

第一章 虚拟现实概述

1.1 前 言

虚拟现实(Virtual Reality,简称VR)是继多媒体以后另一个在计算机界引起广泛关注的研究热点。从表面上看,这些技术之间有许多相似之处:它们都是声、文、图并茂,容易被人们所接受;都可用于娱乐、教育等各方面。但是虚拟现实又要求使用立体眼镜、数据手套等高级设备,更为复杂和昂贵。正因为如此,人们对虚拟现实的认识也就不尽相同。有的认为虚拟现实和现有的一些技术(如计算机图形技术等)没有多少区别;有的认为虚拟现实只是一种游戏,没有什么用处;有的认为虚拟现实离我们太远,好似天方夜谭。那么虚拟现实技术到底是干什么的?它有什么特点?到底能否实现?为什么要研究它呢?为了回答这些问题,我们必须对计算机应用的现状进行分析,即在当前的技术水平下,计算机能为我们做什么事情以及我们是如何使用计算机的。

80年代的计算机能处理文本、图形等数据,可用于管理市场动向、人事档案信息等,因此主要是作为办公室自动化的工具。90年代多媒体技术的出现使计算机具有了综合处理声音、文字、图象等多种媒体信息的能力,使计算机有可能以人们所习惯的声音和图象等手段进行信息交流。因此人们可以通过计算机以及计算机网络进行信息交流,这时计算机就成为人们交流信息的媒介。但是目前的计算机技术还不能处理知识,也就是说计算机还缺乏理解和推理的能力。所以虽然计算机的广泛应用减轻了人的负担,但是在绝大多数情况下计算机系统还只是起着辅助作用,最终的理解和决策要由人来完成。例如,计算机集成制造系统(CIMS)通过计算机技术极大地提高了企业的生产管理水平和本质上讲它是一个人的系统和计算机系统共存的复杂环境。有研究表明,集成制造自动化系统的实施中,70%的障碍来自人的因素,其中包括人-机系统体系结构和人-机交互等。此外,计算机的广泛应用使各个领域每天都在产生大量的数据和信息,例如地球人造卫星获取的地球表面数据,使人们能每隔16到20天就可以更新一次地球表面的数字地图。除了科学计算或测量以外,现代化企业的管理和市场分析也在产生大量的数据。这些大量的数据只有经过处理和理解才能成为有用的信息,但是目前的计算机技术还不能自动地理解。研究表明,人具有很强的在三维空间中进行形象思维的能力,如果把这些计算机收集或产生的数据转换成可视化的直观对象(如形状、符号等),就可以借助人的形象思维能力来快速地理解这些数据,得到有用的信息。例如,复杂的地面坐标数据难以为人所理解,但如果把它们转换成表示高度的地形地图,人就可以一目了然地从中找到有用的信息,如可通行的道路等,这就转换成如何使用户与三维空间中的可视化对象进行交互操作的问题。

从以上分析可看到,计算机的应用增强和扩大了人的认识和改造世界的的能力,提供了前所未有的信息服务。但是计算机并不能代替人的角色,计算机要发挥更大的作用,其关键是以更自然和有效的方式与人进行信息交流,也就是要以人所习惯的方式与计算机进

行通信。人们之间通信的特点可归结如下。

(1) 它是利用视觉、听觉、触觉、力觉、嗅觉等多种感觉,在多种媒体空间中的通信,其中最重要的是视觉和听觉。视觉不但是人类最主要的信息获取手段,而且还是一种思维手段,这就是形象思维。

(2) 它是实时交互的通信,除了要求信息的实时交流以外,手势对引导对方的注意力、说明讨论的对象也起很重要的作用。

计算机的应用离不开人的智能,人-机交互技术是发展计算机应用的关键,虚拟现实就是探讨理想的人-计算机通信方式的技术。

1.2 虚拟现实的概念

从本质上说,虚拟现实就是一种先进的计算机用户接口,它通过给用户同时提供诸如视、听、触等各种直观而又自然的实时感知交互手段、最大限度地方便用户的操作,从而减轻用户的负担、提高整个系统的工作效率。根据 VR 所应用的对象的不同,VR 的作用可以表现为不同的形式,例如将某种概念设计或构思可视化 and 可操作化,实现逼真的透视场效果;达到任意复杂环境下的廉价模拟训练目的等。

虚拟现实的定义可以归纳如下:虚拟现实是利用计算机生成一种模拟环境(如飞机驾驶舱、操作现场等),通过多种传感设备使用户“投入”到该环境中,实现用户与该环境直接进行自然交互的技术,这早所谓模拟环境就是用计算机生成的具有表面色彩的立体图形,它可以是某一特定现实世界的真实体现,也可以是纯粹构想的世界。传感设备包括立体头盔(Head Mounted Display)、数据手套(Data Glove)、数据衣(Data Suit)等穿戴于用户身上的装置和设置于现实环境中的传感装置(不直接戴在身上)。自然交互是指用日常使用的方式对环境内的物体进行操作(如用手拿东西、行走等)并得到实时立体反馈。

与 VR 相类似的一个概念是人工现实(Artificial Reality,简称 AR),它是可以更方便地与用可视化技术建立的三维空间中的物体进行交互的技术。这个空间是人工的,但是物体的控制方法就像物体是在现实空间中一样,所以就称为人工现实。例如,可用 AR 技术来漫游用可视化技术建立的大脑结构。另一个相关概念是遥现技术(Telepresence),它是一种基于 VR 的远程控制、遥操作或遥显示技术。

虚拟现实技术具有以下四个重要特征:

(1) 多感知性(Multi Sensory)

所谓多感知就是说除了一般计算机技术所具有的视觉感知之外,还有听觉感知、力觉感知、触觉感知、运动感知,甚至应该包括味觉感知、嗅觉感知等。理想的虚拟现实技术应该具有一切人所具有的感知功能。由于相关技术,特别是传感技术的限制,目前虚拟现实技术所具有的感知功能仅限于视觉、听觉、力觉、触觉、运动等几种,无论从感知范围还是从感知的精确程度都无法与人相比拟。

(2) 存在感(Presence)

又称为临场感(Immersion),它是指用户感到作为主角存在于模拟环境中的真实程度。理想的模拟环境应该达到使用户难以分辨真假的程度(例如可视场景应随着视点的变

化而变化),甚至比真的还“真”,如实现比现实更逼真的照明和音响效果等。

(3) 交互性(Interaction)

交互性是指用户对模拟环境中物体的可操作程度和从环境得到反馈的自然程度(包括实时性)。例如,用户可以用手去直接抓取模拟环境中的物体,这时手有握着东西的感觉,并可以感觉物体的重量(其实这时手里并没有实物),视场中被抓的物体也立刻随着手的移动而移动。

(4) 自主性(Autonomy)

是指虚拟环境中物体依据物理定律动作的程度。例如,当受到力的推动时,物体会向力的方向移动,或翻倒、或从桌面落到地面等。

根据VR的概念及其上述四个特征,我们不难将VR与相关技术区分开来,例如仿真技术,计算机图形技术以及多媒体技术等。

仿真(Simulation)是一门利用计算机软件模拟实际环境进行科学实验的技术,从模拟实际环境这一点看,仿真技术与VR技术有着一定的相似性。但是,首先在多感知方面,仿真技术原则上以视觉和听觉为主要感知,很少用到其他感知(如触觉、力觉等);在存在感方面,仿真基本上将用户视为“旁观者”,可视场景既不随用户的视点变化,用户也没有身临其境之感;在交互性方面,仿真一般不强调交互的实时性。

计算机图形技术(Computer Graphics,CG)是一门实时图形生成与显示的技术,它具有良好的实时交互性和一定的自主性。但是,在多感知和存在感方面与VR有较大差距。CG主要依赖于视觉和听觉感知,虽然生成的图形可以具有三维立体数据,但由于感知手段的限制,用户并不能感到自己和生成的图形世界融合在一起,比如场景不能随自己的视线改变而改变等。

多媒体技术(Multimedia)是利用计算机综合组织、处理和操作多种媒体信息(如视频、音频、图像、文字等)的技术。虽然具有多种媒体,但是在感知范围上仍没有VR广泛,例如多媒体并不包括触觉、力觉等感知。另外,多媒体处理的对象主要是二维的,因此在存在感和交互性方面与VR有着本质的区别。

尽管VR与上述相关技术有较大差异,但是VR又与它们密切相关。VR是在众多的相关技术基础上发展起来的,但它又不是相关技术的简单组合。从技术上看,VR与各相关技术有着或多或少的相似之处,但是在思维方式上,VR已经有了质的飞跃。由于VR是一门系统性技术,因此它不能像某一单项技术那样只从一个方面考虑问题,它需要将所有组成部分作为一个整体,去追求系统整体性能的最优。

VR技术当前受到重视的原因可归纳如下。

(1) 相关技术日趋成熟,为VR的研究提供了基础,如:

- 实时三维图形生成与显示技术。
- 三维声音定位与合成技术。
- 传感器技术:视觉、触觉、力觉传感器等。
- 识别定位技术:语音、三维景物、表情、手势等。
- 环境建模技术:视觉建模、行为建模、CAD技术等。

(2) 各种传感设备以及计算机价格不断降低,使实现VR的应用成为可能。

(3) VR 提供了别的技术难以实现的巨大可能性,从降低成本、减少危险、提高效率、克服物理条件限制、拓宽应用领域等方面都极具吸引力。

(4) 不论是在商业性,还是在实用性以及技术创新上都富有巨大的潜力。

可以预料 VR 技术可能改变人们的思维方式和生活方式,导致一场重大的技术革命。

1.3 VR 系统的构成

VR 系统可以分为以下 3 类:

(1) 封闭式虚拟现实,即与外部现实世界不产生直接交互。

特点:· 虚拟环境可以是任意虚构的、实际上不存在的世界。

- 目的是:娱乐、或验证某一猜想假设、训练、模拟、预演、检验、体验等。
- 任何操作不对外界产生直接作用。

(2) 开放式虚拟现实,即通过各种传感装置与外界构成反馈闭环。

特点:· 虚拟环境是某一现实世界的真实模型。

- 目的是:通过利用虚拟环境对现实世界进行直接操作或遥操作,以达到克服现实环境的限制(如遮挡、遥远、危险、不便到达、不能到达等)或使操作方便、可靠的目的。例如:提供碰撞报警(利用虚拟环境计算相撞距离),减轻操作员的心理负担,减少操作失误。
- 操作按用户的需要,可以对现实世界进行直接作用或得到各种反馈。

(3) 把(1)与(2)结合,即兼备(1)、(2)的特点,这是一种较为实用的虚拟现实系统。

VR 的系统模型表示见图 1-1。用户通过传感装置直接对虚拟环境进行操作,并得到实时三维显示和其它反馈信息(如触觉、力觉反馈等)。当系统与外部世界通过传感装置构成反馈闭环时,在用户的控制下,用户与虚拟环境间的交互可以对外部世界产生作用(如遥操作等)。

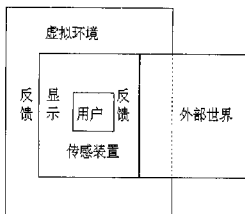


图 1-1 VR 系统模型

VR 系统主要由以下 6 个模块构成(见图 1-2)。

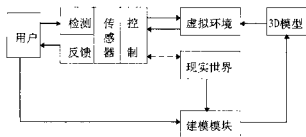


图 1-2 VR 系统的构成

- (1) 检测模块:检测用户的操作命令,并通过传感器模块作用于虚拟环境。
- (2) 反馈模块:接受来自传感器模块的信息,为用户提供实时反馈。
- (3) 传感器模块:一方面接受来自用户的操作命令,并将其作用于虚拟环境;另一方面将操作后产生的结果以各种反馈的形式提供给用户。
- (4) 控制模块:对传感器进行控制,使其对用户、虚拟环境和现实世界产生作用。
- (5) 3D 模型库:现实世界各组成部分的三维表示,并由此构成对应的虚拟环境。
- (6) 建模模块:获取现实世界各组成部分的三维数据,并建立它们的三维模型。

VR 传感装置的类型分类如下:

- (1) 视觉:头盔式立体显示器 HMD(Head-Mounted Display),例如 VPL 公司的 Eye-
phone。它可以分为透过型和非透过型两种。
- (2) 听觉:三维音响输出装置、定位装置。
- (3) 检测手动(包括位置):数据手套(Data Glove),如 DHM(Dexterous Hand
Master)(精密型 DG)、Cyberglove(手指露出型数据手套)等。
- (4) 力反馈、触觉传感器:GROPE 系列手爪等。

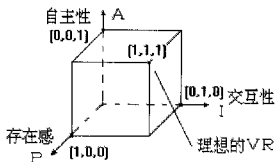


图 1-3 VR 系统的性能评价

(5) 身体运动:数据衣(Data Suit)等。

(6) 语音识别、合成,眼球运动检测等。

根据 VR 技术的特征,可以用系统的存在感 P(Presence)、交互性 I(Interaction)和自主性 A(Autonomy)来评价 VR 系统的性能(见图 1-3),其中 P、I 和 A 的值越大,表示 VR 系统的性能越好。显然,P、I 和 A 的值均取最大时,VR 系统的性能达到极限,即理想的 VR 系统。

1.4 VR 的研究内容

VR 的关键技术和研究内容包括以下几个方面:

(1) 动态环境建模技术

虚拟环境的建立是 VR 技术的核心内容,动态环境建模技术的目的是获取实际环境的三维数据,并根据应用的需要,利用获取的三维数据建立相应的虚拟环境模型。三维数据的获取可以采用 CAD 技术(有规则的环境),而更多的情况则需采用非接触式的视觉建模技术,两者的有机结合可以有效地提高数据获取的效率。

(2) 实时三维图形生成技术

三维图形的生成技术已经较为成熟,而这里的关键是如何实现“实时”生成,为了达到实时的目的,至少要保证图形的刷新频率不低于 15 帧/秒,最好是高于 30 帧/秒。在降低图形的质量和复杂程度的前提下,如何提高刷新频率将是该技术的研究内容。

(3) 立体显示和传感器技术

VR 的交互能力依赖于立体显示和传感器技术的发展。现有的 VR 设备还远远不能满足系统的需要,例如头盔式三维立体显示器有以下缺点:过重(1.5~2kg)、分辨率低(LCD,图象质量差)、延迟大(刷新频率慢)、行动不便(有线)、跟踪精度低、视场不够宽、眼睛容易疲劳等,因此有必要开发新的三维显示技术。同样,数据手套、数据衣等都有延迟大、分辨率低、作用范围小、使用不便等缺点。另外,力觉和触觉传感装置的研究也有待进一步深入,VR 设备的跟踪精度和跟踪范围也有待提高。

(4) 应用系统开发工具

VR 应用的关键是寻找合适的场合和对象,即如何发挥想象力和创造性。选择适当的应用对象可以大幅度地提高生产效率、减轻劳动强度、提高产品质量。为了达到这一目的,必须研究 VR 的开发工具。例如,VR 系统开发平台、分布式 VR 技术等。

(5) 系统集成技术

由于 VR 系统中包括大量的感知信息和模型,因此系统的集成技术起着至关重要的作用。集成技术包括信息的同步技术、模型的标定技术、数据转换技术、数据管理模型、识别与合成技术等等。

1.5 VR 技术的应用

如前所述,VR 本质上是关于人与计算机通信的技术,所以人们认为它几乎可用于支

持任何人类活动,也就是可用于任何领域。当然 VR 的实际应用会受到硬件设备的限制。

VR 技术的前身是图形仿真器(如飞行仿真器),其概念早在 60 年代就已提出,但受限于软、硬件技术的发展,直到 80 年代中才开始逐步兴起,到了 90 年代才开始有 VR 商品系统问世。1992 年出现了世界上第一批 VR 开发工具系统,1993 年出现了众多的 VR 应用系统,并且应用范围也大大扩展。据统计,1993 年世界上出现的 805 个 VR 应用系统中,用于仿真的有 73 个,其它的依次为可视化 67,教育 66,训练 65,娱乐 65,图形 64,军事 52,航天 50,通讯 50,医疗 49,建筑 46,视听艺术 41,商业 40,遥控机器人 39 以及通信 38 个。可以看出,娱乐、视听艺术以及教育方面的应用发展最快(172 项),其次是军事和航天领域(162 项)。随着软、硬件技术的发展,价格的不断下降,除上述应用之外,虚拟现实将进入家庭,并成为丰富生活的重要因素。

值得一提的是医疗领域正日益受到广泛的重视,VR 可以在医疗领域发挥极其重要的作用。例如 VR 可以用于解剖教学、复杂手术过程的规划,在手术过程中提供操作和信息上的辅助、预测手术结果等。此外,远程医疗服务也是一个很有潜力的应用领域,例如在偏远的山区,通过远程医疗 VR 系统,人们可以不用进城也能够接受名医的治疗,对于危重病人还可以实施远程手术;在战场上,可以通过 VR 系统对前线的危急伤员进行远程手术,使他们得到及时的抢救。在远程手术系统中,医生对虚拟病人模型进行手术,他的动作通过高速网络(或卫星)传送到遥远处的手术机器人。手术的实际图象通过机器人上的摄像机传回给医生的头盔立体显示器,并将其和虚拟病人模型叠加,为医生提供有用的信息。美国斯坦福国际研究所(SRI)已研制出一个远程手术实验系统。

在航天领域 VR 技术也大有用武之地。例如,失重是航天技术中的一个必须克服的困难,因为在失重情况下对物体的运动难以预测。因此为了在太空中进行精确的操作,需要进行长时间的仿真训练,以适应失重时操作的特点,现在这种训练一般是水下或特殊的飞机中进行。在均衡浮力的水池中,可以模拟失重状态,但是水的阻力会妨碍对惯性的仿真。在飞机垂直下降时可以产生真正的失重状态,但是时间却非常短(小于 1 分钟)。因此需要寻求新的方法,VR 就是一个合适的选择。美国航天局 NASA(National Aeronautics and Space Administration)在“哈勃太空望远镜修复和维护”计划的仿真训练中就使用了 VR 技术。在训练中,宇航员坐在一个模拟“载人操纵飞行器”功能并带有传感装置的椅子上。椅子上有用于在虚拟空间中作直线运动的位移控制器和用于绕宇航员重力中心调节宇航员朝向的旋转控制器。宇航员头戴立体头盔显示器,用于显示望远镜、航天飞机以及太空的模型,并用数据手套作为与系统进行交互的手段。这样,训练时宇航员就可以在望远镜周围进行操作,并且可以通过虚拟手接触黄色的操纵杆来抓住需要更换的“模块更换仪”(Modular Replacement Instruments;MRI)。抓住 MRI 之后,宇航员可以利用座椅的控制器在空中飞行。座椅上另有三个按钮分别用于望远镜外盖的开/闭、望远镜天线的开/闭以及望远镜太阳能板朝向的调节。经过该 VR 系统的训练,宇航员于 1993 年 12 月成功地完成了将哈勃太空望远镜上损坏的 MRI 用从航天飞机中取出的备件进行更换这一极其复杂而又费时的任务,充分显示了 VR 技术的巨大作用。NASA 计划在该 VR 系统中增加触觉和力觉传感装置,以增加训练的有效性。

虚拟现实技术的研究在中国还刚刚起步,虽然在 VR 硬件(传感器)技术和其它关键

技术方面都需要做大量的研究工作,但是我们不能因此而丧失对 VR 技术及其应用的信念。现在普遍存在着对 VR 的神秘感,认为 VR 是一种昂贵的技术,若不购买大量的先进设备就无法开展 VR 的应用研究;或者认为只有当掌握了相应的硬件技术之后,才有可能进行 VR 的应用开发。显然这是对 VR 的一种不正确的认识。诚然,有了先进的设备可以大大地方便 VR 的应用开发,但是利用现有的最基本的设备(如廉价的立体眼镜、跟踪器、微机等)也可以进行 VR 的应用系统开发。我们总不能等待有了钱再去着手应用系统的开发,那样我们将永远跟在别人的后面,而且越来越落后。VR 应用系统的开发不仅仅是一个设备的问题,好的设备不等于可以开发出好的应用系统。VR 的应用对象的确定本身就是一个需要认真研究的创造性思维过程,只要应用得巧妙,即使是廉价的设备,也可以大幅度提高生产效益或精度,发挥出 VR 的巨大潜力。另外,设备的价格是与生产的数量成反比的,因此 VR 设备的价格也是随时间逐步降低的。当我们的应用研究已达到一定的水平时,相信 VR 设备的价格也会大大降低,那时可以正好拿这些设备为我所用。

第二章 虚拟现实运用的主要技术与设备

2.1 概 述

虚拟现实是多种技术的综合,它运用的主要技术包括:实时三维计算机图形技术;广角(宽视野)的立体显示;观察者(头、眼)的跟踪;手及手势、人体姿势的跟踪;立体声;触觉、力觉反馈;语音输入输出。

虚拟现实不仅仅只是一个较大、较快或较高质量的常规计算机图形系统,因为在通常的计算机图形系统中,用户只是一个外部的观测者,通过屏幕(或窗口)去观看一些人造的环境(图形)。而在VR中,上述技术通常被用来把用户变成人游环境中的一个参与者。在本节中,我们对这些主要技术进行一个简单的介绍,让读者对关键技术及难点有一个初步的了解。在后面的章节中,我们将结合具体的设备对某些技术进行较详细的介绍。在以后的详细介绍中,我们并不完全根据上面所列的次序来进行。我们将围绕视觉、听觉、触觉(力觉)和辨认,即“看”、“听”、“摸”和“认”来进行介绍及讨论。注意,这里的辨认是指计算机对用户输入的辨认,包括对手势、姿势以及语音的辨认。

2.1.1 实时三维计算机图形技术

比较而言,从计算机模型去产生图形图象并不是太难的事情。如果有足够准确的模型又有足够的时间,我们可以产生物体在不同光照条件下相当精确的图象,但是,这里的关键是实时,我们可能什么都有了,但没有足够的时间,例如在飞行模拟系统中,图象的刷新非常重要,最好每秒要显示25~30帧,至少也不能少于15帧。由于我们既要求图象的质量,又要求图象的速度,再加上非常复杂的虚拟环境,问题就会相当困难。

现代的图形(计算机)工作站每秒钟大约可以描绘35万块独立的多边形表面,同时还可以显示中等质量的表面明暗程度的变化,这里还包括把看不见的表面从图形中去掉以及其它一些光线的作用,例如反光的亮区等等。如果以最低标准(每秒15帧)来计算,我们只能描绘大约由2万3千块多边形表面所组成的世界。再考虑到表面的纹理、影调(明暗变化)等,多边形的块数还要减少。现实场景仅用两万块左右的多边形来表示是远远不够的。如果在PC机上,情况就更差了。你可以看到一些在PC机上的“虚拟现实”的图形全是由一些大块大块的多边形组成,与现实场景相差甚远。一般来说,如果能够达到显示几十万块多边形所组成的虚拟环境就相当满意了。

虽然计算机图形技术在VR中起着非常重要的作用,但我们不准备在本书中详细讨论,因为这方面的著作已不在少数。有兴趣的读者可以参阅有关文献。

2.1.2 广角(宽视野)的立体显示

人都有两只眼睛,每只眼睛都能看到周围的世界,但两眼的位置略有不同。两只眼睛得到的略有不同的图象在脑子里融合起来,形成一个周围世界的整体景象,这个景象中包括了距离远近的信息。当然,距离的信息也可以从其它的方法获得,例如眼睛聚焦的远近、物体大小的比较等。在VR中,双目立体视觉也起了很大的作用。通常,用户的两只眼睛看到的不同图象是分别产生的,然后显示在不同的显示器上。有的系统采用单个显示器,但用户戴上特殊的眼镜后,一只眼睛只能看到显示的奇数帧,而另一只眼睛只能看到偶数帧,而奇、偶帧之间的不同(也就是视差)就产生了立体感。

这些技术在第二章节中还要详细介绍。简单的VR系统的显示可以是普通计算机的屏幕,但有人认为,既然要让用户觉得处于虚拟环境中,用户的视场在水平方向至少要有 110° ,垂直方向 60° ,还应该有不小于 30° 的双目立体重叠。由于在VR中两眼看到的图象是分别产生的,因此首先遇到的就是同步问题。至于各种光学设备的校正以适应不同人的两眼之间距离的差别等,人们也在进行研究。

当前,有很多头戴显示器采用的是电视广播的标准,图象的质量比通常看到的计算机图形显示差得多,再加上采用了液晶显示器,实际有效的分辨率就更差了。如果还需要把图象拉伸成为宽视角的景象,结果可能是景象变成了由一块一块平面拼起来的图象。

2.1.3 用户(头、眼)的跟踪

在人造的环境中,每个物体相对于系统的坐标系都有一个位置与姿态,而用户也是如此。用户看到的景象要由用户的位置和头(眼)的方向来确定。在计算机图形技术中,视场(视角)的改变是通过鼠标或键盘来实现的,因此用户的视觉系统与其运动感知系统之间的联系被分割了。

如果利用头部跟踪来相应地改变图象的视角,那么上面所说的图象视角变化与人的运动感知系统之间就可以联系起来,这就会增加感觉上的逼真程度。头部跟踪的另一个优点是,用户不仅可以从双目立体视觉中去认识环境,而且可以通过头部的运动去观察认识环境,这往往可以提供更多的关于距离的信息。

虽然有很多人在研究眼球的跟踪,但到目前为止,大部分研究的成果还没有用在VR上。

2.1.4 手及手势、人体姿势的跟踪

在与计算机的交互中,键盘和鼠标是目前最常用的工具。键盘对于文字输入是极其有效的,但对图形图象就显得不大好用。鼠标对于二维空间操作起来很有效,但对于三维空间不是很方便。在三维空间中,因为有六个自由度,我们很难找出一个比较直观的办法把鼠标的平面运动映射成三维空间中的任意运动。现在,已经有一些设备可以提供六个自由度,其中比较有名的是3Space数字化仪和SpaceBall空间球。

3Space数字化仪用的是磁场原理,我们在本章还要介绍。SpaceBall是一个球,当它被操作者握在手掌中时,操作者对球施加的力被测量出来,并被映射成三维空间中的旋转。

与平移量。

目前,在VR中最“时髦”的输入设备可能就是数据手套和数据衣了,但它们昂贵的价格不是一般应用所能接受的。

2.1.5 立体声

人类能够很好地判定声源的方向。在水平方向上,我们是靠声音的相位差及强度的差别来确定方向的,因为声音到达两只耳朵的时间或距离会有不同。你可能熟悉用耳机去听立体声音乐,你的左右耳听到的不同声音是事先在不同位置录制的,所以给你一种声音的方向感,但是当你的头部转动时,你听到的声音的方向就随之改变。但在VR系统中,我们希望声音的方向与用户头部的运动无关。

2.1.6 触觉(力)反馈

在一个VR系统中,用户可以看到了一个虚拟的杯子。你可以设法去抓住它,但是你的手没有像真正接触杯子时的那种感觉,虚拟杯子的“表面”不能够阻止你的手穿过它。这是一个相当困难的问题。

一种办法是在手套的内层安装一些可以振动的触点来模拟触觉。其它方法包括在手套中加一种液体橡胶的夹层,或者使一些物质在电场的作用下从液态变成固态等。但要达到真正的触觉效果也很困难,我们在触觉及力觉一节中再详细介绍。

2.1.7 语音输入输出

有人说目前计算机讲的话听起来像一个醉汉,这说明质量还需要改进。当然计算机要听懂用户的话,则是另一个难题。用人的自然语言作为计算机的输入,除了可靠性以外,输入效率也是值得考虑的。

2.2 视觉显示技术及设备

我们的脑子每秒钟要接受大量的视觉信息,并处理这些信息以获得对周围环境的理解。至今为止,我们还不很清楚人类究竟是怎样来处理视觉信息并进行理解的,我们也不准备详细讨论这个问题。但是,从虚拟现实视觉显示的角度来讲,因为我们试图用计算机产生的图形图象来造成用户身临其境的“假象”,就有必要介绍一下人是怎么从二维(平面)图象中来获得立体感的。

为了造成视觉上的虚拟环境,我们还必须将获得的或创造出的虚拟环境的图象显示到我们的眼前,显示可以是二维或三维的。所以,在这一节里,我们还将介绍一些典型的VR显示设备,并对一些关键技术进行讨论。

2.2.1 立体视觉

虚拟现实视觉技术只有在对人类的立体视觉有所了解的情况下才能够发展,那么,人是怎样从视觉中获得立体信息的呢?可能有不少人认为,我们看周围的环境有立体感或者

说有深度(距离)信息,是因为我们用两只眼睛的缘故。这个理解在一定程度上是对的,但不完全。我们感知到的深度信息有两种,一种是物体离我们有多远,这是绝对距离;另一种是两个或几个物体之间的相对位置,或相对距离。据研究,我们的脑子能从四个方面获得深度(距离)线索。这四个方面是:静态图象中的深度线索、由运动造成的深度线索、生理上的深度线索以及双目立体视觉深度线索。这里,我们用“线索”这个词,表示我们对于究竟怎样获得深度信息还不大清楚。如果我们用“信息”,似乎太直接了当了。

(1) 静态图象中的深度线索

当我们用一只眼睛去看周围的世界或者去看一张图画、照片等时,我们也能发现关于深度的信息。这是通过下面五个方面来获得的:一是图象中物体的明暗变化以及阴影等;二是重叠,当你看到甲物体的一部分被乙物体遮住了,你就知道乙物体比较靠近你;三是尺寸的大小,同样的物体离观察者越近,尺寸越大。但是,这种线索只有观察者对物体尺寸熟悉时才能利用;第四种是利用物体上直线的投影角度来判断距离,这叫做线性透视,图 2-1 就是一个明显的例子;第五种线索是纹理。例如,你看一大片稻田时,在近处的稻秧组成的图案就很粗,甚至可以分清这株和那株,但在远处,稻秧就成了很密的一片了。

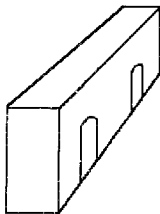


图 2-1 线性透视例子

(2) 运动深度线索

当用一只眼睛看几棵树的上部时,如果它们之间没有相互遮挡的话,你很难说出哪棵树近、哪棵树远。但当左右移动一下头部时,树的远近就很容易判断了,因为近处的树在你看来移动得多或者说得快,而远处的树则动得慢一些。

(3) 生理上的深度线索

从上面的介绍来看,如果全靠静止图象中的线索或者头部运动的话,人有一只眼睛就够了。但是实际上,近处物体上最重要的深度信息要靠两只眼睛来获取。从生理上来探讨,当你看近处物体时,因为两只眼睛要看同一处,眼球就要向内(向脸的中心线)转动。物体离你的脸越近,眼睛的转动就越多。如果看的物体在六七米之外,两眼视线就基本平行了。所以,眼球转动的多少就提供了深度的信息。

从眼睛聚焦的过程来看,为了要看清一个物体,我们的眼睛要调节晶状体使得物体的像能清晰地聚焦在我们的眼底上,这种调节过程能给我们一些深度的线索,你可以试验这种方法,但获得的深度信息往往不大准确。

(4) 双目视觉线索

当我们用双眼看同一景物时,由于左右眼的位置不同,所以看到的景象不完全一致,这种差别我们称之为双目视差。虽然我们还不清楚大脑是如何把两幅不完全相同的图象融合起来而加以理解的,但是我们确实可以从双目视差中获得立体信息。在虚拟现实的显示技术中,利用双目视差来造成立体感是一个重要的方面。

2.2.2 视觉显示设备

为了显示虚拟环境,必须把显示器、专用的光学系统以及位置和方向的跟踪系统组合在一起,而且这些部件必须配合工作得像一个整体一样。目前,视觉显示硬件虽然有不同形状和大小,而且在适用性和结构上有很大的不同,但它们所用的元件大致上是同一类型的。目前,常见的二种基本显示方法是:头盔立体镜、头盔单目镜和可移动的视觉显示器。我们将在下面对几种典型的设备进行介绍,说明它们是怎样工作的,显示的图象又是怎样呈现在我们眼前的。我们注意到,不少视觉显示装置是和耳机以及姿态识别或跟踪系统结合在一起的。

从应用上来讲,视觉显示装置又可以分为半投入式和全投入式两种。使用半投入式的显示器时,用户可以同时看到现实世界和计算机所产生的图象;当使用全投入式显示器时,用户只能看到计算机生成的图象。

(1) 达摩克里斯之剑头盔显示器

假如让你来设计一个包括图象显示和头部位置、方向追踪的设备,那么,你可能想到一个像头盔一样的东西,再加上一副特大的护目镜。你已经不是最早有这种想法的人了,这种设备叫做头盔式显示器 HMD。1968年,当 Ivan Sutherland 在哈佛大学当学生时,就首先提出了头盔显示器的概念。Sutherland 被认为是在计算机图形学领域的先驱者,他设计了达摩克里斯之剑头盔显示器。

达摩克里斯之剑是“危险临头”的意思,因为这个设计用了从天花板上悬挂下来的机械臂作为位置追踪器,而这只机械臂就像一把剑正对着人的头。Sutherland 将两个小的阴极射线管(CRT)显示器装在水盔的两边,用来显示立体图象的线框图。这些图象由玻璃透镜反射,经简单的光学元件进入眼睛。这个系统能够显示具有简单几何形状物体的线框图,用户看到的线框图叠加在真实环境之上(图 2-2)。也就是说,这是一种半投入式的显示器,用户既看到了周围的世界,又看到了虚拟物体的线框图。

(2) NASA 头盔显示器

NASA 是美国国家航空航天管理局的简称。1984年,NASA 的科学家们在加利福尼亚州的一个研究中心开始致力于开发他们自己的立体镜头盔显示器,用于遥控机器人和空间站的研究。在资金不足条件下,NASA 的科学家们只得考虑低成本的头盔显示器。他们最初使用了一对廉价的黑白袖珍电视机上的液晶显示器(LCD),液晶显示器的对角线的尺寸为 2.7 英寸(约 7cm),并有 320×240 象素的分辨率。这些液晶显示器不仅比



图 2.2 Sutherland 的 HMD

微型阴极射线管便宜得多,而且它们工作时所需要的能量极低,这有利于控制系统的重量和能量消耗。

(3) UNC 头盔显示器

虽然北卡罗来纳州大学(University of North Carolina,UNC)的研究人员早在 80 年代初就开始研制头盔显示器,但他们在 1989 年与美国空军技术学院共同研制的头盔显示器最为有名。这种创新式的头盔显示器安装在自行车头盔的周围(见图 2-3)。显示器用

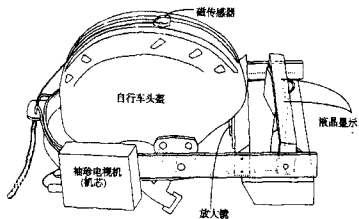


图 2.3 UNC 开发的头盔显示器

了两台液晶显示器并安装有荧光背景照明,专用的透镜位于头盔的前方,支撑小臂与头盔的两侧相连接。控制显示的电子装置安装在头盔背面的一个小盒内。头盔显示器上各零部件的重量平衡是很关键的,如果不平衡,它就会掉下来或戴不稳。

有趣的是,UNC的研究人员在设计头盔显示器时遇到了一个障碍:鼻子。如果在每只眼睛的前方安装一块大的液晶显示板,由于要碰到鼻子,用户当然不舒服;如果移动液晶显示板使它远离眼睛,结果是液晶显示板的中心不易与眼睛的视轴对准。UNC的研究人员想出了解决办法,改变液晶显示板的角度,使之避开鼻子。这种办法也有问题,当倾斜着观察环境时,就必须对图象进行光学上的校正,否则虚拟环境看上去是斜的。在UNC系统中,用户的头部位置和方向由磁性定位系统跟踪。

基于Sutherland的视场叠加显示的设计思想,UNC的研究人员改进了设计。在新的方案中,液晶显示器不是垂直或成一个角度置于用户的眼前,而是水平地伸出到头盔之外,大约与眉毛持平(见图2-4)。显示的图象是向下投射的,经过一个放大镜到达安装成45°角的半镀银镜面。该头盔显示器不仅可以看到镜子表面反射过来的图象,而且可以看到镜子远处的“真实世界”。这种显示器具有近25°的水平视场。

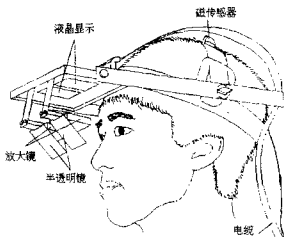


图2-4 改进后的UNC开发的头盔显示器

UNC的设计表明,在实验室里研制显示器系统(不是为了商业目的)时,可采用拼接液晶显示器和头盔的方法,其成本还不到商业系统的一半。

(4) VPL的EyePhone系统

我们还有必要谈一下EyePhone(眼-耳机)头盔显示器。VPL公司的EyePhone系统采用有360×240像素分辨率的彩色LCD显示,并用一种叫做LEEP的立体广角镜片,以提供100°的水平视场。显示器安装在脸部的前面,并有橡胶密封器遮挡住周围的光线。方向的检测使用了磁性定位系统。新型EyePhone头盔显示器的总重量仅为2.5磅(约1kg多一点)。

(5) PrivateEye 单目镜显示器

单目镜显示器也值得一提,尤其是-一种叫 PrivateEye 的产品。整个系统的大小约为 $3 \times 3.3 \times 9 \text{cm}^3$, 系统套在用户的头上, 而显示器悬挂在用户的一只眼睛的前面(见图2-5)。用户看到的是一个大约 12 英寸的计算机屏幕, 安装在面前大约 70cm 远的半空中。

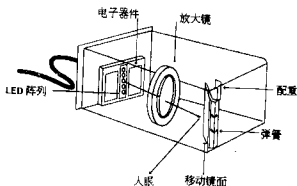


图 2-5 PrivateEye 的结构图和外视

(6) BOOM 可移动式显示器

无可否认, 头盔显示器在实际使用时不是很方便。例如, 一位建筑师如果戴上头盔显示器去观察建筑物的透视图, 打电话时需要摘掉它, 使用时又要把它戴上, 他会感到十分烦琐。针对虚拟现实的某些应用, 人们提出了更为灵活性的设计, 即可移动式显示器。用户可以把显示器方便地置于眼前, 而且可以很快地移开。例如, 由 NASA 研制的双目全方位显示器 BOOM (Binocular Omni Orientation Monitor) 就是典型的可移动式显示器。除

了BOOM中用的LEEP光学元件外,BOOM的设计与头盔显示系统之间几乎没有相似之处。BOOM没有使用在大多数头盔显示器中已成为标准的液晶显示器,而是用小型的阴极射线管。阴极射线管能产生比液晶显示器小得多的像素,所以图象比较柔和。BOOM所用的阴极射线管的对角线尺寸约为6.5cm,分辨率为1280×1024象素,显示彩色图象。

BOOM显示器装在一只小黑盒内,而小黑盒则挂在由两段构成的一个机械臂上。可以通过两个孔去观察小盒里的显示器,并可抓住附在小盒上的把手并移动小盒。整个系统是精确平衡的,所以当 you 你放开显示器时,显示器会平稳地悬在空中。

可移动式显示器对于大多数应用是优于头盔显示器的,这不仅是由于它易于“戴上”和“放下”,而且由于其位置及方向的跟踪是通过测量机械臂关节的角度变化来实现的,因此非常精确。这种精确性对于科学可视化大有好处。

2.2.3 显示方法及效果

大部分的VR显示设备采用液晶显示器。例如,NASA的液晶显示器图象由76,800个象素构成,这些象素通常是不足以产生一幅清晰的图象的。观察者与显示器之间距离的缩小又进一步显著地减低了图象的清晰度。大多数微型液晶显示器的设计是考虑从大约60~70cm处观察的,在该距离处由于单个象素很小,所以图象看起来还比较柔和清晰,但当接近液晶显示器观察时,图象的清晰度就会降低,因为这时可以看到构成图象的单个象素。在看报纸上的相片时,可遇到同样的情况。从正常的观察距离观看时,图象看上去清晰、鲜明。当靠近看时,可以看清楚相片实际上是由很多小点组成,这时图象也就不再显得清晰了。然而,与阴极射线管相比,液晶显示器的优点(如轻便、廉价、体积小等)依然超过了它的缺点,所以人们在头盔显示器中仍然使用液晶显示器。

用于UNC系统中的彩色液晶显示器,其对角线的尺寸为3英寸(约7.5cm)。早期的头盔显示器仅使用了简单的放大镜,用户仅具有大约55°的水平视场,这对于大部分的VR应用是不够的。后来,他们利用了LEEP(Large Expanse Extra Perspective)光学元件来有效地扩大了观察的范围,明显地提高了虚拟环境中的视觉效果。LEEP用了6个镜片(每边3个),可以把2.7英寸的液晶显示范围扩展到120°的水平和垂直视场。

上面介绍的PrivateEye显示器采用了十分特别的方法来产生图象,它使用了排列成一行的280个发射红光的发光二极管(LED)来代替微型阴极射线管或液晶显示器。在发光二极管的对面是一面摆动的镜子。在镜子的每一个摆动周期中,每个发光二极管闪烁720次。发光二极管的闪烁和镜子的运动是精确同步的,其结果看起来像是在一个有280个水平象素和720个垂直象素的屏幕上显示一样。可以调整显示器上的光学元件来改变屏幕在空间的位置,看到的图象可以出现在20cm以外甚至2m左右。

头盔显示器也可通过光纤来显示图象。光投影装置通过光纤束把图象送到头盔,经过光学系统的投射产生一个广角立体显示。将射束分裂器(镀银的透明镜子)放在眼睛的前方,用户可同时看到真实世界和计算机所产生的图象。

还有一些显示器的设计是直接由摄像机得到图象,而不是直接从计算机得到生成的图形。两个视频摄影机安装在一个可以转动的架子上,经过控制台把图象送入显示系统。用这种系统去显示计算机图形的最简便方法是将摄影机分别对准两个监视器,每个监视