

硕 士 研 究 生 读 书 报 告



题目 游戏中的多来源声音渲染

作者姓名 其佈尔

作者学号 21651057

指导教师 李启雷

学科专业 软件工程

所在学院 软件学院

提交日期 二○一七年4月

Rendering for Large Multi-Source Scenes

A Dissertation Submitted to

Zhejiang University

in partial fulfillment of the requirements for

the degree of

Master of Engineering

Major Subject: Software Engineering

Advisor: Li Qilei

By

Qibuer

Zhejiang University, P.R. China

2017

摘要

在虚拟现实和分布交互仿真应用中，为增加参与者或观察者的沉浸感，生成多维的感官信息(如视觉、听觉、触觉以及动感等信息)更能符合人的生理和心理感知。本文论述了虚拟环境中的声音渲染技术。首先从声音的传播和人的感知角度，论述了声音渲染的数学模型，然后论述了其研究内容和实现的关键技术，如声音的建模分析、三维空间定位、距离感知、与图象信息的融合以及声音渲染的验证等。文章最后给出了用于虚拟座舱的一种声音渲染方案。

**关键词：**虚拟现实；声音渲染技术；声音空间化；头部相关传递函数；空间立体声

Abstract

The immersion of the participator or viewer in virtual environment can be enhanced greatly if their visual, auditory, tactile and other sensory organs are stimulated simultaneously .This paper addresses issues of the sound rendering technology providing the auditory cue , particularly the real-time digital acoustic synthesis for reproduction via headphone .Based on the sound' s propagation and the human' s psychoacoustics, a mathematical model has been made to explain the principle of real-time three-dimensional sound synthesis and manipulation .The analysis and modeling of the sound is the first step for synthesis, an interactive , real-time ability to spatialize the sound and simulate the distance is essential , the fusion between the audio and other (primarily visual)modality contexts can enhance the immersion and imagination .The paper includes an application example in a form of the real time digital acoustic synthesis for the virtual cockpit. The off-line work and the real time pipeline procedure are described.

**Keywords：**acoustic effects; sound rendering; virtual reality

# 1 引言：

游戏场景以及交互虚拟环境中的几何和视觉复杂度在过去几年里迅速增加，近几年视觉渲染和硬件技术的进步使得市场上的GPU能够在一定互动比下生成好质量的画面，这种进步也使得人们更关注其他领域，比如声音渲染，让虚拟环境更加真实和融合。但是，互动比下复杂场景中真实声音效果的实现仍然存在挑战。场景的高听觉复杂度主要由以下几个因素影响：

声音来源数量庞大：场景可以有很多声音来源，从几百种到几千种。这些声音来源与路上的车，商场或体育场的人群或工厂里机器的噪声相似。

物品数量庞大：场景大部分由成百的静态和动态物品组成，这些物品与建筑模型，长几十米或几百米的室外场景相似。

声音效果：模拟各种声音效果也十分重要，比如早期的反射，之后的回声、衍射、扩散等。

高听觉复杂度导致在声音传播以及声音渲染的难度增加。广义上说，声音传播方法可以分为基于波和几何学技术。基于波的方法在数字上解决了声音波方程，能够精确模拟所有的声音效果。但是，这种方式局限于直邮几件物品并且没有大量声音来源静态场景。几何学传播技术基于光波理论，能够用来计算早期反射和只有几种声音来源的动态场景中的衍射。

最重要的挑战是在动态场景互动比下模拟后期回声。这种回声与在多次反射并伴有多次延迟后到达听者的声音和脉冲回应的尾部相似。从感觉上来说，这种回声能够告诉你周围环境的大小及其对声音的吸收率。很多真实世界的场景中，包括音乐厅、森林、城市的街道或者山脊，都各有各的回声。但是这种真实的声音元素-回声，计算代价非常高。在一个典型的房间大小的环境中使用光波追踪，只计算1到2秒的回声长度就需要中等大小房间声音反射的计算量。

# 2 带通分频：

为了更好地模拟声音的传播，原始信号被分为高、中、低三个频段，具体实现可借助于IIR方法。这种方法的原则就是利用输入和已有的输出来构造新的输出。各常参数可以方便地用IIR的构造方法求得，由此可以看出，IIR方法的速度是很快的。因为它是线性的，而且使用简便，其效果也是比较令人满意的。

三维场景的建造和浏览

本系统是面向网络应用的，其可以使两个以上的用户在同一个虚拟场景中进行三维漫游，同时可以对话，也可以看到对方。其中程序分配是：有一个服务器端程序，它不停侦听来自各个客户端的消息，这些消息包括用户的动作以及声音信息，然后把这些动作发送给其他用户，以实现多用户的动作同步和语音同步；有多个相同地位对等的客户端程序，它们负责接收用户的动作信息和语音信息。

在客户端使用的是纯Java程序包JavaGL，由于它的执行效率比较高，因此对于一个不是很复杂的三维场景渲染基本能够达到实时。客户端程序是通过分析接收到的键盘按键信息来控制三维场景的视角和位移变换，以达到三维漫游的目的。

每次漫游时，一旦有一个用户有动作，那么客户端程序就会自动把这些情况实例打成上述对象包裹，发往服务器端程序；然后服务器端程序就会对这样的包裹进行分析，并作出相应的决定，如果有必要，就把这个包裹转发给其他用户。比如用户发生了一次位移动作，用户的新位置坐标将被添入上述对象中，同时类型信息变为“300”，表示该包裹传送的是三维信息，然后发往服务器；服务器接收到包裹之后，如果发现该包裹应该广播到其他所有用户即进行传送，于是其他用户都接到了同样的信息，它们分别在自己的三维场景中，把刚才那个用户的动作表现出来。如此一来，多个用户就可以真正同步的共享同一个三维场景。其他的信息也是通过这样的方式交流的。

# 3 声音传播过程的模拟

上述所做的工作就是把声音原始信号分成高、中、低3个频段，进而按能量排序分成不同能量的各个分量，然后针对每个频段自身特点分别传送，最后把高、中、低的各个分量在听着的位置进行叠加，最终得到听者应该听到的声音信号，用若干根声线表示声音相应的高、中、低频能量分量，每条声线也有自身的数据结构。

对于各个成分的分配，在现代声学中也没有相应的精确定义，因此需要构造一个模型从能量分布形状上去逼近。图像按照极坐标的方式画出来，此图像的形状比较像在人头部上方俯视的中频发射图像，其中每一条声线的方向代表相应声音分量的方向，而每一条声线的长度则代表相应声音分量的大小。

具体传播的计算过程如下：

低频部分：由于可以认为低频声音除了被墙壁吸收之外，还可以毫无阻挡地透过墙壁，于是，低频成分的传播变得简单化了，即只需要在声源到听者的射线上，考虑能量的递减过程即可。该能量事按照距离平方的倒数递减的。

高频部分：可认为高频声音除了被墙壁吸收之外，其余部分均在墙面之间进行镜面反射，直到声音到达听者或者衰减到足够小的时候才停止反射的计算过程。

中频部分：它是3个部分中最复杂的。由于中频声音遇到墙壁时，能量分为被墙壁吸收、被墙壁反射、透过墙壁继续传播3部分，所以可干脆把中频部分除了被墙壁吸收以外的能量再分为两部分，即反射分量和透射分量。然后分开来做类似于低频部分和高频部分的处理。

# 4 总结：

每个系统时采用B-S结构框架，服务器端程序使用Java语言开发的Windows侦听程序，浏览器端是在网页中用Java语言开发的Applet程序，这样就充分利用了Java语言各方面的优势。

在实时性方面，基本可以达到要求。因三维场景的实时渲染可以达到18帧/s以上，而声音的网络传播和渲染每次也可以通过调整声线数量控制在0.07s之内，所以在感官上，整个系统可以保证用户有实时的体验。

参考文献

[1] Interactive Sound Propagation and Rendering for Large Multi-Source Scenes. Carl Schissler, Dinesh Manocha. 2017.

[2] 声音渲染技术及其在虚拟环境下的应用. 罗福元, 王行仁, 彭晓源. 2015.