

请各所（系、中心）根据学生人数和教师数量，组成开题评审委员会。开题评审委员会由副高职称（含）以上的教师任主任委员，讲师（含）以上担任委员，人数不少于3人。开题应重点审查题目是否可以构成复杂系统，是否可以支撑毕业要求指标点。

开题报告模板附后。

北京理工大学

本科生毕业设计（论文）开题报告

学 院： 计算机学院

专 业： 计算机科学与技术

班 级： 07111506

姓 名： 王铎暉

指导教师： 刘庆晖

校外指导教师： 刘天添

二〇一九年一月三日

1. 毕业设计（论文）选题的内容

如今，从视觉上进行的物体/交互模拟已经比较成熟，甚至已经延伸至液体，软物体如布料，沙子，熔浆等的模拟，但从听觉的角度上，适配当今交互系统的声音模拟还比较欠缺，主要都采取的是对已记录声音的再处理来进行声音模拟，对物体本身材质，交互时的物理信息(如速度，碰撞点，形状，碰撞力等)考虑较少，以及缺乏从信号处理方面考虑声音的最基本生成属性，如增益，频率等。本课题主要研究基于物理模型的实时声音模拟，从听觉感知的角度对场景应用中的物体交互提供实时，准确，以及更真实的声音。本课题主要研究目标是针对以碰撞和滚动为代表的冲击声音和持续声音实时模拟。强调实时的重要性首先来源于当今游戏和实时模拟项目，特别是VR，AR 相关领域中，要求模拟音效必须实时配合画面，但同时分给声音模拟的运行时间和可调用资源也是有限的（给声音模拟调用的存储资源往往小于 25MB，对碰撞声音的资源调配往往小于 2MB^[4]），所以我们需要比较快速的，基于物理模型的模拟，来满足在实时仿真环境中比较好的沉浸式体验与真实感。同时，建立合理的声音模型，是从较底层解决控制声音各种参数（频率，音量，声音增益等）的高效方案。

2. 研究方案

2.1 本选题的主要任务

根据声音建模与合成，环境音效添加的划分，本课题题的主要任务分为以下三个部分：

(1) 因为声音来源于振动，而振动产生的各种波形的叠加就是声音。所以根据弹簧质点阻尼模型，首先进行振动的模拟，来预先得到相应的波形，利用的是傅里叶变换/拉普拉斯变换^[2,3]建立弹簧阻尼模型偏移量与频率之间的关系。

(2) 利用声波传递函数^[2,3]对声音的传播进行模拟，从而搭建由振动波形到被听者听到之间声音传递的考量。包括了对振动物体的材质，形状，以及声源等参数的考虑；或继续利用弹簧质点阻尼模型^[1,5]，将碰撞时的力量转变为波形的增益混入之前震动模拟得到的波形来模拟声音，利用杨氏模量控制振动物体的材质等信息。

(3) 考虑环境音效的模拟/添加^[8]，如声音扩散效果，声音混响效果，声音的空间效果，音量与听者位置信息的变换等。

在完成声音模拟后，重点在于对各种材质，各种物体的测试与优化。从最开始的球状物体之间的碰撞，延申至马克杯，桶，地面，桌子等之间的碰撞声音场景模拟。最后再在现有基础上，进行滚动声音这种连续性声音的模拟，重点要考虑的是物体运动速度和离听者的距离之间的信息。因为强调的是实时模拟，所以我们同样需要完成对声音模拟整体性能的分析，比如占用 CPU 的情况，占用内存情况，以及 FPS 等^[1,2,4]。

2.2 技术方案的分析、选择

(1) 技术方案的分析

对于计算机图形学中的模拟与仿真，首先来自于对各种物理模型实时控制的可视化需求以及场景模拟，方案设计的数字化需求。人们可以通过建模，参数化等方式，将真实世界的各种现象变成基于数字，基于计算的控制。因此，许多物理模型被提出来解决，优化模拟与仿真中遇到的一系列问题。从最基本的弹簧阻尼模型^[1,2,3]入手，一维的弹簧阻尼模型是比较容易建立起相应的形变量与力的关系的，但是模拟强调时间，当引入了时间的变量，仅仅是力与偏移量是无法解决任意时刻弹簧的变化以及未来时间弹簧阻尼模型变化的趋势的，因此引入了时间卷积/积分^[2,3]的方法，这是基于数学与物理的准确模型求解。但同时我们知道，想要解决这样的积分问题，得到准确解是不可能的，只能依靠设置单位时间的采样大小，利用离散化的方法，迭代的方法进行不断逼近和减小误差。所以在计算机模拟领域中，积分与迭代的求解成为了优化模拟的最终效果和提升模拟速度的重点方向与基本方法。对于声音模拟，更需要重点考虑的是频率的问题，因为声音的产生来源于多个不同频率，不同振幅的正弦波^[8]叠加，它所涉及到的最核心的一个数学理论是欧拉公式，它通过将三角函数与复指数函数^[2,3]相关联，将声音模拟从数字信号处理的角度，拉到了物理建模，数学计算的角度。对声音模拟进行物理模拟的重要意义在于，如果对于场景模拟中对所需要的声音都采用真实世界采样的方法是很难满足对各种声音的需求的。比如，对于一个简单的马克杯落地的声音场景模拟而言，根据地板和马克杯材质和撞击位置的不同会发出各种不同的声音：木制马克杯和陶瓷马克杯，有地毯或木制地板，马克杯底和杯柄着地发出的声音都是非常不同的，如果全部采用采样的方法将是非常巨量的工作。基本的计算机音乐知识已经告诉我们，声音

的产生来源于正弦波的合成，且大量是谐音(Harmonic Sound)的处理^[8]，因此如果能够将声音成功的抽离并建立对应的波形模型等，将能够数字化，快速的进行声音的合成与模拟，比如对于木制和陶瓷马克杯的声音，只需要改变对于波形的频率即可，或者在更复杂的建模基础上，改变一些衰减系数^[8]就能够完成了，这样的技术方案将极大的丰富场景模拟中被使用声音的真实性与多样性。

对于声音模拟，主要考虑两个方面：声音生成^[2,6](Sound Generation)和声音传播^[3,7](Sound Propagation)。声音生成在上节已经讨论过，来源于物体表面的振动，声音的传播则需要考虑的是声音的频率和听者位置的问题。首先对于振动的建模基于物体的网格对象进行建立，考虑三角网格物体的各个顶点对外界作用力的振动模型进行对于波形频率的求解。然后再结合经典的声音传播公式，将各个波形在不同位置的情况进行最终能听到的声音的合成。而本课题就是针对上述的技术方案进行研究并进行优化来模拟更高质量的，实时的，可从底层较易改变的声音，然后希望最后能够将整个模拟的结果放在 Unity3D，甚至是通过插件的形势放入 Unity3D 然后进行比较复杂的场景模拟(如 AR, VR)来进行最终的模拟声音效果和性能的检验。

(2) 技术方案的选择

在技术方案的选择方面，本课题最核心的问题其实是在实时上，所以上述的模拟方案，最终都是需要向实时的要求上进行妥协与平衡，比如我们可能会牺牲一些不重要的波形，来提升声音模拟的速度。具体的技术方案主要有两种：

1. 从振动模型入手的声音模拟及加速方案

首先利用弹簧质点阻尼模型建立基本的振动模型，得到针对不同网格下，不同材质的物体的不同位置的声音波形模式(模式的数量往往有几千个)。为了达到实时的要求(至少 30fps)，通过合理的方案对最终合成声音的波形模式数量进行删减和控制，使得最终实时合成的声音只有几百个波形模式，从而加速声音模拟。主要的加速方案有两种：模式压缩^[1](Mode Compression)和模式截断^[1](Mode Truncation)，其中模式压缩是从听觉感知角度，将模式进行合并。因为人的听觉感知只能基本在低于 2KHz 的波形模式^[1]下，分清 1Hz 的模式区别，所以在大于 2KHz 的情况下，可以将一些很多相近的频率用一个中间频率进行代表，

从而压缩了模式数量。模式截断是停止多余的波形模式混合与叠加，在它对声音的贡献小于一定数值时。这样的数值主要是针对波形的振幅^[6]来确定，因为振幅很小，证明它对最终声音质量影响也很小，所以可以在振动发出一定时间后（小于 0.2s）就不考虑它对合成声音的作用，从而减少模式的数量。

同时，在场景模拟中，往往不是一个声音在场景中发出声音，往往是大量的，复杂的碰撞，滚动声音等构成了整个场景中的总体声音情形。这时候同样需要对计算资源进行合理调配来完成声音模拟的最大配置，因为从听觉感知的角度，听者总是对正面面对的物体给予了最大的关注度^[1,4]，所以分配给听者直面的物体最大的计算资源模拟出质量最高的声音，同时对在背面的物体基于较少的计算资源模拟出质量较差的声音，对听者在模拟的场景中能够感受的总体声音影响是比较小的，但却能够明显的加速声音模拟。

2. 从已记录的声音入手的声音模拟及加速方案

对于已记录的声音进行建模，是一种基于数字信号处理的一种快速声音模拟的手段，它通过提取已记录声音中能展现最终声音的特征型波形，并建立合理的衰减手段（如指数型衰减模式^[8]）来对声音重新建模来满足多种材质/场景的需求，同时也能够从振动波形的基本角度上对声音进行处理。这种声音模拟方式对声音的采样有一定的依赖，但是一种比较快速的实时声音模拟方案，它需要重点解决的问题侧重于比直接使用已记录的声音占用内存更少，以及能够使得已记录声音也能产生更真实的变化，而不是失真。对于场景中多种发生物体的声音模拟，也可以同样采用优先队列^[4]的方式，给听者直接朝向的物体给予较多的资源进行计算模拟。

2.3 实施技术方案所需的条件

本课题涉及的声音模拟主要在于将预先推导的时间积分公式进行编程实现，已算的预处理，以及将模型结果用于可视化的 Demo 中，在声音模拟阶段对计算机硬件要求不高，对硬件要求较高的阶段在于后期如果需要引入 VR 等大型场景模拟测试，需要较高配置的显卡条件，具体所需软硬件条件如下。

本课题研究的问题所需要的相关软件条件：

- (1) 编程语言：C++/ CHUCK/ C#
- (2) 图形化编程接口：OpenGL
- (3) 测试开发引擎：Unity3D, SteamVR
- (4) 开发系统环境：Ubuntu, Windows

本课题研究的问题所需要的相关硬件条件：

- (1) 当前主流 PC 一台：16GB 内存 + NVidia 1060 及以上
- (2) HTC Vive/Vive Pro 一套

2.4 存在的主要问题和关键技术

(1) 存在的主要问题

对于实时的声音模拟，目前存在的主要问题是对于一些比较厚的，饱满的物体(如：马克杯，塑料桶，篮球等)的声音模拟其实是比较好的,但对于一些比较薄的物体，比如盘子，弦等的声音模拟，受限于他们不太清晰的网格化结构，声音模拟的效果并不是很好^[2,3]，这是本课题想要解决的主要问题之一，同时对不规则的复杂物体的声音模拟可能需要探索更多的优化加速方案来达到实时的要求。当前较多的声音模拟比较多的是基于碰撞等的冲击声音的模拟，缺乏对滚动等持续性的声音模拟，它涉及到物体的持续性移动对声音的影响，以及产生的声音是否具有可重复性^[1,8]等问题。因此，本课题主要希望在完成碰撞声音的实时模拟基础上，能够完成基本的对滚动声音的模拟，如简单的金属小球在桌面滚动场景的声音模拟。

(2) 技术关键

当前的技术关键在于对于当前已有工作中以及优化后的的振动模型积分公式在保证较好声音质量的同时，达到实时模拟的要求。同时，从对已记录的声音也建立相应物理模型，完成一个声音库^[4,6]的积累，使得满足较大场景模拟的需要。最后，对场景中声音模拟的测试需要尽量满足多种物体，多种材质的需求，才能体现模型的有效性与普适性。涉及到的理论知识与编程方面内容较多，较复杂，涉及各种数学公式(如快速傅里叶变换^[2,3]，赫姆霍兹方程^[2,3]等)，迭代方法（时间卷积），矩阵处理(稀疏矩阵分解，力矩阵对角化^[1]等)的实现。对于

最后增加环境音以及空间音效等不是本课题的主要研究目标，因为有了基本的声音模拟后，可以直接利用现有的游戏引擎等环境音效插件直接添加上述效果来适配根据听者位置移动的场景模拟。

2.5 预期能够达到的研究目标

预期在完成基于物理模型的声音模拟建模基础上，从基本的规则物体(球类物体)撞击声音入手，延展至更多的不规则物体(马克杯，盘子，水桶等)的撞击声音模拟，最后，能够完成基本的持续滚动声音模拟。

最后提交的成果应为：将上述模拟情况处理成较好的 Demo 方便进行展示，并对模拟的性能进行分析，形成较完整的本科毕业设计报告。在时间充足的情况下，将现有的研究成果转换为 Unity3D 的辅助插件融入到 Unity3D 引擎中，进行更多的场景模拟测试。

3. 课题计划进度表

2019.1.7 - 2019.1.20	完成开题报告相关工作，进行相应文献阅读与方法收集工作,进行振动模型时间积分公式推导与理解，寻找现有相应的编程方案并测试。
2019.1.7 - 2019.2.24	完成实时声音模拟的相应课程内容，并针对现有方法和代码实现进行复现工作，掌握开题报告中描述理论的实现基本方法与流程。
2019.2.25 - 2019.3.7	开始刚体的碰撞声音模拟(从规则物体开始)，建立起基本的声音模拟物理模型，决定实时模拟加速方案。
2019.3.7 - 2019.3.21	完成刚体的碰撞声音模拟，并延展至不规则物体，决定 Demo 制作的场景模拟设计方案
2019.3.22 - 2019.4.7	准备并提交毕业设计中期检查相关材料，开始搭建 Demo 及尝试进行 Unity3D 插件的转换工作
2019.4.7 - 2019.4.21	首先在 OpenGL 的图形化界面下进行声音模拟的可视化场景操作，并进行对已记录声音的物理建模尝试
2019.4.22 - 2019.5.7	进行用户测试和相应性能分析，进行多种不规则物体的碰撞声音模拟，以及滚动声音模拟探索。
2019.5.5 - 2019.5.16	将环境音效与空间音效加入已完成的聲音模拟中，搭建一个比较好场景来展示整体声音模拟成果，最好能放入 VR 环境中。
2019.5.16 - 2019.5.22	完成所有 Demo 的搭建与优化，争取完成建立 Unity3D 插件工作，并开始建立自己的声音素材库。
2019.5.23 - 2019.6.11	筹备毕业设计答辩，完成毕业设计论文撰写，提交毕业设计各阶段资料，完成毕业设计整体工作。

4. 参考文献

- [1] Raghuvanshi, N., & Lin, M. C. (2006, March). Interactive sound synthesis for large scale environments. In *Proceedings of the 2006 symposium on Interactive 3D graphics and games*(pp. 101-108). ACM.
- [2] Li, D., Fei, Y., & Zheng, C. (2015). Interactive acoustic transfer approximation for modal sound. *ACM Transactions on Graphics (TOG)*, 35(1), 2.
- [3] Doug L. James, Timothy R. Langlois, Ravish Mehra, and Changxi Zheng. 2016. Physically based sound for computer animation and virtual environments. In *ACM SIGGRAPH 2016 Courses (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 22, 8 pages. DOI: <http://dx.doi.org/10.1145/2897826.2927375>
- [4] Lloyd, D. B., Raghuvanshi, N., & Govindaraju, N. K. (2011, February). Sound synthesis for impact sounds in video games. In *Symposium on Interactive 3D Graphics and Games* (pp. PAGE-7). ACM.
- [5] Ren, Z., Yeh, H., & Lin, M. C. (2013). Example-guided physically based modal sound synthesis. *ACM Transactions on Graphics (TOG)*, 32(1), 1.
- [6] Yamamoto, Kazuhiko, and Takeo Igarashi. "Interactive physically-based sound design of 3D model using material optimization." *Symposium on Computer Animation*. 2016.
- [7] Cao, Chunxiao, et al. "Interactive sound propagation with bidirectional path tracing." *ACM Transactions on Graphics (TOG)* 35.6 (2016): 180.
- [8] Cook, P. R. *Real sound synthesis for interactive applications*. AK Peters/CRC Press.