

# 《虚拟现实：理论、技术、开发与应用》

## notes

---

collected by wxb

## Chapter 1

---

### 1.1 Virtual Reality Definition

以计算机技术为核心，通过将虚拟信息构建、叠加，再融合于现实环境或虚拟空间，从而形成交互式场景的综合计算平台，这便是“泛虚拟现实技术”的核心。具体来说，就是建立一个包含实时信息、三维静态图像或者运动物体的完全仿真的虚拟空间，虚拟空间的一切元素按照一定的规则与用户进行。

### 1.2 Three Core Elements

#### Immersion 沉浸性

虚拟环境的真实程度

#### Interaction 交互性

手势、动作、表情、语音、眼球、脑电波

#### Imagination 多感知性

视觉、听觉、触觉、动觉

### 1.3 Relative Techs

#### Computer Simulation

描述性技术，定量分析方法

**Definition:** 通过建立某一过程或某一系统的模式来描述该过程或系统，然后用一系列有目的、有条件的计算机仿真实验来刻画系统的特征，从而得出数量指标，为决策者提供关于这一过程或系统的定量分析结果，作为决策的理论依据。

#### Computer Graphics

计算机图形学是一种使用数学算法将二维或三维图形转化为计算机显示器的栅格形式的科学。

#### AI 人工智能

人工智能能够创造接受感知的事物，而VR是创造被感知的环境。人工智能的事物可以在VR环境中进行模拟和训练。

#### 5G

在VR/AR技术中，语音识别、视线跟踪、手势感应等都需要低时延处理，这也同时要求网络时延必须足够低。所以，超高速、超低时延的5G网络就为VR走进人们的日常生活铺平了道路。

## Big Data

VR使未来的数据分析变成交互式的，大数据将变为沉浸式。

## Chapter 2

### 2.1 Tech Structure of VR



### 2.2 实物虚化

实物虚化是在虚拟世界中描绘现实世界中的事物的过程。

#### 2.2.1 几何造型建模

- 形状：多边形、三角形、顶点等
- 外观：表面纹理、颜色、光照系数

用于存储虚拟环境中几何模型的模型文件需要包含几何信息的数据结构、相关的构造与操纵该数据结构的信息。

#### 人工几何建模

- 通过图像编程工具或虚拟现实建模软件建模
  - OpenGL, Java3D, VRML
- 利用交互式的绘图、建模工具进行建模
  - AutoCAD, 3ds Max, Maya, Autodesk 123D

#### 数字化自动建模

三维扫描仪。

#### 2.2.2 物理行为建模

##### 1. 物理建模

物理建模是对虚拟环境中的物体的质量、惯性、表面纹理（光滑或粗糙）、硬度、变形模式（弹性或可塑性）等物体属性特征的建模。

结合计算机图形学、物理学知识，关注力反馈问题。

#### 分形技术

自相似结构常用于对拥有不规则的复杂外形的物体的建模

- 优点：用简单的操作就可以完成复杂的不规则物体建模
- 缺点：计算量较大，难以满足实时性需求，在虚拟现实系统中一般仅用于静态远景建模

## 粒子系统

粒子系统主要由大量称为粒子的简单体素构成，每个粒子具有位置、速度、颜色和生命期等属性，这些属性可根据动力学计算和随机过程得到。

常用粒子系统建模制作的效果有火、爆炸、烟、水流、火花、落叶、云、雾、雪、尘、流星尾迹等。

## 2. 行为建模

让物体的运动和行为模式符合客观规律。

为满足虚拟现实的自主性，除了对用户行为直接做出反应的行为模型以外，还需要考虑与用户输入无关的行为模型。

## 2.3 虚物实化

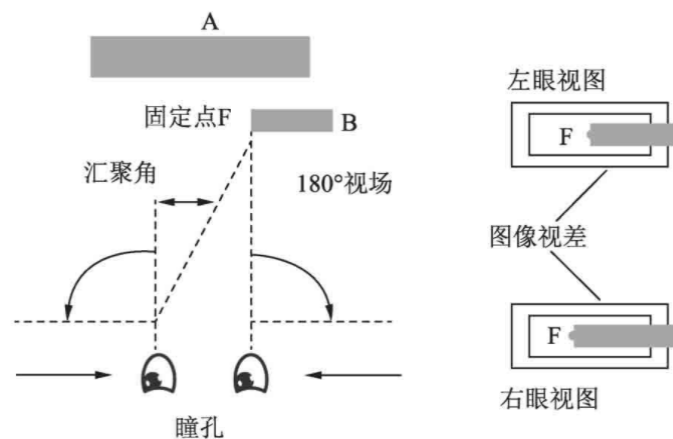
虚物实化则是将建模好的虚拟场景呈现给用户的过程。

### 2.3.1 视觉绘制

#### 人类视觉系统

Field Of View(FOV): 眼睛能观察到的范围称为视场

立体成像原理：内瞳距 -> 图像视差



单眼观察：运动视差，深度感知

深度感知：近场，图像视差；远场，固有线索如线性透视、阴影、遮挡等

#### 立体显示

- 基本思路：产生两幅轻微位移的图像输送到两只眼睛
- 技术差异：如何使得两只眼睛在看同一个画面时接收到不同的图像

#### 分色技术

- 原理：让某些颜色的光只进入左眼，另一部分只进入右眼
- 操作：先将要立体显示的画面分成有细微偏移的两个图像，偏左边的画面去除绿色或蓝色，偏右边的画面则去除红色，再将两者叠加显示在屏幕上
- 例：红蓝3D眼镜 左边镜片为红色，右边镜片为绿色或蓝色

## 分光技术

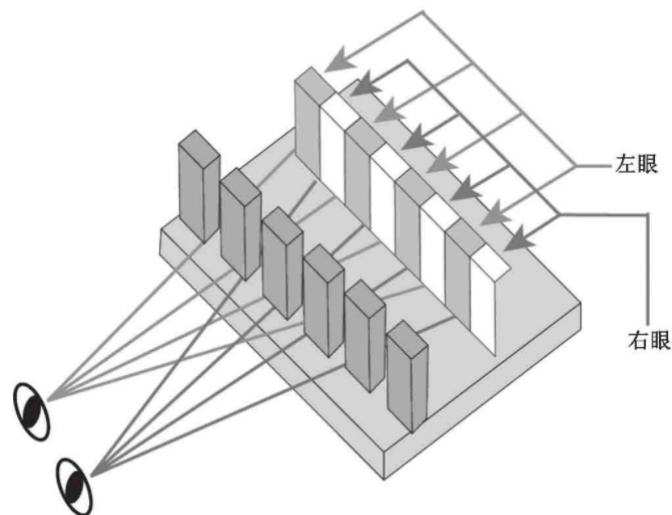
- 原理：让两只眼睛分别接收到偏振方向不同的光。偏光眼镜的左、右两片镜片都为偏光镜，它可以滤过特定偏振方向的光，而将其他的光阻隔。
- 操作：将要展示给两眼的画面偏振方向互相垂直，叠加到屏幕上
- 例：偏光眼镜

## 分时技术

- 原理：将两个有偏移的画面在不同的时间播放。在画面刷新和眼睛遮挡频率很快的情况下，根据人眼视觉暂留的特性就能合成连续的画面。
- 操作：在画面第一次刷新时播放左眼的画面，并遮住右眼，下一次刷新时播放右眼的画面，遮住左眼
- 例：液晶快门眼镜

## 光栅技术

- 原理：在显示器前端加上光栅，让左眼透过光栅时只能看到部分画面，右眼也只能看到另外一半画面



## 真实感实时绘制

### 真实感绘制技术

- 真实感：几何真实感、光照真实感（光源照射下的明暗变化或镜面反射）、行为真实感
- 任务：模拟真实物体的物理属性，即物体的形状、光学性质、表面纹理和粗糙程度，以及物体间的相对位置、遮挡关系等
- 常用技术
  - 纹理映射：提升虚拟物体表面细节处的真实度
    - 主要方式：在虚拟物体的几何表面贴上细腻的纹理图像
  - 环境映射：实现虚拟物体在光照等的作用下拥有的物体表面镜面反射、明暗变化和规则透视等与真实世界相一致的效果
    - 主要方式：用纹理图像贴于物体表面的方式来完成
  - 反走样技术：用提高图像像素密度的方式来减少图像失真的方法
    - 主要方法：以两倍的分辨率绘制图形，再由像素值的平均值计算正常分辨率的图形；或计算每个相邻元素对一个像素点的影响，再将其加权求和得到最终像素值

## 实时性绘制技术

1. 画面更新速度快，人眼察觉不出画面闪烁
2. 对虚拟世界中物体的姿态和位置进行实时计算并动态进行绘制
3. 对用户的输入及时作出相应的反应

流程：先进行几何外形和轮廓的绘制，然后利用纹理映射、环境映射来渲染真实感，最后进行实时的画面输出

提升绘制效果

- 提高图形显示能力（计算机运行速度）
- 降低场景复杂度
  - 预测计算：根据各种运动的方向、速率和加速度等运动规律，在下一帧画面绘制之前用预测、外推的方法推算出手部跟踪系统及其他设备的输入，从而减少输入设备产生的延迟。
  - 脱机计算：在实际应用中尽可能将一些可预先计算的数据进行预先计算并存储在系统中，运行时直接调用。
  - 3D剪切：针对可视空间剪切，将一个复杂的场景划分成若干子场景。通过剔除虚拟环境在可视空间以外的部分，有效地减少在某一时刻所需要显示的多边形数目，从而降低场景的复杂度，减少计算量。
  - 可见消隐：系统仅显示用户当前能“看见”的场景，当用户仅能看到整个场景很小部分时，系统仅显示相应场景，从而大大减少所需显示的多边形的数目。
  - 细节层次模型Level Of Detail (LOD)：对场景中不同的物体或物体的不同部分，采用不同的细节描述方法。在适当条件下显示细节简化后的模型，减少需要显示的多边形数目。

### 2.3.2 并行绘制

同时对多个图形，或者同一图形的不同部分进行绘制，充分利用计算资源，避免计算资源闲置。

#### 图形绘制流水线

##### 1. 应用程序阶段

- CPU/GPU
- 任务：建模、加速计算、动画、人机交互响应用户输入（鼠标、追踪器、数据手套）

##### 2. 几何处理阶段

- Geometry Engine, GE
- 任务：从三维坐标变换为二维屏幕坐标的过程，包括模型变换（坐标变换、平移、旋转和缩放等）、光照计算、场景投影、剪裁和映射

##### 3. 光栅化阶段

- 硬件实现，光栅化单元 Rasterizer Units, RU
- 任务：把几何处理阶段输出的几何图形信息（坐标变换后加上了颜色和纹理等属性）转换成视频显示器需要的像素信息，即把几何场景转化为图像。
- 重要功能：反走样

## 并行绘制方法

### 流水线并行

- 优点：最常见、最易实现；不同的阶段可能会占用不同的软件、硬件资源
- 缺点：后一阶段需要前一阶段输出的结果来作为输入才能开始执行；流水线整体的速度将受限于最慢的那个阶段所需的时间，并且划分过多的阶段也会影响并行效果

## 数据并行

- 操作：数据会被划分成子数据流，在一些相同的处理模块上对这些子数据流进行处理
- 优点：绘制流水线阶段数不会影响并行的效果；在数据相关性较弱的绘制过程中能达到较高的并行度，效果明显，并且数据并行方法具有很好的扩展性，常被用来构建大规模的并行绘制系统
- 缺点：受制于系统中的通信带宽，以及相同处理模块的数目

## 作业并行

- 主要应用在流水线中有独立分支的情况
- 瓶颈：流水线中独立分支的数量、独立模块之间的差异性

## 并行图形绘制系统的实现

### 基于高端多处理器和高性能图形工作站

传统；价格高昂

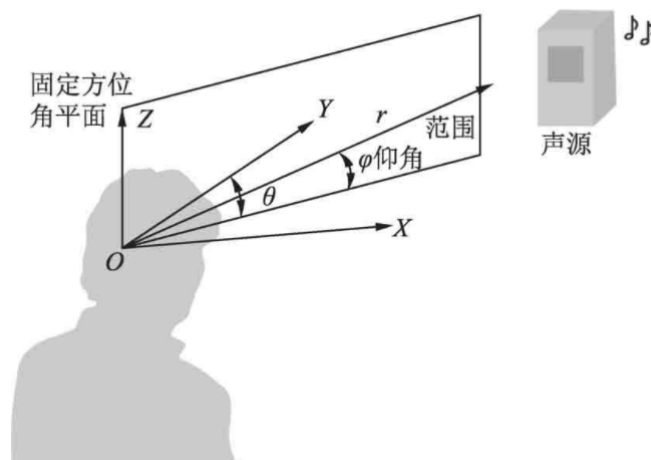
### 基于PC集群

- 新趋势，价格低廉，性能客观
- 例：美国Alamos国家实验室海洋模拟数据；中国基于VTK，人体头部CT扫描

## 2.3.3 声音渲染

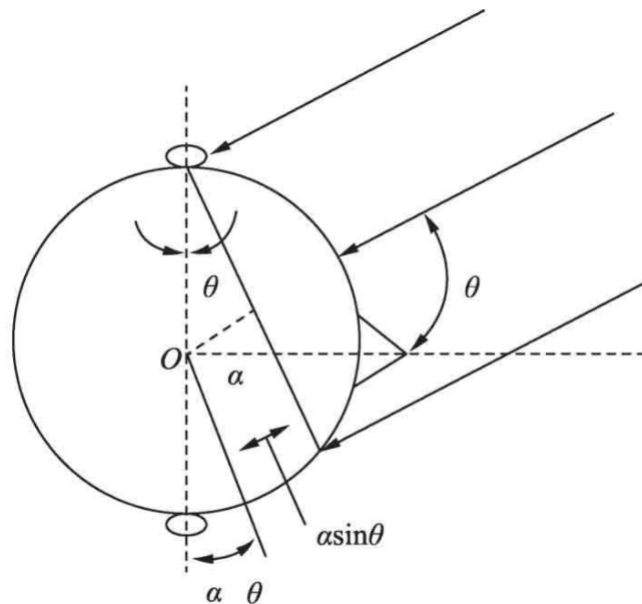
### 人类听觉系统

大脑根据声音的强度、频率和时间线索判断方位角 $\theta$ 、仰角 $\phi$ 和范围 $r$ ，从而确定声源位置。



### 方位角线索

两耳时差 Interaural Time Difference:  $ITD = \frac{a}{c}(a\theta + a\sin\theta)$



两耳强度差 Interaural Intensity Difference, IID

### 仰角线索

声音经外耳反射后进入内耳，来自用户前方的声音与头顶的声音有不同的反射路径。声音在被外耳反射时，一些频率被放大，另一些被削弱。声音和耳郭反射声音之间的路径随仰角而变化，大脑则通过感受不同频率声音被放大或削弱的程度来推断声源的仰角。

### 距离线索

大脑利用对给定声源的经验知识和感觉到的声音响度估计声源的距离。

- 平移头部时，声音方位角变化越大，距离越近
- 声源的声音与经周围环境第一次反射后的声音强度之比。声音的能量以距离的平方衰减，反射的声音不会随距离变化发生太大变化

### 头部相关传输函数 HRTF

头部相关脉冲响应 HRIR -> HRTF (傅里叶变换)

<https://www.cnblogs.com/h2zZhou/p/8405717.html>)

### 三维虚拟声音技术

#### 三维虚拟声音

- 立体声技术：将采集到的立体声经过一定处理，最后重放的时候还能够恢复一定程度的立体感。最常见的方式就是采用多扬声器的方式来营造声音从四面八方传来的空间立体感。
- 立体声来自听者面前的某个平面，而三维虚拟声音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、后方或前方。

### 原理

1. 双耳效应
2. 耳廓效应
3. 人耳的频率滤波效应。人耳的声音定位机制与声音频率有关，对20 ~ 200Hz的低音通过相位差定位，对300~4000Hz的中音通过声强差定位，对高音则通过时间差定位。
4. 头部相关传输函数

## 2.3.4 力触觉渲染

### 人类的触觉系统

#### 触觉感受器

人体的皮肤一般通过分辨触点的位置或接触的时间来区分多次接触。皮肤对触点位置感知的精细度取决于皮肤中传感器的密度。

#### 本体感受器

神经末梢位于骨骼关节中，感受器放电的振幅是关节位置的函数，它的频率对应于关节的速度。身体定位四肢的精确性取决于本体感受的分辨率，即能检测出的关节位置的最小变化。

#### 温度传感器

### 力触觉反馈

#### 接触反馈 Touch Feedback

用户在虚拟环境中能够感知到虚拟对象接触表面的几何外形、纹理、硬度、光滑度和温度等物理属性信息

#### 力反馈 Force Feedback

在与虚拟环境进行交互的过程中，用户产生一定的输入到虚拟环境中，例如用手抓握手柄等操作，虚拟现实系统会给出一定的力的作用的反馈

## 2.4 其他技术

坐标系名称	坐标系描述
世界坐标系	由于摄像机存在于真实世界中，因此需要使用一个基准坐标系来表示它和空间的任意点在真实世界中的位置，这个坐标系称为世界坐标系
图像坐标系	原点是光轴与成像平面的交点， $X$ 轴与 $Y$ 轴分别与摄像机坐标系的 $X$ 、 $Y$ 轴重合
摄像机坐标系	原点位于光学中心， $Z$ 轴与光轴重合

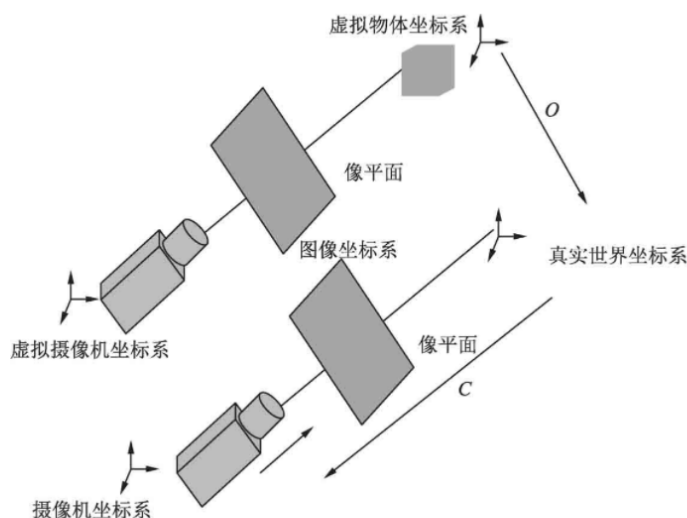


图2-21 三维注册坐标系的关系



表2-2 3种注册技术的比较

注册技术	原 理	优 点	缺 点
基于硬件跟踪设备的注册技术	根据信号发射源和感知获取的数据求出物体相对空间的位置和方向	系统延迟小	设备昂贵，对外部传感器的校准比较准，且受设备和移动空间的限制，系统安装不方便
基于视觉跟踪的注册技术	根据真实场景图像反求出观察者的运动轨迹,从而确定虚拟信息“对齐”的位置和方向	无须特殊硬件设备，注册精度高	计算复杂性高，造成系统延迟大；大多数都用非线性迭代，误差难控制，鲁棒性不强
基于混合跟踪的注册技术	根据硬件设备定位用户头部运动的位置和姿态，同时借助视觉安抚对配准结果进行误差补偿	算法鲁棒性强，定标精度高	系统成本高,系统安装烦琐，移植困难

标定技术就是确定摄像机的光学参数、集合参数、摄像机相对于世界坐标系的方位以及摄像机与世界坐标系的坐标转换。

计算机视觉技术通过对三维空间中目标物体几何信息的计算实现识别与重建，从而让摄像机获取真实场景中的图像信息。