

07 | 操控数字世界：用手与身体取代键盘

2022-05-11 方军

《说透元宇宙》

课程介绍 >



讲述：山荣

时长 18:30 大小 16.94M



你好，我是方军。

在前面两讲，我们探讨了构建立体互联网的前两个步骤“扫描实体世界”和“渲染数字世界”，这一讲，我们来聊聊实现立体互联网至关重要的第三步——操控数字世界，也就是，让我们与数字世界真正互动起来。

人机交互：从虚拟现实手套说起

你可能已经发现了，虽然有不少人喜欢把“元宇宙”直接等同于 VR，但我们这个专栏已经到了第七讲，却从来没有直接讨论过 VR。这主要是因为，元宇宙不只是让你戴上 VR 头盔进入虚拟世界。当然，VR 的确是跟元宇宙有些关联的，尤其是在人与虚拟世界交互这个部分。在我看来，狭义的虚拟现实包括两个方面。

第一，利用人类视觉的特点，在两个眼睛前呈现略有不同的画面，欺骗双眼，让我们的大脑以为画面是立体的。

第二，在“立体画面”的基础上，让我们可以走入数字世界，用手甚至整个身体与数字世界互动。

很多人认为第一点也就是欺骗双眼很重要，但其实后者更重要。在法国多位学者共同编写的《虚拟现实与增强现实：神话与现实》一书中，他们强调了虚拟现实的目标。

我们首先要提醒自己的是，VR 的目标是让用户在执行虚拟任务的同时，相信自己是在现实世界中执行任务。

你看，他们说的关键词是“执行虚拟任务”，通俗一点说就是行动，与数字世界互动。

很多人在讨论虚拟现实时把视觉作为重点，关注各种头戴式显示器、头盔、眼镜，但其实让我们可以用某种方式在虚拟世界里行动才是关键。我的这个观点受到了“虚拟现实之父”杰伦·拉尼尔的启发，他创办了最早的虚拟现实公司 VPL。看他的自传《虚拟现实：万象的新开端》时，我发现，相比很多人以为重要的视觉，杰伦·拉尼尔更重视“触觉”。

在我看来，重视触觉的理由可以说得很简单，也可以说得很高深。简单地说，在 20 世纪 80 年代，由于计算机根本无法创造出逼真的三维画面，因此当时的他不得不选择触觉，或者更准确地说是手的操控动作作为关注的重点。

要说得更高深一点的话，我可以直接引用杰伦·拉尼尔的话，他特别善于概括技术的意义，他说：

输入比（你感觉到的）现实更重要，你在 VR 中输入的就是你自己。

如果你不能伸手触摸虚拟世界并影响它，你在其中就是一个二等公民。

对元宇宙来说，我们用计算机创造的、让我们看到的数字世界并不是最重要的，最重要的是你在这个数字世界中的行动。Decentraland 是一个总被提起的元宇宙，但我和不少朋友一直在批评它，因为它搞了很多的建筑和会展，但没法让我们这些用户在其中方便地活动。这带来的效果就是拉尼尔说的，Decentraland 看起来再惟妙惟肖，但我们身在其中却觉得不自在，因为你没法影响它，你感觉自己是个“二等公民”。

在虚拟现实领域，人们一直在探索操控数字世界的方法。如果你跟我一起去看虚拟现实技术的早期发展历史，你会看到各种被称为数据手套的设备。这些数据手套上装了很多传感器，戴上它你就可以操作计算机中的三维物体。

VPL 公司有一个早期试验产品名叫“抓住（GRASP）”，对，就是抓住一个东西的“抓住”。在屏幕前，只要你戴上数据手套移动双手，屏幕里的手就会跟着移动。你的手，实际上就是屏幕里的手，可以抓住电脑屏幕里的形状和模型，这就是这个产品叫“抓住”的原因。

有了“抓住”这个与虚拟世界交互的原型产品之后，人们开始畅想各种各样的可能性：你可以抓住一个棒球，把球扔出去；可以拿起木槌，弹奏虚拟的木琴；可以抓住虚拟的岩壁，玩起攀岩。

VPL 公司在市场上比较受欢迎的虚拟现实产品主要由两大部分组成。第一部分是头戴式的显示设备，我们可以用它模拟三维立体的视觉效果，但它的实用功能几乎都是靠第二个部分来实现的，也就是数据手套、手柄等各种操控设备。比如，他们为福特等汽车公司制造了虚拟现实版的汽车驾驶模拟器。用户可以头戴显示头盔，手上戴着数据手套，用手套操作虚拟的方向盘。

又比如，他们生产了一款滑雪模拟器。用户头戴显示头盔，脚踩在带有传感器的滑雪板上，两手撑着滑雪杆，这样，用户在室内就可以拥有在雪道上滑雪的感觉。

刚刚这么长的一段讨论，要传递的信息是简单的：**元宇宙不只是一个看起来惟妙惟肖的数字世界，同样重要的是我们要能操控这个数字世界。**

但我们有哪些方式可以操控数字世界呢？接下来，我们还是以一系列消费级的计算机产品为例，来看看我们与数字世界的交互方式的进化。我喜欢说，能看到多远的过去，就能看到多远的未来，因此这里在谈交互时，我们也来一起看看历史。

在相当长的时间里，操控计算机的方式是键盘，在图形化界面兴起后，鼠标成为辅助的设备，等到智能手机普及之后，我们用手指的触摸取代了鼠标。人们也曾试图把笔引入进来，但到目前为止，笔触也只在很小的范围内使用，如数码插画师使用的绘画板与笔，又比如我们配合 iPad 使用的 Apple Pencil。

早期任天堂也曾经制作过游戏手套，但现在在游戏操控方面，游戏玩家还是更喜欢使用游戏手柄。数据手套这样的产品只是局限在科研、工业模拟等极少数场景。另外，现在在技术上深度

相机已经可以很好地捕捉我们的身体动作，比如把它放在微软 **Kinect** 的体感设备上，方便我们用身体去操控游戏。但 **Kinect** 的停产表明，这一类互动技术想要普及还有很长的路要走。

微软的 **Hololens** 可能是目前消费级产品中最顶级的“广义 VR 设备”，当然，微软自称他们做的是略有不同的“混合现实”（**MR**），这两个概念都不难理解而且有很多交叉点，我们就不强行区分了。我们可以透过 **Hololens** 眼镜同时看到实体世界和虚拟世界，它的传感器能够准确地捕捉到我们的手，然后让我们用手去操控数字世界。

Hololens 很酷，但如果你实际用用它，也会发现它的互动操作部分离理想状态还很远。当你要输入文字时，你的眼前会出现一个像实体键盘一样的虚拟键盘，如果你只是打一个字母，按键会轻轻向下然后弹起，同时伴随着按键音，感觉还挺真实。看起来不错嘛。但是，如果你试着连续打字，按错键也是常有的事，这是我们使用实体键盘时根本都不会遇到的障碍。

在展望人机交互的未来时，还有人会提到“脑机接口”，也就是用人的脑电波活动直接操控计算机，不过，这样的技术离普通人的生活更加遥远，我们先不讨论了。

看完我们和数字世界互动方式的变化过程，我来做个小小的分类和总结。我们跟数字世界的互动有三个层次。

- 第一种是由计算机创造的，可以“被动观看”的数字世界，也就是说，它没有交互。这样被动观看的世界我们早就在电影、电视里看到很多了。
- 第二种是可以互动观看的三维数字世界，我们可以走入其中，从不同的角度观看它。这种互动看起来即将成为现实，比如在 **Decentraland** 这样的网页版虚拟世界中，我们能够在虚拟的建筑群中走动，它给你营造一种沉浸其中的感觉。
- 第三种是我们可以自然地与之互动的三维数字世界。我们想自然地生活在数字世界中，方便地操控其中的物品。

我们当然希望元宇宙来了之后，能够马上彻底改变我们与数字世界的交互方式，我们希望自己不再需要键盘、鼠标、触摸屏。但是，这条路还有很远要走，目前几乎没有可行的思路。

说到这里，你可能会想，操控数字世界看起来没什么好方法了，那就这样吧，跳过跳过，我们进入下一个话题吧。

且慢，其实，在过去一年多元宇宙应用的大发展中，我们发现了一个新的方向。这一节课的主题是讨论我们与数字世界互动，而从最新的变化中我们看到，**元宇宙交互的未来或许应该从“人机交互”转向“人人交互”**。

人人交互：让机器和人做各自擅长的事

之前，一提到操控数字世界，人们关注的都是怎么让人操控虚拟的三维物品，也就是让人去操控机器，各种设备的研发重点都是“人机交互”。

但是，当越来越多的人进入了数字世界之后，我们何必纠结于让人与虚拟的场景、虚拟的物品交互呢？在数字世界中，如果人和人互动起来，在投入地交谈和活动时，我们可能根本注意不到周边的环境。

这两年 Facebook 和微软所讨论的元宇宙的重要应用场景，是利用虚拟现实或增强现实来完善会议系统。在 Facebook 展示的元宇宙会议系统中，你可以看到，虚拟人物形象在数字世界中代表你参与会议。那些虚拟形象相当简陋，只有上半个身子，根本不像真实的人。

2021 年底，在这款产品刚推出时，网上出现了很多嘲讽的声音：“就这？这就是元宇宙吗？”“它和电影里高度拟真的元宇宙差别实在太大了。”微软的元宇宙会议系统略微酷炫一点，但其实也是半斤八两。不过，你可别被这些观点影响，Facebook 和微软的做法可能是权宜之计，但它可能反映了一条更加现实的元宇宙技术发展方向。

我们常常在想，构建一个数字世界，如何做到尽可能逼真呢？构建了一个个三维立体的数字物体，如何让人用手更好地跟它互动？人们一直在研究这两个问题，试图找到好的答案。但现在，在这个元宇宙刚开始兴起的时刻，我们不如重新定义一下和数字世界互动的问题，把它变成：如果我们构建数字世界的主要目标，是创造一个让人与人更好地互动的场景，我们应该怎么做？

是的，比起看到逼真的世界，比起用手去操控数字物品，我们可能更渴望跟其他人互动、交流。Facebook 和微软的做法正反映了在探索与数字世界的交互方式时的一个重大的思维转向。之前，我们的关注重点是“人机交互”，也就是怎么让人和虚拟世界的数字物品进行交互。现在，我们应该在元宇宙视角下重新审视这个问题，因为**真正重要的是人与人的互动**。

其实，互联网发展至今给我们的经验是，在网络世界，不管我们是产品设计者、还是网络产品的用户，我们都更看重人与人的互动。因此，现在我们探索元宇宙，关注重点也应该是“人人

交互”。

从“人机互动”向“人人互动”的这个转变让元宇宙应用快速落地的可能性大幅度提高了。因为当重新设定了问题后，你立刻会发现，对于“人人交互”，技术条件的准备要充分得多。

让用户和其他用户在数字世界中更好地交互，这是互联网过去几十年一直在做的事情。我们可以通过文本聊天、语音聊天还有视频聊天和他人互动，相关的技术都已经很成熟了。同时，由于交互的对方是人，所以交互对象可以很容易地理解我们传递的信息。对比而言，要让机器准确理解语气、表情就难得多，更不用说让机器在理解之后，在数字人的脸上渲染出微妙的表情了。

现在，Facebook 以 VR 和元宇宙之名在做的，就是在现有互联网的人人交互的基础上再往前走小半步：在数字世界用一个相对简陋的三维人像代表你。也许将来，它们会再进一步，加上动作、表情等功能，来传递一些现有的聊天软件不能很好传递的信息。虽然现在我们并不能确信 Facebook 让我们用数字人像开会这个功能是不是对的，但我们可以确信，从“人机交互”向“人人交互”是正确的大方向。

把关注重点从人机互动转向人人互动，自然而然地解决了好几个过去难以跨越的难题。

第一个是用计算机模拟反馈的问题。之前我们在讨论人机交互时有一大难点，那就是很难制造反馈，很难让用户感受到一个接近真实的外部世界。这很不容易，因为骗过人的触觉比骗过视觉还要难。但当我们转变到“人人交互”，由于做出反馈的是远方的真实的人，我们的感觉也一下子变得真实起来。

比如说，要真实地模拟拳击的场景，你的拳头打出去需要感受到阻力并反弹回来，你的手还应该感觉到疼痛。以今天的技术，想让这个反馈足够真实仍然相当困难。但是，让你感觉到对面的人在夸奖你却很简单。对方在屏幕上也许只是一个三维数字人像，但他激动地竖起大拇指，你立刻就能理解对方的赞美，而不会深入去考虑那个头像是不是很逼真。

第二个是计算机图形能力问题。虽然现在计算机渲染三维图像的能力越来越强，但是，受到渲染时间、呈现设备的限制，我们能看到的三维立体世界在视觉上还很简陋。微软的 Hololens 眼镜能够呈现的全息图像基本上还只是光影线条，你要相当有想象力才能把它们想象成真实的物体。

技术当然会持续进步，但只有等到计算机的图形能力达到完美的時候，元宇宙才能开始发展吗？当然不是。当元宇宙转向“人人交互”，我们就能够接受相对低画质的数字世界了。人人交互的认同感弥补了画质的不足。

比如，现在有的虚拟世界是由“体素”也就是一个个小的立方体块构成的。如果你只是作为游客去参观，你会觉得它的视觉效果难以忍受。但是，如果你被邀请去参加一个活动，你在现场看到了一些熟悉的朋友，你还可以用文字和简单的动作和对方打招呼，这些小小立方体组成的装饰、人物突然就变得完全可以接受了。

第三个问题是，并非所有问题机器都能处理。我们一直在试图让机器处理所有问题，但事实是，有太多事情机器做不到。比如，让机器模拟真人说话，往往只要几句话它就会被识破，也就是说，能够通过所谓“图灵测试”的机器到现在为止其实还不存在，那些宣称通过了图灵测试的，其实都是加了非常严格的限制条件。

其实，我们何必让机器完成所有的事呢？为什么不可以让机器做机器擅长的事，让人做人擅长的事呢？在转向“人人交互”之后，我们可以不必纠结了。在聊天工具、会议软件、游戏里，它们基于不同的目的、用不同的方式把人连接到了一起，在里面，人在跟其他人互动。在元宇宙里，我们也是在跟进到元宇宙里面的其他人交流。除了一些特定的场景，通常情况下，我们没有必要总是想用机器去模拟和取代人。

互联网在处理人与人的交互方面有太多的经验，有了这个经验宝库，元宇宙一下子就变得容易实现了。元宇宙的发展不是要抛开现有网页、App 建立全新的三维立体的数字世界，然后把所有的互联网用户迁移过去。它应该是现有互联网的自然延伸，它致力于把一部分有必要、有需要的物体变成三维立体的，然后延续互联网中人与人的交互，给所有人更好的数字生活体验。

比如，外卖背后的技术产品处理的也是人与人的交互，我作为用户点餐，餐厅接受订单后匹配骑手，外卖 App 给出反馈，骑手取餐后，我可以看到骑手在地图中移动，看到自己叫的外卖越来越近。和外卖相关的这些产品并没有试图让机器解决所有问题。

未来如果技术继续进步，我们可以让机器做得更多。但在技术、场景还不成熟的情况下，让机器和技术完成它们擅长的，让人做自己擅长的，就可以最快地做出一款足够好用的产品。

总之，从“人机交互”转向“人人交互”让我们一起思考与讨论了一个基本问题，我们用计算机构建数字世界的目的是什么？我们的目的是，让人与人更好地连接。

总结

讲到这里，我们立体互联网部分的课程就要结束了，我来总结一下这部分内容。在第 5、6、7 讲中，我们分别讨论了扫描实体世界、渲染数字世界、操控数字世界三种底层技术手段。

立体互联网是用技术建立实体世界的模型，是用技术渲染出以假乱真的三维数字世界，也是我们可以操控、交互的数字世界。

在这节课关于交互的讨论中，你应该也已经接收到了我的看法：交互的重点不是让人去操控数字世界中的物品，而是用技术让人与人的交流互动变得更加方便、真切。也就是，从“人机交互”走向“人人交互”。

在元宇宙，人与人的交流互动并不只是说话、聊天，人与人的交流互动还包括一起协作创造，包括各种各样的经济活动。这就自然地把我们引向了下一部分对于价值互联网的讨论。


我们试图将社会、经济、文化、组织也都复制到数字世界，最终形成我们可以工作与生活在其中的元宇宙梦想之地，这需要价值互联网的相关技术与应用的支撑。立体互联网和价值互联网加起来，才是元宇宙。

课后题

在这一讲的最后，给你留一道思考题。对于我们说的元宇宙的“人人交互”，你有什么看法？欢迎留言说出你的想法。

分享给需要的人，Ta 订阅超级会员，你最高得 50 元

Ta 单独购买本课程，你将得 20 元

 生成海报并分享

 赞 1  提建议

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

精选留言 (4)

写留言



rondo

2022-05-16

非常赞同人人交互是元宇宙的核心的观点。



六维

2022-05-12

新的思路：人人交互，也许是立体互联网走向下一步的一个阶梯。



GAC·DU

2022-05-11

感觉人人交互像是妥协，或者说是过度，终究要让元宇宙这个局进行下去。蓝图很壮观，可现实是硬件哭诉“臣妾做不到啊”，等我再发育一下。立体互联网是路漫漫其修远，在这种情况下价值互联网如何切入，等立体互联网吗？还是各玩各的呢？



peter

2022-05-11

老师的课很精彩，但我接触本专栏比较晚，没有同步阅读，现在是同步了。下面是01课到05课的一些问题：

Q1：安卓上有AR应用开发吗？

Q2：安卓手机有类似苹果手机那样的三维图形建模能力吗？

Q3：安卓手机有深度相机系统吗？

Q4：我的华为手机上，应用市场上没有《模拟飞行》游戏，该游戏没有手机版本吗？

该游戏是运行在游戏机上吗？

Q5：现在有探险类的游戏或应用吗？05课中提到了“去元宇宙中的亚马逊雨林探险时”，现在市面上有这样的应用吗？

