

第 3 章

虚拟现实的相关技术

实现 VR 系统除了需要功能强大的硬件支持,对相关的软件技术也提出了很高的要求。它必须能够充分地发挥硬件设备的潜能,有效地管理和使用系统的各种资源。VR 系统的相关技术主要包括计算机图形学、虚拟环境建模技术、视觉实时动态绘制技术、人机自然交互技术、虚拟声音技术等。

【学习目标】

- (1) 掌握虚拟现实技术的计算机图形学基本原理。
- (2) 了解虚拟现实技术的虚拟环境建模技术。
- (3) 了解虚拟现实技术的视觉实时动态绘制技术。
- (4) 了解虚拟现实技术的人机自然交互技术。
- (5) 了解虚拟现实技术的虚拟声音技术。

3.1 虚拟现实中的计算机图形学原理

计算机图形学技术是实现虚拟现实系统的重要理论基础之一,计算机图形学是研究如何在计算机环境下生成、处理和显示图形的一门学科。它所研究的主要内容是在计算机环境下景物的几何建模方法、模型的处理方法、绘制技术、图形输入和控制的人机交互界面等。它可以把用户脑海中的构想有效地转化为用户能观察的视觉图像,再结合传感技术、用户交互技术,使得计算机系统构造出一个用户能看得到、触摸得到、感受得到并能沉浸其中的虚拟环境。本节将简要介绍虚拟现实系统中的计算机图形学原理。

3.1.1 三维图形绘制原理

在现实世界中,我们所观察到的对象都是三维的,这些对象包含深度信息,而计算机只能绘制二维图像。为了在计算机屏幕上显示三维环境中的实体对象,计算机需要把相应的三维实体对象映射到二维屏幕上,实现对三维实体对象的显示,此过程需要对模型进行坐标转换、视点转换、投影转换和几何变换等多种图形变换,这一过程类似于照相机的拍照过程,如图 3-1 所示。

此过程通常要经历几个步骤。

- (1) 将照相机固定在三脚架上使之对准场景(视点变换)。
- (2) 让拍摄的场景位于取景框中的合适位置(模型变换)。

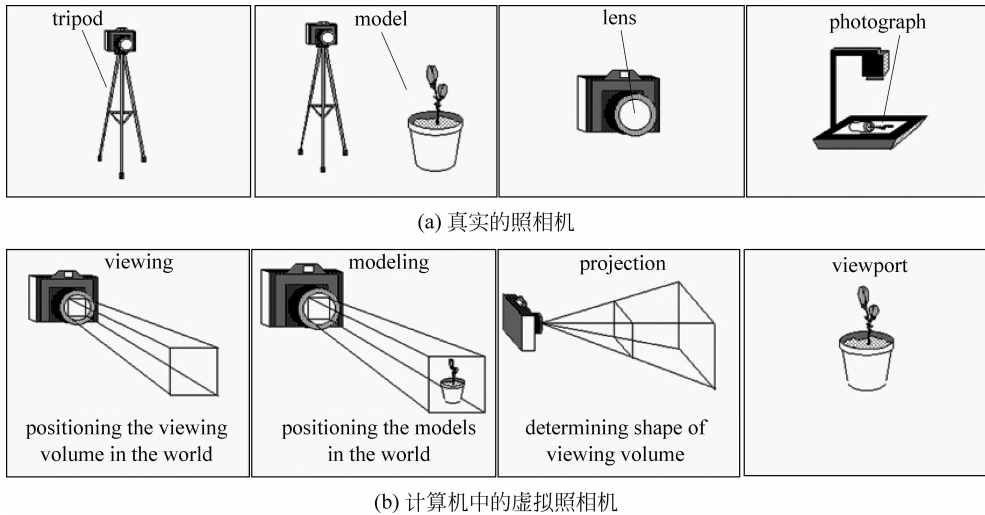


图 3-1 真实的照相机与计算机中的虚拟照相机对比

(3) 选择照相机镜头或调整放大倍数(投影变换)。

(4) 决定相片的大小(视口变换)。

综上所述,一个三维物体从建模到最后在计算机屏幕上显示的大致流程如图 3-2 所示。

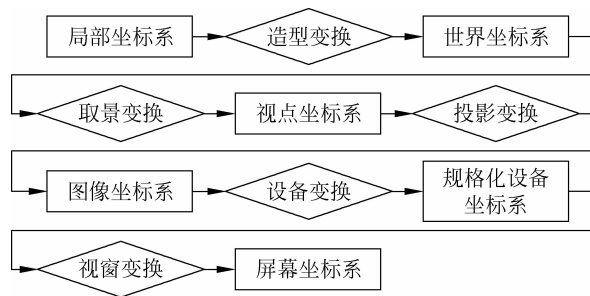


图 3-2 三维物体显示流程图

其中,视点变换相当于调整照相机的位置方向,模型变换则指定模型的位置和朝向,如对模型进行旋转、平移、缩放或执行这些操作的组合。投影变换则相当于选择照相机镜头,可以将投影变换视为指定视野或视景物。即确定哪些物体在视野内以及这些物体在视野内的大小。投影变换和视点变换一起决定场景将如何映射到计算机屏幕上,投影变换指定映射的方式,而视口变换指定场景将被映射到什么样的屏幕区域中,视口指定图像占据的计算机屏幕区域。

3.1.2 坐标系

为了在虚拟环境中定位虚拟实体对象,一个重要的概念即坐标系。目前,虚拟环境中常用的坐标系(图 3-3)分为以下几种:

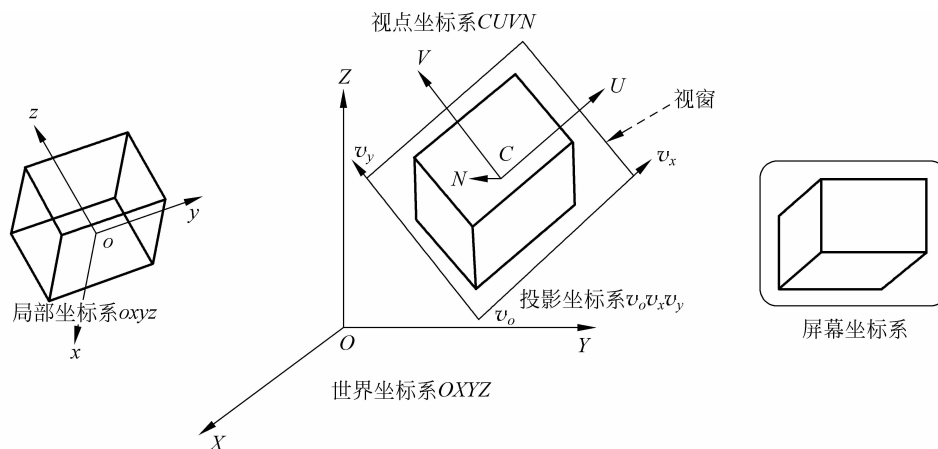


图 3-3 虚拟环境中的坐标系

1. 世界坐标系

世界坐标系(World Coordinate System, 又称全局坐标系)主要用于虚拟场景中的所有图形对象的空间定位和定义,包括观察者的位置、视线等。计算机图形系统中涉及的其他坐标系统都是参照它进行定义的。

2. 局部坐标系

局部坐标系(Local Coordinate System, 又称造型坐标系)主要为考察物体方便起见,独立于世界坐标系来定义物体几何特性,通常是在不需要指定物体在世界坐标系中的方位的情况下,使用局部坐标系。如在定义局部物体时,通过指定局部坐标系的原点在世界坐标系中的方位,然后通过几何变换,就可很容易地将局部物体放入世界坐标系内,使它由局部上升为全局。

3. 观察坐标系

观察坐标系(Viewing Coordinate System, 又称视点坐标系)通常是以视点的位置为原点,通过用户指定的一个向上的观察向量来定义整个坐标系统,默认为左手坐标系。观察坐标系主要用于从观察者的角度对整个世界坐标系内的对象进行重新定位和描述,从而简化几何物体在投影面的成像的数学推导和计算。

4. 成像面坐标系统

成像面坐标系统是一个二维坐标系统,主要用于指定物体在成像面上的所有点,往往是通过指定成像面与视点之间的距离来定义成像面,成像面有时也称投影面,可进一步在投影面上定义称为窗口的方形区域来实现部分成像。

5. 屏幕坐标系统

屏幕坐标系统也称设备坐标系统,它主要用于某一特殊的计算机图形显示设备(如光栅

显示器)的表面的点的定义。在多数情况下,对于每一个具体的显示设备,都有一个单独的坐标系,在定义了成像窗口的情况下,可进一步在屏幕坐标系中定义称为视图区的有界区域,视图区中的成像即为实际所观察到的。

总之,为在三维空间创建和显示一个或多个几何物体,首先必须建立世界坐标系,然后指定视点的方位、视线及成像面的方位。为了观察到物体的成像,还必须在各坐标系之间视图变换之后,再进行投影变换。

3.1.3 透视投影

人们观察自然界的物体时,所得视觉映像同观察点和观察方向有关。同样,要用计算机生成一幅三维视图,也需要确定观察点和观察方向,还需要将观察范围以外的部分图形裁剪掉。而且,由于图形输出设备通常都是二维的,为了现实虚拟物体,必须将三维几何模型经过投影变换,转换为显示器二维平面透视投影图。从三维物体模型描述到二维图形描述的过程称为投影变换。

投影变换分为平行投影和透视投影两种。

(1) 平行投影。平行投影可以看成投影中心在无限远处的投影,如图 3-4 所示。平行投影物体的相对度量保持不变(例如两个等长线段的投影结果仍然是等长的),适用于建筑和机械设计。

(2) 透视投影。投影射线汇聚于投影中心,或者说投影中心在有限远处的投影。即从空间选定的一个投影中心到物体上每点连成直线从而构成了一簇射线,射线与选定的投影平面的交点集便是物体的投影,如图 3-5 所示。透视投影符合人类的视觉特点,产生的投影效果更为真实。

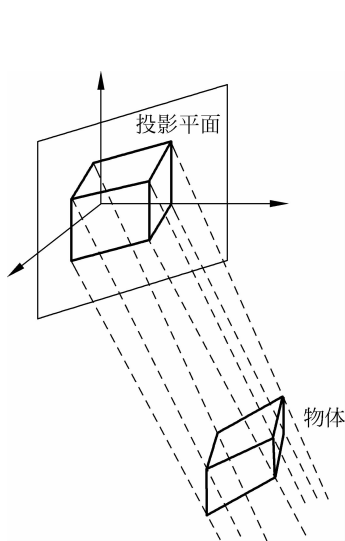


图 3-4 平行投影变换示意图

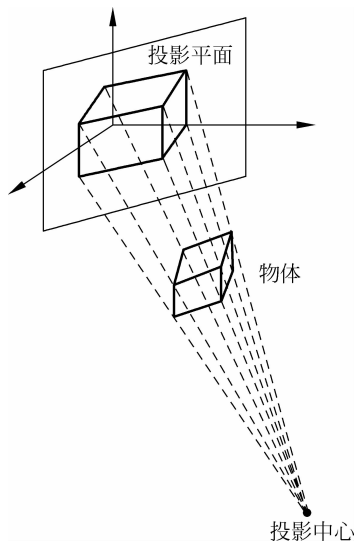


图 3-5 透视投影变换示意图

3.1.4 三维裁剪

通过跟踪从物体上的点到观察者的直线可以在投影平面上捕捉到物体的透视投影。直线与投影平面的交点即为物体的投影点。只有当物体位于观察者的视场内时,这一过程才有效。所以当物体位于观察者的后方、上方、下方、左方和右方的情况可以不予考虑。然而有些情况下,物体的一部分是可见的,而其他部分是不可见的。这意味着物体要被裁剪掉不可见的部分。并且对左、右眼要分别进行三维裁剪。三维裁剪算法是计算机图形系统的基本特征,这里不作详述。

3.1.5 三维消隐

消隐是在一定观察方向下消除不可见的线和面,有时也称为可见性测试。对于一个三维物体,当沿着投影视线观察它时,由于物体中各种表面或其他物体的遮挡,使某些线段或面成为不可见,这些不可见的线段或面就称为隐藏线或隐藏面。为了使计算机显示图像有一定的逼真感,在确定了观察点及观察方向后,首先要解决的问题,就是决定显示对象的哪些部分是可见的,哪些部分为自身或其他物体所遮挡是不可见的,即找出隐藏线和隐藏面。消除这些不可见的部分,只显示可见的线、面,可以使所显示的图形没有二义性。

虽然各种消隐算法在可见性测试和不可见面消除方法上区别不大,但这些消隐方法有时还可以一起被统称为不可见面的消除,简称消隐。在三维造型技术、真实感图形的显示、虚拟场景的显示,以及在地形和地图的绘制中,消隐都起到至关重要的作用。所以研究和实现消隐算法,并根据场景的复杂度和各个不同应用领域的场景来提高消隐算法的速度是很有必要的,它对整个三维图形显示、真实感图形的显示以及各种地形地貌图形的显示是很有意义的。

3.2 虚拟现实建模技术

虚拟现实是一种逼真地模拟人在自然环境中视觉、听觉、嗅觉、运动等行为的一种全新的人机交互技术,其最终目标是使用户置身于一个由计算机生成的虚拟环境中。建模是对现实对象或环境的逼真仿真,虚拟对象或环境的建模是虚拟现实系统建立的基础,也是虚拟现实技术中的关键技术之一。建模是对现实对象或环境的虚拟,对象建模主要研究对象的形状和外观的仿真。环境建模主要涉及物理建模、行为建模、声音建模等。

评价虚拟现实建模的技术指标包括以下内容。

(1) 精确度。它是衡量模型表示现实物体精确程度的指标,也是表现场景真实性的主要元素之一。

(2) 操纵效率。在实际运用过程中,模型的显示、运动模型的行为、在有多个运动物体的虚拟环境中的冲突检测等都是频度很高的操作,必须高效实现。

(3) 易用性。创建有效的模型是一个十分复杂的工作,建模者必须尽可能精确地表现物体的几何和行为模型,建模技术应尽可能容易地构造和开发一个好的模型。

(4) 实时显示。在虚拟环境中,模型的显示必须在某个极限帧率以上,这往往要求快速显示。

3.2.1 几何建模技术

虚拟对象基本上都是几何图形构成的。采用几何建模方法对物体对象虚拟主要是物体几何信息的表示和处理,描述虚拟对象的几何模型(例如多边形、三角形、顶点以及它们的外表(纹理、表面反射系数、颜色)等),即用一定的数学方法对三维对象的几何模型进行描述。物体的形状由构成物体的各个多边形、三角形及顶点来确定;物体的外观则是由表面纹理、材质、颜色、光照系数等决定的。

1. 形状建模

要表现三维物体,最基本的是绘制出三维物体的轮廓,利用点和线来构建整个三维物体的外边界,即仅使用边界来表示三维物体。三维图形物体中运用边界表示的最普遍方式是使用一组包围物体内部的表面多边形来存储物体的描述,多面体的多边形表示精确地定义了物体的表面特征,但对其他物体,则可以通过把表面嵌入到物体中来生成一个多边形网格逼近,曲面上采用多边形网格逼近可以通过将曲面分成更小的多边形加以改进。由于线框轮廓能快速显示以概要地说明表面结构,因此,这种表示在设计 and 实体模型应用中普遍采用。通过沿多边形表面进行明暗处理来消除或减少多边形边界,以实现真实性绘制。

形状建模通常采用的方法如下。

1) 人工几何建模方法

(1) 对于对象的形状建模常常可以利用现有的图形库来创建。常用的图形库有图形核心系统(Graphical Kernel System,GKS)、程序员级分层结构交互图形系统(Programmer's Hierarchical Interactive Graphic System,PHIGS)、开放式图形库等。利用这些图形库建模具有编程容易、效率较高等优点。

(2) 利用建模软件进行建模,如 AutoCAD、3ds MAX 等,这些软件具有可视化、交互性强等特点,可以方便地创建虚拟对象的几何模型。

2) 自动几何建模方法

自动化的建模方法很多,最典型的是利用三维扫描设备对实际物体进行三维建模。如三维扫描仪又称为三维数字化仪,是一种将真实世界的立体彩色图形转换为计算机能直接处理的数字信号的装置。它在 VR 技术、影视特技制作、高级游戏、文物保护等方面有着广泛的应用。事实上,在 VR 系统中,靠人工构造大量的三维彩色模型费时费力,且真实感差。利用三维扫描技术可为 VR 系统提供大量的与现实世界完全一致的三维彩色模型数据。

2. 外观建模

对象的外表是一种物体区别于其他物体的质地特征,VR 系统中虚拟对象的外表真实感主要取决于它的表面反射和纹理。一般来讲,只要时间足够宽裕,用增加物体多边形的方法可以绘制出十分逼真的图形表面。但是 VR 系统是典型的限时计算与显示系统,对实时性要求很高。因此,省时的纹理映射(Texture Mapping)技术在 VR 系统几何建模中得到了广泛的应用。用纹理映射技术处理对象的外表,一是增加了细节层次以及景物的真实感,二是提供了更好的三维空间线索,三是减少了视景多边形的数目,因而提高了帧刷新率,增强了复杂场景的实时动态显示效果。

1) 纹理映射

所谓纹理映射,就是把给定的纹理图像映射到物体表面上,并不是特定的几何模型,使用纹理映射可以避免对场景的每个细节都使用多边形来表示,进而可以大大减少环境模型的多边形数目,提高图形的显示速度。

纹理映射的过程如图 3-6 所示。

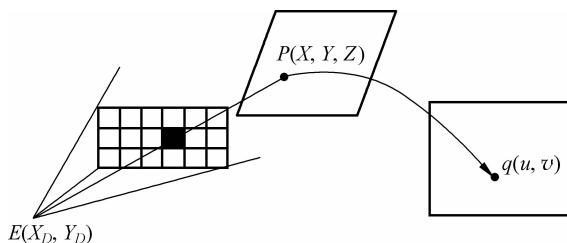


图 3-6 纹理映射过程示意图

$E(X_D, Y_D)$ 代表眼点, $P(X, Y, Z)$ 代表物体上的点, $q(u, v)$ 代表纹理上的像素点。所以,纹理映射实际上是屏幕空间、物体空间和纹理空间之间的一系列变换过程。虚拟对象的纹理可通过拍摄对应物体的照片,然后将照片扫描进计算机的方法得到,也可用图像绘制软件建立。

从物体表面的质地特征来看,纹理映射分为颜色纹理映射和凹凸纹理映射。前者是通过颜色色彩或明暗度的变化来表现物体的表面细节,后者则是通过对景物表面各采样点法向量的扰动来表现物体几何形状凹凸不平的粗糙质感。

从具体算法来看,纹理映射可分为标准纹理映射和逆向纹理映射。标准纹理映射是对纹理表面均匀扫描,并直接映射到屏幕空间。逆向纹理映射是对屏幕上的每一个像素,通过逆映射寻找到物体空间上的对应点,再在纹理空间找到相应的像素点,取得纹理值经滤波后显示该像素。

纹理映射技术应用很广,尤其描述具有真实感的物体。例如绘制一面砖墙,就可以用一幅真实的砖墙图像或照片作为纹理贴到一个矩形上,砖墙就很逼真。如果不用纹理映射的方法,则墙上的每一块砖都必须作为一个独立的多边形来画,这样会增加许多的计算量。纹理映射也常常运用在其他一些领域,如飞行仿真中常把一大片植被的图像映射到一些多边形上用以表示地面,或用大理石、木材、布匹等自然物质的图像作为纹理映射到多边形上表示相应的物体。

2) 光照

当光照射到物体表面时,可能被吸收、反射或者折射。被物体吸收的部分转化为热,而那些被反射和折射的光传到我们的视觉系统,使我们能看见物体。为了模拟这一物理现象,我们使用一些数学公式来近似计算物体表面按照什么样的规律、什么样的比例来反射或者折射光线。这种公式称作明暗效应模型。

假设物体不透明,那么物体表面呈现的颜色仅仅由其反射光决定。通常,反射光由三个分量表示,分别是环境反射光、漫反射光、镜面反射光。

(1) 环境反射光。环境反射光在任何方向上的分布都相同。环境反射光用于模拟从环境中周围物体散射到物体表面再反射出来的光。环境反射光可以用下面的公式表示:

$$I = K_a I_a$$

其中, K_a 是环境反射常数, 与物体表面的性质有关; I_a 是入射的环境光光强, 与环境的明暗有关。

(2) 漫反射光。漫反射光的空间分布也是均匀的, 但是反射光的光强与入射光的入射角的余弦成正比。通常可以用下面的公式表示:

$$I = K_d I_i \cos\theta$$

其中, K_d 是漫反射常数, 与物体表面的性质有关; I_i 是入射光的光强; θ 是入射角, 如图 3-7 所示。

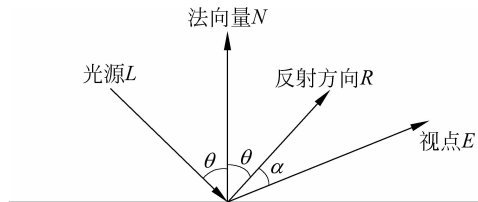


图 3-7 入射方向、反射方向及实现方向示意图

(3) 镜面反射光。镜面反射光为朝一定方向的反射光, 它遵循光的反射定律。反射光和入射光对称地位于表面法向量的两侧。对于纯镜面, 入射光严格地遵守光的反射定律单向反射出去。然而真正的纯镜面是不存在的, 一般光滑表面, 实际上是由许多朝向不同的微小平面组成的, 其镜面反射光存在于镜面反射方向的周围。常常使用余弦函数的某次幂来模拟一般光滑表面反射光的空间分布, 光照处理算法表示为

$$I = K_s I_i \cos^n \alpha$$

在计算机图形学中, 光滑的曲面常用多边形进行逼近表示, 因为处理平面比处理曲面容易得多。但是, 这样就会失去原本曲面的光滑度, 呈现多边形。这种现象是因为不同平面的法向量不同, 形成不同平面之间的不连续的光强跳跃。

图 3-8 是光照示意图, 图中白色小球是一个点光源, 光线在立方体和球体两个对象上发生反射, 产生明暗效果。

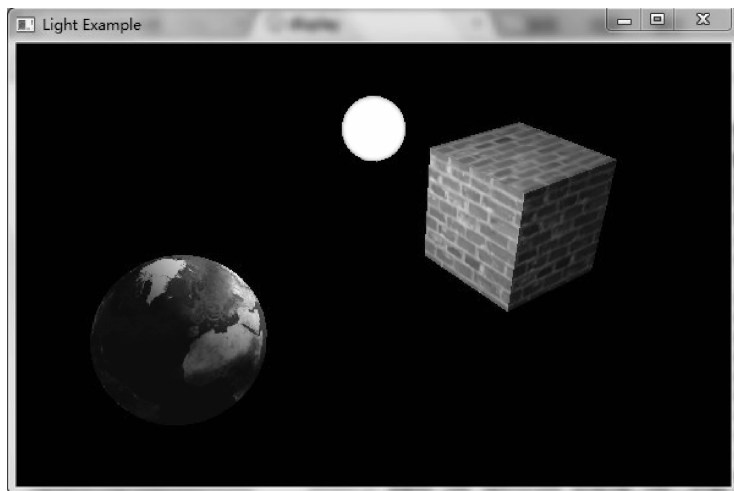


图 3-8 光照示意图

3.2.2 物理建模技术

在虚拟现实系统中,虚拟对象必须像真的一样,这需要体现对象的物理特性,包括重力、惯性、表面硬度、柔软度和变形模式等,这些特征与几何建模相融合,形成更具有真实感的虚拟环境。例如,用户用虚拟手握住一个球,如果建立了该球的物理模型,用户就能够真实地感觉到该球的重量、硬软程度等。

物理建模是虚拟现实中较高层次的建模,它需要物理学和计算机图形学的配合,涉及力学反馈问题,重要的是重量建模、表面变形和硬度的物理属性的体现。分形技术和粒子系统就是典型的物理建模方法。

1. 分形技术

自然界存在的典型景物如高山、沙漠、海滨、白云,这些都是大自然多姿多彩的美丽景色,也是传统数学难以描述的怪异曲线、曲面。在虚拟现实系统的虚拟世界中,必然要出现这些怪异的曲线、曲面,既然传统的数学对其难以描述,必然要借助新的数学工具。分形理论认为,分形曲线、曲面具有精细结构,表现为处处连续,但往往是处处不可导,其局部与整体存在惊人的自相似性。因此,分形技术是指可以描述具有自相似特征的数据集。自相似特征的典型例子是树。若不考虑树叶的区别,当我们靠近树梢时,树的细梢看起来也像一棵大树。由相关的一组树梢构成的一根树枝,从一定距离观察时也像一棵大树。这种结构上的自相似称为统计意义上的自相似。

自相似结构可用于复杂的不规则外形物体的建模。该技术首先用于水流和山体的地理特征建模。例如,我们可以利用三角形来生成一个随机高程的地理模型,取三角形三边的中点并按顺序连接起来,将三角形分割成4个三角形,同时,我们给每个中点随机地赋一个高程值,然后递归上述过程,我们就可以产生相当真实的山体了。

分形技术的优点是简单的操作就可以完成复杂的不规则物体的建模,缺点是计算量太大。因此,在虚拟现实中一般仅仅用于静态远景的建模。

2. 粒子系统

所谓的粒子系统,就是将人们看到的物体运动和自然现象,用一系列运动的粒子来描述,再将这些粒子运动的轨迹映射到显示屏上,在显示屏上看到的就是物体运动和自然现象的模拟效果了。

粒子系统是一种典型的物理建模系统。其基本思想是:采用大量的、具有一定生命和属性的微小粒子图元作为基本元素来描述不规则的模糊物体。在粒子系统中,每一个粒子图元均具有形状、大小、颜色、透明度、运动速度和运动方向、生命周期等属性,所有这些属性都是时间 t 的函数。随着时间的流逝,每个粒子都要经历“产生”、“活动”和“消亡”三个阶段。

利用粒子系统生成画面的基本步骤如下。

- (1) 产生新的粒子。
- (2) 赋予每一新粒子一定的属性。
- (3) 删去那些已经超过生存期的粒子。

(4) 根据粒子的动态属性对粒子进行移动和变换。

(5) 显示由有生命的粒子组成的图像。

粒子系统采用随机过程来控制粒子的产生数量,确定新产生粒子的一些初始随机属性,如初始运动方向、初始大小、初始颜色、初始透明度、初始形状以及生存期等,并在粒子的运动和生长过程中随机地改变这些属性。粒子系统的随机性使模拟不规则模糊物体变得十分简便。

粒子系统应用的关键在于如何描述粒子的运动轨迹,也就是构造粒子的运动函数。函数选择的恰当与否,决定效果的逼真程度。其次,坐标系的选定(即视角)也有一定的关系,视角不同,看到的效果自然不一样。

在虚拟现实,粒子系统常用于描述火焰、水流、雨雪、旋风、喷泉、火灾、战场硝烟、飞机尾焰、爆炸烟雾等现象。

3.2.3 运动建模技术

几何建模只是反映了虚拟对象的静态特性,而 VR 中还要表现虚拟对象在虚拟世界中的动态特性,而有关对象位置变化、旋转、碰撞、伸缩、手抓握、表面变形等方面的属性就属于运动建模问题。

1. 对象位置

对象位置通常涉及对象的移动、伸缩和旋转。因此往往需要用各种坐标系来反映三维场景中对象之间的相互位置关系。例如,假设我们开着一辆汽车围绕树驾驶,从汽车内看该树,该树的视景就与汽车的运动模型非常相关,生成该树视景的计算机就应不断地对该树进行移动、旋转和缩放。

2. 碰撞检测

在虚拟世界中,必须对用户和虚拟对象的移动加以限制,否则就会出现两个对象自由穿透的奇异情景。因此,碰撞检测技术也是 VR 系统中不可缺少的关键技术之一。有了碰撞检测,在虚拟环境中进行漫游时,才可避免诸如观察者穿墙而过、3D 游戏中被距离很远的子弹击倒等现实中不真实情况的发生。

碰撞检测技术不仅要能检测是否有碰撞的发生、碰撞发生的位置,还要计算出碰撞发生后的反应。由于碰撞检测需要具有较高的实时性和精确性,如必须在很短的时间(如 30~50ms)内完成,其技术难度很高。目前较成熟的碰撞检测算法有层次包围盒法和空间分解法等。

1) 层次包围盒法

利用体积略大而形状简单的包围盒把复杂的几何对象包裹起来,在进行碰撞检测时,首先进行包围盒之间的相交测试,若包围盒不相交,则排除碰撞可能性;若相交,则接着进行几何对象之间精确的碰撞检测。显然,包围盒法可快速排除不相交的对象,减少大量不必要的相交测试,从而提高碰撞检测的效率。常用的包围盒箱不仅仅是矩形,还可以是圆球、圆柱等。边界箱的选择和需要碰撞检测的虚拟对象有关,尽量做到算法简单,检测精度较高。层次包围盒法应用较为广泛,适用于复杂环境中的碰撞检测。

2) 空间分解法

空间分解法是将整个虚拟空间分解为体积相等的小单元格,所有对象都被分配在一个或多个单元格之中,系统只对占据同一单元格或相邻单元格的对象进行相交测试。这样,对象间的碰撞检测问题就被转化为包含该对象的单元格之间的碰撞检测。当对象较少且均匀分布于空间时,这种方法效率较高;当对象较多且距离很近时,由于需要进行单元格更深的递归分割,这样需要更多的空间存储单元格,并需要进行更多的单元格相交测试,从而降低了效率。因此,空间分解法适用于稀疏环境中分布比较均匀的几何对象间的碰撞检测。

3.3 实时动态绘制技术

3.3.1 实时动态绘制技术的基本原理

实时动态绘制技术是指利用计算机为用户提供一个能从任意视点及方向实时观察三维场景的手段,它要求当用户的视点改变时,图形显示速度必须跟得上视点的改变速度,否则就会产生迟滞现象。实时动态绘制技术所期望的是图像帧速高而等待时间短。

三维立体图形所含信息量比二维平面图形要大得多,特别是在进行消隐、浓淡、阴影及纹理等处理时,都必须通过大量、快速的计算来实现,而且虚拟环境越复杂,每秒产生的帧数就越少。因此,当生成虚拟环境的视图时,必须要设计出好的数据空间和视频图像,使计算机系统尽快运行,每秒生成足够数量的新帧,从而保证系统刷新频率不低于 20~30 帧/秒。

此外,等待时间是实时动态绘制技术的另一个关键指标,它指用户动作的开始与在显示器上出现对这一动作的响应反馈之间的时长。除了与图像帧速有关外,等待时间还要再计入跟踪定位系统的过渡时间以及应用程序的运行时间。对于 VR 系统来说,如果等待时间过长,新视点场景就不能及时更新,会产生人动而景未动或人未动而景动的视觉交叉错位现象。这不但大大降低了用户的沉浸感,严重的还将产生在 VR 技术中常说的“仿真病”现象,即由于人脑对于视觉与体感上的差错无法协调而导致出现眩晕、呕吐等症状。

3.3.2 实时动态绘制技术简介

一般来说,实时动态绘制技术可分为基于图形和基于图像的两种绘制技术。

1. 基于图形的实时动态绘制技术

在 VR 系统中,为保证三维几何图形的实时快速生成,至少要求图形的刷新频率不低于 20~30 帧/秒,它取决于画面的照明度、阴影、纹理和图形的复杂度等因素。因此,如何选择合适的算法来降低场景的复杂度是关键问题。目前,用于降低场景的复杂度、提高实时动态绘制速度的常用方法有场景分块、可见消隐及细节选择等。

1) 场景分块

场景分块(World Subdivision)是指把一个复杂的场景划分为多个相互之间几乎或完全不可见的子场景。例如,可以把一个大的建筑物按房间划分成多个子部分,当用户在某个房间浏览时,只能看到房内的场景及与之相连的房间,而与它相距较远的和其他楼层的房间场

景则处于不可见的状态。这样系统就能大量地减少在某一时刻需要显示的多边形数目,从而有效地降低了可视场景的复杂度。但是,这种方法的缺点是仅对封闭空间有效,对开放空间则难以处理。

2) 可见消隐

在三维场景的绘制过程中,基于给定的视点和视线方向,决定场景中哪些物体的表面是可见的,哪些是被遮挡而不可见的,称之为场景的可见消隐(Visibility Culling)。使用此方法能使系统仅显示用户当前能“看见”的场景,它与场景分块有所不同的是,场景分块仅与用户所处的场景位置有关,而可见消隐则与用户的视点关系密切。当用户“看见”的场景占整个场景的比例很小时适用此法,而当用户“看见”的场景比较复杂时,这种方法就不起作用了。

3) 细节选择

即使采用了场景分块技术和可见消隐技术,有时用户能“看见”的场景仍会相当复杂,为此产生了细节选择(Level of Detail, LOD)方法。所谓细节选择,即为每个物体建立多个相似的模型,不同模型对物体的细节描述不同。对物体细节的选择越精确,模型也就越复杂。VR 系统将根据物体在屏幕上所占区域的大小及用户视点等因素自动地为各物体选择不同的细节模型,从而减少所需显示的多边形数目。例如,当观察者离一棵树很远时,系统可以选择一个较为简单(如只能显示出树型而分辨不出树叶)的模型来代表它;而随着观察者的逐步接近,系统就将采用分段更替的方法,选择越来越精确的模型对它加以描述。与前两种技术相比,细节选择是一种更有前途的方法,因为它不仅可以用于封闭空间模型,也可以用于开放空间模型。但是,这种方法也对场景模型的描述及其维护提出了较高的要求。如图 3-9 所示为头部的 LOD 模型。

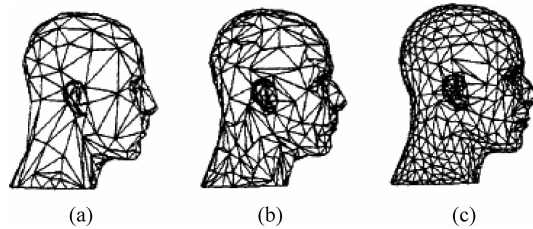


图 3-9 LOD 示意图

2. 基于图像的实时动态绘制技术

基于几何图形的实时动态绘制技术有许多优点,特别是场景中的观察点和观察方向可以随意改变而不受限制。但其造型过程复杂、工作量大,且对每一个观察点或观察方向都需要进行成像计算。因此,近年来的新研究热点是基于图像的实时动态绘制技术。用图像合成代替几何建模来创建复杂的场景,不但真实感强,而且整个过程都可以在二维空间中进行,其绘制时间不取决于场景的复杂度,而只与显示分辨率有关。

目前,基于图像的绘制技术主要有全景技术和图像的插值及视图变换技术两种,此外还有基于分层表示及全视函数等方法。

1) 全景技术

全景技术是指在一个场景中的一个观察点,用相机每旋转一下角度就拍摄一张照片所

得到的一组照片,再在计算机中采用各种工具软件拼接成一个全景图像。它所形成的数据较小,对计算机配置要求低,适用于桌面式 VR 系统。建模速度虽快,但一般一个场景只有一个观察点,因此交互性较差。

2) 图像的插值及视图变换技术

全景技术只能在指定的观察点进行漫游。现在,研究人员研究了根据在不同观察点所拍摄的图像,交互地给出或自动得到相邻两个图像之间对应点,采用插值或视图变换的方法,求出对应于其他点的图像,生成新的视图,根据这个原理可实现多点漫游的要求。

3.4 人机自然交互技术

虚拟现实技术的研究目标是消除人所处的环境和计算机系统之间的界限,即在计算机系统提供的虚拟空间中,人可以使用眼睛、耳朵、皮肤、手势和语言等各种感觉器官直接与之发生交互,这就是虚拟环境下的自然交互技术。

3.4.1 手势识别技术

手势是一种较为简单、方便的交互方式,系统只需跟踪用户手的位置以及手指的夹角就有可能通过已接收的手势下达指令。目前,能识别手势的典型交互设备是数据手套,它能够实时捕捉手的运动,能对较为复杂的手的动作进行检测,包括手的位置、方向和手指弯曲度等,并可根据这些信息对手势进行分类,因而较为实用。但是,数据手套价格昂贵,而且相应的测量装置也限制了人的自由运动。与之相比,采用摄像机输入手势则是一种更为先进的方法。它由摄像机连续拍摄手部的运动图像后,先采用轮廓提取的办法识别出手上的每一个手指,进而再用边界特征识别的方法区分出一个较小的、集中的手势。这种识别方式的最大优点就是不干扰用户。

3.4.2 面部表情识别技术

表情是人类用来表达情绪的一种基本方式,是非语言交流中的一种有效手段。人脸表情识别也是人工心理理论研究的重要组成部分。人脸表情含有丰富的人体行为信息,对它的研究可以进一步了解人类对应的心理状态。计算机和机器人如果能够像人类那样具有理解和表达情感的能力,并能够自主适应环境,这将从根本上改变人与计算机之间的关系,使计算机能够更好地为人类服务。迄今为止,计算机的表情识别能力还与人们的期望相差较远,但仍然能从现阶段的研究成果中感受到这项技术的魅力所在。计算机面部表情的识别技术通常要分为三个步骤进行,即表情的跟踪、表情的编码和表情的识别。

3.4.3 眼动跟踪

在虚拟世界中生成视觉的感知主要依赖于对人头部的跟踪,即当用户的头部发生运动时,生成虚拟环境中的场景将会随之改变,从而实现实时的视觉显示。但在现实世界中,人们可能经常在不转动头部的情况下,仅仅通过移动视线来观察一定范围内的环境或物体。在这一点上,单纯依靠头部跟踪是不全面的。为了弥补这一缺陷,我们在 VR 系统中引入眼

动跟踪技术。

眼动跟踪技术的基本工作原理是利用图像处理技术,使用能锁定眼睛的特殊摄像机,通过摄入从人的眼角膜和瞳孔反射的红外线连续地记录视线变化,从而达到记录、分析视线追踪过程的目的。

常见的视觉追踪方法有眼电图、虹膜-巩膜边缘、角膜反射、瞳孔-角膜反射、接触镜等。

视线跟踪技术可以弥补头部跟踪技术的不足之处,同时又可以简化传统交互过程中的步骤,使交互更为直接。因而,目前多被用于军事(如飞行员观察记录等)、阅读以及帮助残疾人进行交互等领域。

3.4.4 触觉、力觉反馈传感技术

触觉、力觉反馈传感技术是运用先进的技术手段,将虚拟物体的空间运动转变成特殊设备的机械运动,在感觉到物体的表面纹理的同时,也使用户能够体验到真实的力度感和方向感,从而提供一个崭新的人机交互界面。即运用“作用力与反作用力”的原理来达到传递力度和方向信息的目的。在 VR 系统中,为了提高沉浸感,用户希望在看到一个物体时,能听到它发出的声音,并且还希望能够通过自己的触摸,来了解物体的质地、温度、质量等多种信息,这样才全面地了解该物体,从而提高 VR 系统的真实感和沉浸感,并有利于虚拟任务的执行。如果没有触觉、力觉反馈,操作者无法感受到被操作物体的反馈力,得不到真实的操作感,甚至可能出现在现实世界中非法的操作。

触觉感知包括触摸反馈和力量反馈所产生的感知信息。触觉感知是指人与物体对象接触所得到的全部感觉,包括触摸感、压感、振动感、刺痛感等。触摸反馈一般指作用在人皮肤上的力,它反映了人触摸物体时的感觉,侧重于人的微观感觉,如对物体的表面粗糙度、质地、纹理、形状等的感觉。而力量反馈是作用在人的肌肉、关节和筋腱上的力,侧重于人的宏观整体感受,尤其是人的手指、手腕和手臂对物体运动和力的感受。如果用手拿起一个物体时,通过触摸反馈可以感受到物体是粗糙的或坚硬的等,那么通过力量反馈,才能感受到物体的质量。

由于人的触觉相当敏感,一般精度的装置根本无法满足要求,所以触觉与力觉反馈的研究相当困难。目前大多数 VR 系统主要集中并停留在力觉反馈和运动感知上面。其中,很多力觉系统被做成骨架的形式,从而既能检测方位,又能产生移动阻力和有效的抵抗阻力。而对于真正的触觉绘制,现阶段的研究成果还很不成熟;对于接触感,目前的系统已能够给身体提供很好的提示,但不够真实;对于温度感,虽然可以利用一些微型电热泵在局部区域产生冷热感,但这类系统还很昂贵;对于其他一些感觉(如味觉、嗅觉和体感等),至今仍然对它的理论知之甚少,有关此类产品相对较少。虽然目前已研制成了一些这样的触摸和力量反馈产品,但它们大多还是粗糙的、实验性的,距离真正的实用尚有一定的距离。

3.5 三维虚拟声音技术

在虚拟现实系统中,听觉信息是仅次于视觉信息的第二传感通道,听觉通道给人的听觉系统提供的是声音显示,也是创建虚拟世界的一个重要组成部分。而虚拟环境中的三维虚

拟声音与人们熟悉的立体声音有所不同。立体声虽然有左右声道之分,但就整体效果而言,立体声来自听者面前的某个平面,而三维虚拟声音则是来自围绕听者双耳的一个球形中的任何地方,即声音出现在头的上方、后方或者前方。因此在虚拟环境中,能使用户准确判断出声源的准确位置,符合人们在真实世界中听觉方式的声音系统称为三维虚拟声音。

3.5.1 三维虚拟声音的特征

三维虚拟声音的特征主要包括全向三维定位特性和三维实时跟踪特性。

(1) 全向三维定位特性(3D Steering)。是指在三维虚拟环境中把实际声音信号定位到特定虚拟声源的能力。它能使用户准确地判断出声源的精确位置,从而符合人们在真实世界中的听觉方式。

(2) 三维实时跟踪特性(3D Real-Time Localization)。是指在三维虚拟环境中实时跟踪虚拟声源的位置变化或虚拟影像变化的能力。当用户转动头部时,这个虚拟声源的位置也应随之变动,使用户感到声源的位置并未发生变化。而当虚拟发声物体移动位置时,其声源位置也应有所改变。因为只有声音效果与实时变化的视觉相一致,才可能产生视觉与听觉的叠加和同步效应。

例如,假想在虚拟房间中有一台正在播放节目的电视。如果用户站在距离电视较远的地方,则听到的声音也将较弱,但只要他逐渐走近电视,就会感受到越来越大的声音效果;当用户面对电视时,会感到声源来自正前方,而如果此时向左转动头部或走到电视左侧的话,他就会立刻感到声源已处于自己的右侧。这就是虚拟声音的全向三维定位特性和三维实时跟踪特性。可以说,一套性能良好的三维声音系统将能使所有虚拟声音的体验与人们在现实生活中取得的经验相同。

3.5.2 语音识别与合成技术

在虚拟现实系统中,语音应用技术主要是指基于语音进行处理的技术,主要包括语音识别技术和语音合成技术,它是信息技术处理领域的一项前沿技术。

1. 语音识别技术

语音识别技术是指计算机系统能够根据输入的语音识别出其代表的具体意义,进而完成相应的功能。一般的方法是事先让用户朗读有一定数量文字、符号的文档,通过录音装置输入到计算机,于是计算机就准备好了用户的声音样本。以后,当用户通过语音识别系统操作计算机时,用户的声音通过转换装置进入计算机内部,语音识别技术便将用户输入的声音与事先存储好的声音样本进行对比。系统根据对比结果,输入一个它认为最“像”的声音样本序号,这就可以知道用户刚才念的声音是什么意思,进而执行此命令。因此通过语音识别技术,计算机可以“听”懂人类的语言。

一个完整的语音识别系统可大致分为以下三个部分。

(1) 语音特征提取。其目的是从语音波形中提取出随时间变化的语音特征序列。

(2) 声学模型与模式匹配(识别算法)。声学模型通常将获取的语音特征通过学习算法产生。在识别时将输入的语音特征同声学模型(模式)进行匹配与比较,得到最佳的识别

结果。

(3) 语言模型与语言处理。语言模型包括由识别语音命令构成的语法网络或由统计方法构成的语言模型,语言处理可以进行语法、语义分析。对小词表语音识别系统,往往不需要语言处理部分。

一般来说,语音识别的方法有3种:基于声道模型和语音知识的方法、模式匹配的方法以及利用神经网络的方法。

(1) 基于声道模型和语音知识的方法起步较早,在语音识别技术提出的初期,就有了这方面的研究,但由于其模型及语音知识过于复杂,现阶段没有达到实用的阶段。

(2) 模式匹配的方法发展比较成熟,目前已达到了实用的阶段。在模式匹配方法中,要经过特征提取、模式训练、模式分类和判决三个步骤。常用的技术有动态时间归正、隐马尔可夫理论和矢量量化技术三种。

(3) 利用神经网络的方法是20世纪80年代末期提出的一种新的语音识别方法。神经网络本质上是一个自适应非线性动力学系统,模拟了人类神经活动的原理,具有自适应性、并行性、鲁棒性、容错性和学习特性,其强大的分类能力和输入输出映射能力在语音识别中都很有吸引力。但由于存在训练、识别时间太长的缺点,目前仍处于实验探索阶段。

2. 语音合成技术

语音合成技术是将计算机自己产生的或外部输入的文字信息(如文本文件内容、Word文件内容等),按语音处理规则转换成语音信号输出,使计算机流利地读出文字信息,使人们通过“听”就可以明白信息的内容。也就是说,使计算机具有了“说”的能力,能够将信息“读”给人类听。这种将文字转换成语音的技术称之为文语转换技术(Text To Speech, TTS),也称为语音合成技术。

一个典型的语言合成系统可以分为文本分析、韵律建模和语音合成三大模块。主要功能是根据韵律建模的结果,从原始语音库中取出相应的语言基元,然后利用特定的语音合成技术对语音基元进行韵律特性的调整和修改,最终合成出符合要求的语音。

常用的语音合成方法,按照合成方法分类,分为参数合成法、基音同步叠加法和基于数据库的语音合成法。参数合成法是通过调整合成器参数实现语音合成的。基音同步叠加法是通过时域波形拼接实现语音合成的。基于数据库的语音合成法是采用预先录制语音单元并保存在数据库中,再从数据库中选择并拼接出各种语音内容来实现语音合成的。

按照技术方式分类,分为波形编辑合成、参数分析合成以及规则合成三种。波形编辑合成是将语句、短语、词或音节作为合成单元,这些单元被分别录音后进行压缩编码,组成一个语音库。重放时,取出相应单元的波形数据,串接或编辑在一起,经解码还原出语音。这种合成方式也称为录音编辑合成。参数分析合成是以音节、半音节或音素为合成单元。按照语音理论,对所有合成单元的语音进行分析,提取有关语音参数,这些参数经编码后组成一个合成语音库。输出时,根据待合成的语音信息,从语音库中取出相应的合成参数,经编辑和连接,顺序送入语音合成器。在合成器中,通过合成参数的控制,将语音波形重新还原出来。规则合成存储的是较小的语音单位,如音素、双音素、半音节或音节的声学参数,以及由音素组成音节,再由音节组成词或句子的各种规则;当输入字母符号时,合成系统利用规则

自动地将它们转换成连续的语音波形。

本章小结

本章主要介绍了虚拟现实系统的相关技术,主要包括计算机图形学的基本原理、建模技术、绘制技术、人机交互技术和虚拟三维声音等。叙述了各种技术的主要概念、特性和基本原理,由于篇幅的限制,这里对这些技术只是进行了简要介绍,虚拟现实技术是集各种技术之大成者,在虚拟现实系统的开发中,需要深入了解更多的技术详情,还需要查阅更多的资料。

练习与思考

一、填空题

1. VR 系统的相关技术主要包括_____、_____、_____和_____等。
2. 虚拟环境中常用的坐标系分为_____、_____、_____和_____。
3. 环境建模主要涉及_____、_____和_____等。
4. 实时动态绘制技术可分为_____和_____两种绘制技术。

二、简答题

1. 简述三维图形绘制的基本原理。
2. 评价虚拟现实建模的技术指标包括哪些?
3. 自然交互技术主要包括哪些内容?
4. 三维虚拟声音的主要特征有哪些?
5. 简述碰撞检测的原理及主要方法。