Shannon-Weaver 的交流

一个重要的因素是有多少人会通过媒体进行互动。从两个人开始说起,有史以来最强大的数学模型之一是 Shannon-Weaver 通信模型，数十年来一直是工程通信系统设计的基础;见图 10.13。该模型涉及发件人和收件人。通信系统对来自发送者的消息进行编码，然后通过噪声信道发送。在另一端，系统解码该消息并且它到达收件人。收件人可以提供反馈，以指示邮件是否已完整收到。这种通信模式引发了信息理论领域，它为通信信道带来了明确的带宽概念，并揭示了数据压缩的局限性。

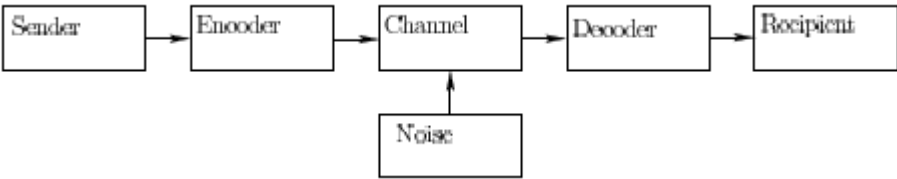


图 10.13: 经典的 Shannon-Weaver 传播模型（从 1948 年开始）。发送者向编码器提供消息，编码器通过受噪声干扰的信道发送消息。另一方面，解码器将该消息转换成适合接收机的格式。这个模型作为工程领

域的交流理论的基础。

这个模型的强大之处在于它能够在数学上量化人类的相互作用，但它也不足以涵盖 VR 中可能存在的各种相互作用。通过再次遵循通用模拟原理，现实世界中存在的任何类型的人际交互都可以引入 VR。Shannon-Weaver 模型受到诸如 19 世纪电报或 20 世纪手持式接收器（或对讲机）之类的交互机制的启发。在这些情况下，人类彼此完全隔离，并且该技术提供了类似于写信的突发信息。我们已经从文本到音频再到视频通信，并且可以通过将触摸和嗅觉等其他感官的显示结合进来，进一步延伸；同时，也有很多机会使用合成模型，可能与摄像头和麦克风捕获的实际信息相结合。简单的手势和行为举措可以提供交互的微妙但重要的组成部分，而这些组成部分并未被经典通信模式所捕捉。

尽管存在虚拟现实的缺点，但请记住，Shannon-Weaver 模型能够为计算机网络和系统提供强大的带宽和延迟分析，最终支持任何形式的社交互动。因此，它对 VR 系统中可以或不可以实现的内容具有深远的影响。发生这种情况是因为所有“通信”都会转换为通过电缆或网络连接发送的位流。一个关键的问题是确保有针对性的社交互动 VR 体验在计算机网络上舒适，可靠地得到支持。

从头像到视觉捕捉

其他人如何在 VR 中看到你？这是最有趣的问题之一，因为它取决于社会背景和技术限制。这存在一个明确的可能性范围。在某种极端情况下，用户可以通过虚拟形象代表自己，虚拟形象是一种 3D 表现形式，可能完全不符合用户可看到、可听到的形象以及行为特征；如见图 10.14。在另一个极端情况下，用户可能会使用成像技术进行拍摄，并在虚拟世界中以高度精确的 3D 表示进行复制；见图 10.15。在这种情况下，虚拟世界的人物形象像是直接从真实世界传送过来一样。这样的想法还存在许多其他可能性，值得考虑和权衡。



图 10.14: Second Life 提供的初始人物造型集合。



图 10.15: 2016 年来自微软的全息通信研究。通过增强现实眼镜（Hololens）可以看到，人的 3D 表示被实时提取并叠加在世界中。

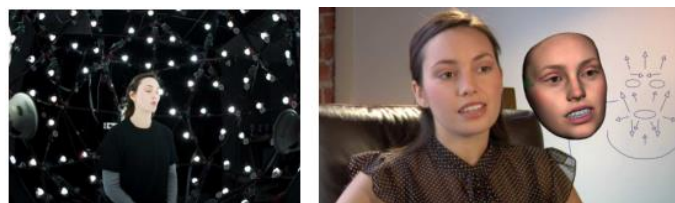


图 10.16: 2009 年的数字艾米莉项目: (a) 真人成像。 (b) 几何模型与复杂的渲染技术融合以制作动画，产生逼真的面部运动。

VR 中的虚拟形象的一个主要吸引点在于匿名性，它提供了在社交环境中扮演不同角色或展现不同个性特征的机会。在一种称为变形人效应的现象中，人们已经观察到，一个人的行为会根据虚拟人物的虚拟特征而改变，这与人们在穿着不同制服或服装时会表现出不同的形象类似[360]。用户可能想要生活在幻想中，或尝试从不同的角度看待世界。例如，如果人们能够在一个从种族，性别，身高，体重，年龄等方面都不同的世界进行体验，便可能会产生同理心。

用户也可能想要尝试其他形式的体验。例如，一群孩子可能想要居住在动物的尸体中，并进行正常的生活。畅想一下，如果人们把你看作是外星人、昆虫、汽车，甚至是奶酪。1986 年，当皮克斯在动画短片 *Luxo Jr* 中拥有台式电脑的生活时，人们感到非常惊讶。好莱坞电影在过去的几十年里充满了动画角色，我们有机会体验其中的一些人，同时居住在虚拟世界！现在考虑这种方式的物理可行性。基于目前的技术，可以独立考虑三种主要的相似性：

1. **视觉外观:** 在可见特征方面，虚拟形象与真实人物看起来有多接近？
2. **听觉外观:** 来自虚拟形象的声音与人的声音，语言和言语模式相匹配多少？
3. **行为表象:** 虚拟形象的动作与身体语言，步态，面部表情和其他动作的匹配程度如何？

第一种相似性可以通过在虚拟世界中创建一个运动学模型（回忆第 9.4 节）来进行与人的匹配，该运动学模型的大小和移动性与实际的人相对应。此外也可以执行其他比较简单的匹配，例如头发颜色，肤色和眼睛颜色。为了进一步提高真实感，可以使用纹理映射将皮肤和衣服映射到虚拟形象上。例如，用户脸部的图像可以被纹理映射到虚拟形象脸部。也可以

通过构建合成模型来进行高精度匹配。迄今为止进行的一些最佳合成匹配是由 USC 创新技术研究所的研究人员完成的；见图 10.16。如第 1.1 节所述，一个令人沮丧的问题是恐怖谷理论。人们常常会描述计算机生成的动画，比如看僵尸或谈论尸体等，这些动画往往趋近于人类自身的现实。因此，不完美的匹配通常比视觉外观上的“几乎相同”的匹配要好得多。对于听觉部分，Second Life 和类似系统的用户喜欢发短信。这种互动在 VR 中就好像是他们正在讨论。发短信有助于确保匿名。在 VR 中记录和再现语音很简单，使听觉外观与视觉外观相匹配更加简单。我们必须注意以适当的区域化来呈现音频，以使其他人看上去是从虚拟人物的嘴巴里说出来的；具体细节请参阅第 11 章。目前，已可以通过使用实时语音更改软件（如 MorphVOX 或 Voxal Voice Changer）进行录音，轻松保留匿名消息，这可能比在某些设置中发短信的方式更受欢迎。

最后要注意的是，人的行为体验可以在 VR 中完美匹配，而虚拟人物会具有完全不同的视觉外观。这是动作捕捉系统产生的主要原因，其记录真实演员的动作，然后用动画制作虚拟人物。请注意，电影制作通常是一个漫长的离线过程。在低成本的虚拟现实系统中，目前还无法实现与人的视觉和行为外观实时准确匹配的功能。此外，尽管最近在研究在这方面取得了一些进展[180]，但如果人脸的一部分被头戴式耳机覆盖，其捕获将会很困难。

不过，当前的跟踪系统可以在某些情况下提供精确匹配的行为外观。例如，头部追踪可以直接与头像相关联，以便其他人可以知道其头部转向的位置。用户还可以理解头部运动或手势的意义，如“是”或“否”。图 10.17 显示了一个简单的 VR 体验，用户可以一边看电影，一边由被跟踪的虚拟头像（他们也可以相互交谈）代替。在一些系统中，可以使用眼动追踪，以使用户可以看到虚拟人物正在看的位置；然而，在某些情况下，这可能会再次陷入恐怖谷理论。如果跟踪手部运动，可以使用如图 10.12 所示的控制器完成，这种情况仍然可以使用户沉浸于虚拟世界。



图 10.17: Oculus Social Alpha，这是三星 Gear VR 的应用程序。多个用户可以在虚拟世界中见面并进行社交。在这种情况下，他们一起在剧院看电影。他们的头部动作是使用头部追踪数据提供的。他们也可以用本地化的音频互相交谈。

从一对一到整个社会

现在考虑不同规模的社交互动。我们在现实世界中进行的绝大多数一对一互动是与我们认识的人进行的。同样，通过科技进行交互时也是如此，无论是通过短信，电话还是视频聊天。我们通过科技进行的大多数互动都是针对性的，因为参与有特定的目的。这表明，虚拟现实技术可以用于将视频聊天提升到一个新的水平，在这个阶段，两个人会感觉他们在虚拟世界中是面对面的。但要注意的是，在现实世界中，我们可能会有一些无意的互动，而不是只进行有针对性的社交。

一对一交流的一个重要方面在于两个人之间的关系是对称的还是互补的（来自 Paul Watzlawick 的交际公理）。在一个对称的关系中，两个人的地位是平等的，而在一种互补的关系中，一个人处于更高的地位，就像老板和员工或父母和孩子一样。这极大地影响了互动的风格，特别是在有针对性的活动中。

现在考虑现实世界中的一小群人之间的交互。可能是一个家庭，一帮同事共同吃饭。也可能是孩子们一起在操场上。也可能是朋友和家人聚集在一起度假或庆祝生日。这种互动类型的 VR 版本可以专注于有针对性的活动。也许你是一个不能亲自出席的人，而是通过一些 VR 界面与团队“出去”。也许有一次会议，一些人需要远程参加，这一点在目前通过被称为“电话会议”的方式进行处理，由网络传输语音和视频。但最接近虚拟现实的常见情景是人们在网络视频游戏中发生的一些社交互动。他们可能会组队并通过短信或语音进行互动。随着人数增加到十几个，互补关系的情况导致人们需要进行自我介绍或采访。这有一些例子如老师向一个班的学生讲课，政治家在一群记者面前讲话等。在这些场景中，领导者已被明确指派与团队沟通。这种设定在 VR 中可通过拍摄全景视频来重现。或者，整个活动直接在虚拟世界中进行。在对称关系的情况下，人们可能会在大型聚会中并以小组形式进行对话。这也可以在 VR 中复制。

改善的社交互动

本书中两个共同的主题是 VR 可以产生比现实更好的体验，并且我们的感知系统会适应新的刺激。因此我们会很自然地想知道如何通过 VR 改变或改善社交互动。转型社会交往的概念已经被引入 [16]。一个发人深思的例子如图 10.18 所示。在一个虚拟的世界里，老师可以在讲课时，注视着每个学生。这在现实世界中是不可能的，但在 VR 中很容易制作，因为每个学生都可以看到不同版本的虚拟世界。当然，学生可能会认为老师不可能观察所有人，但是她正在关注自己的可能行会对学习成果产生重大影响。虚拟世界的课堂上看起来可能只有少数学生，而实际上会有成千上万的学生参加。所以还有多少社交互动机制可以在现实世界中实现？我们的大脑能够多快适应它们？在什么样的环境中会更喜欢这种互动，以便在现实世界中相遇？这些问题可以在未来的虚拟世界中得以解决，同时其应会带来更多激动人心的社交新机制。

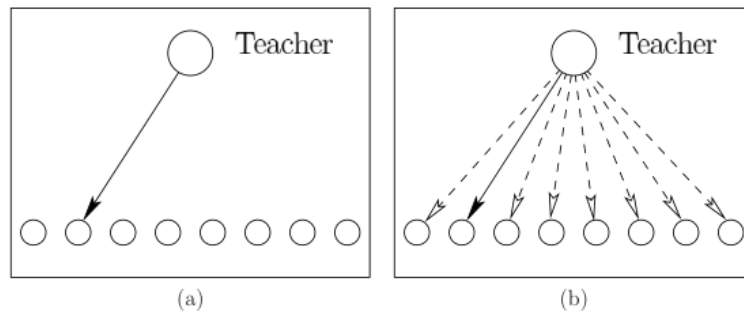


图 10.18: (a) 显示了一个普通教室的自上而下的描述, 教师可以直接看一个学生。(b) 在一个 VR 教室里, 至少从每个学生的角度来看, 老师可以同时看每个学生。

10.5 其他交互机制

本章涵盖了三个互动机制的呈现方式: 运动, 操纵和社交。这些方式从数十年的研究和开发中涌现出来, 但并未完全涵盖各种相互作用。许多系统要求构建不属于三种方式的自定义交互机制。此外, 随着目前广泛使用低成本的 VR 系统, 我们预计新的互动机制将会出现。这里介绍几个其他交互机制和相关的挑战。

与信息 and 媒体的互动

在遵循通用模拟原则的条件下, 互联网内容可以以多种方式进入 VR。第 1.2 节中的图 1.7 显示了虚拟影院中的电影屏幕。在这种情况下, 可能需要简单的交互来暂停或更改电影。作为一个更复杂的例子, 网络浏览器可以出现在虚拟世界中的公共显示器上或现实世界中用户熟悉的任何其他设备上。或者, 虚拟屏幕可以直接浮在用户面前, 同时提供稳定, 熟悉的背景: 见图 10.19。

数十年来, 人们通常使用两个输入设备与计算机和浏览器交互, 一个用于打字, 另一个用于其他的指向。在个人电脑的情况下, 这已经具体为了键盘和鼠标的形式。使用现代智能手机时, 人们则需要在小型触摸屏上键入内容, 或者使用语音或轻扫式等替代方法。他们通过手指触摸进行指向, 而另外一只手指进行屏幕缩放等操作。



图 10.19: 在 HTC Vive 中查看 Valve Valve 游戏应用商店。

文本输入和编辑

智能手机上的打字选项足以让用户输入检索词或简短信息, 但对于撰写小说而言, 这些

选项还不足够。对于目前坐在电脑面前撰写报告、计算机程序、报纸文章的专业人员而言，需要什么类型的界面才能吸引他们在 VR 中工作？

一种选择是跟踪真实的键盘和鼠标。跟踪指尖可能需要提供视觉反馈。这就需要开发相应的系统，将桌面和周围环境神奇地转化为任何东西。就像在桌面系统上使用背景图像一样，令人轻松的全景图像或视频可以在工作时包围用户。对于实际工作部分而言，VR 应在在周围和不同深度出现多个屏幕或窗口，而不是仅在用户面前放置一个屏幕。

这很容易借用现有桌面窗口系统的界面概念进行类比，但仍有许多研究在设计和评估全新的界面，以提高写作时的生产率和舒适度。虚拟现实中的文字处理过程如何？用于编写和调试软件的集成开发环境（IDE）是什么样的？如果键盘和鼠标被其他接口取代，那么用户可能甚至不需要坐在办公桌上工作。这其中的一个挑战是让用户学习一种提供与使用键盘相媲美的文本输入速度的方法，但同时使他们能够更舒适地工作。

3D 设计和可视化

能够居住在 3D 虚拟世界的优势是什么？除了视频游戏外，其他几个领域也促进了计算机图形学的发展。在进行计算机辅助设计（CAD）之前，建筑师和工程师花费了很多时间用铅笔和纸张辛苦地在纸上画出精确的线条。计算机已被证明是设计不可或缺的工具。数据可视化在过去几年中一直是计算机领域的关键应用。例如医疗，科学和市场数据。有了这些用途，我们被迫通过在屏幕上操纵 2D 投影来查看设计和数据集。

VR 则可以提供设计或数据集 3D 版本交互和查看的能力。而且被观察事物可以从外向内看，也可以从内部看。如果仔细研究第 6 章中的感知概念，那么设计的对象或环境与真实环境之间的差异可能比以往任何时候都小。在构建物理原型之前，查看 VR 中的设计可以被看作是一种虚拟原型。这可以在产品开发周期中实现快速，低成本的突破。

交互机制是实现基于 VR 的设计和可视化的根本挑战。什么能让建筑师，艺术家，游戏开发人员，电影制作人员或工程师长时间舒适地构建 3D 世界？什么工具可以让人们在 3D 投影世界中准确地操纵高维数据集？

未来

即使仅应用通用的模拟原理，也可以设想更多的交互形式。视频游戏已经提供了许多通过标准游戏控制器进行交互的想法。除此之外，任天堂 Wii 遥控器在制作诸如保龄球或网球等运动的虚拟版本时尤其有效。什么样的新互动机制对 VR 来说是舒适和有效的？如果显示器可以呈现出视觉以外的感觉，则会出现更多的可能性。例如，如果一个人穿着一套能够对身体施加适当力量的西服，你能否在虚拟世界中感受到另一边某人的一个拥抱呢？

进一步阅读

有关人体运动控制和学习的概述，请参阅书籍[196,271]。[61]中介绍了 VR 环境中的本体感受问题。有关运动的更多信息，请参见[31]的[52]和第 6 章和第 7 章。为了掌握更多关

于机器人的问题，请参见[203]。

以下创新性的论文研究了手眼协调的局限性：[56,70,346]。[229]中引入了幂法则，这表明反应时间的对数会随着练习量线性减少。将 Fitts 定律与指向设备操作联系起来的研究包括[72,193,194,302]。有关人机交互的广泛报道，请参阅[37,40]。有关通过头像进行社交互动的其他参考资料，请参[20,210,328]。