**摘 要**

在基于视觉的仿真与模拟研究被广泛应用于动画制作，虚拟环境搭建，人像建模，危险实验预模拟等领域的大环境下，研究针对其他感官系统 (如听觉，触觉，嗅觉等)的仿真与模拟成为利用计算能力探索搭建更真实虚拟世界的新趋势。得益于现今游戏引擎 (如Unity3D, Unreal, Blender等)和虚拟现实设备(如Oculus Rift, HTC Vive等)的普及，本工作提出了一种高效的基于物理仿真来生成针对刚性物体碰撞声音的方法，实时，且能够应用于虚拟现实环境中。我们通过利用弹簧质点阻尼模型来提取刚性物体的震动信息 (一般用正弦波表示)，并合理抓取如碰撞力，碰撞点的物理信息来进行声音合成。

然后，本工作基于根据已记录声音进行实时碰撞声音模拟的方法，在虚拟现实环境中进行了一个经典心理学实验，主要是利用了3D声音能够提供给听者位置信息的功能，来测试声音对人认知物体间运动因果关系的影响：当一个运动的物体碰撞一个静止物体时，观察者会立刻认知到相应的因果事件，那就是运动的物体造成了这个静止物体向前运动。但是如果这个碰撞事件同时伴随着碰撞声音，和存在第二个物体运动的时间延迟，观察者对整个因果事件的判断就会受到影响。因为在真实世界动态改变声音各种参数和设置时间延迟比较困难，所以本工作利用了沉浸式的虚拟环境来进行探究，最终结果证明了碰撞声音发生位置的偏离也将削弱人对物体间运动因果关系的认知。

**关键词：** 声音合成，虚拟仿真测试, 刚体模拟

**Abstract**

Under the over-all situation that lots of research in the areas of simulation based on visual effects is massively utilized for animation production, virtual environments setup, human modeling, pre-simulated hazard environments and so on, research into simulation based on other sense organs (e.g., audio, haptic, olfactory feelings) has gradually become a new trend to build a more immersive virtual world by exploiting computing resources than only focuses on visual parts. Nowadays, game engine such as Unity3D, Unreal Engine, and Blender combing consumer-grade virtual reality devices like Oculus and HTC Vive, which provide users high-quality immersive feelings in virtual world, are increasingly popular. We take advantage of these in our work. We present an efficient method to produce impact sound between rigid-body objects based on physical simulation and these sounds are synthesized in real time, which can also be put into virtual environments. By utilizing mass-spring-damping system to extract vibrations (sinusoid waves) on rigid-body objects, we can synthesis sounds integrating more physical information such as impact point and force.

Then, we conduct a classical psychological experiment in virtual environments based on the method of pre-recorded sounds. We mainly take advantage of 3D sounds that can offer position information to listeners to test influence of sound that affects humans’ causal perception on launching events, which means that when a moving object collide with stationary one, an observer will immediately perceive a causal event that a moving object cause the stationary one to move forward. However, when the second object’s motion is delayed, or is accompanied by a collision sound, such causal impressions attenuate and strengthen. We research into this experiments wholly in virtual environments because it is really difficult to dynamically setup time delay of objects’ motions and parameters of sounds. The final results approve that the position of impact sound will attenuate people’s causal perception on such launching events.

Keywords: Sound Synthesis, Virtually Simulated Test, Rigid-body Simulation

**目录**

[第一章 引言 4](#_Toc6748834)

[第二章 相关工作 6](#_Toc6748835)

[第三章 技术方法 7](#_Toc6748836)

[第四章 测试及实验 8](#_Toc6748837)

[第五章 结果与讨论 9](#_Toc6748838)

[第六章 总结 10](#_Toc6748839)

[参考文献 11](#_Toc6748840)

第一章 引言

1.1 声音模拟

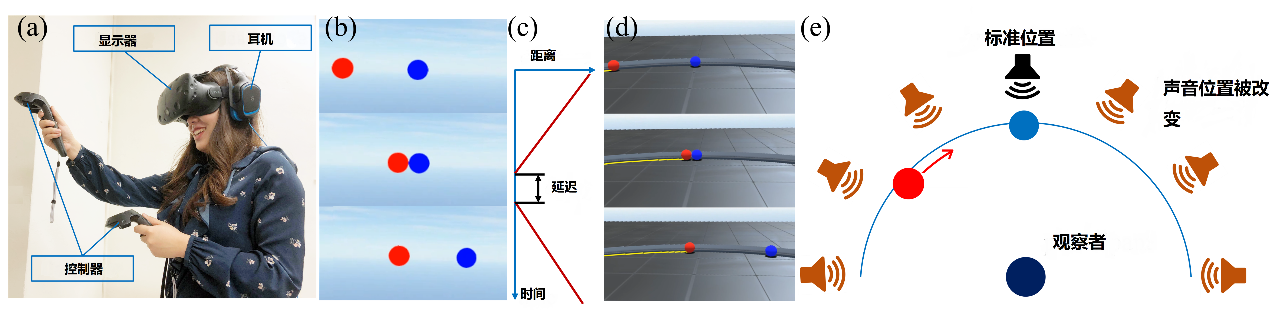
当前游戏引擎下，大多数的交互引用都是利用的已经记录好的声音来给场景中模拟的物体交互提供声音的反馈[3]。它主要利用了这样的声音实时性很好且产生声音的过程非常快，使得能够比较好的应用于大型模拟场景中，但是这样的声音很难考虑到物体交互间比较复杂的物理性质，仅仅能够改变一下声音总体的大小和音高等。以碰撞盘子状物体发出的声音为例，盘子声音的大小和音色在敲击盘子边缘和中心时肯定是不同的。所以，基于模拟与仿真对声音进行建模的方法通被提出且应用于生成合理的，考虑了较多物理因素(如碰撞力的大小/方向)的声音。本工作建立了对应刚体的弹簧质点模型并结合有限元建模[2, 4, 5]的方法来抓住刚体表面比较小规模的震动从而自然生成刚体之间的碰撞声音。本工作基于一个重要理论[1]：即可以用一系列的正弦波形（无论谐波与否）的叠加（傅里叶合成）来表示合理的刚体碰撞声音。我们利用从偏微分 (PDE) 转换为常微分方程 (ODE)的方法推导出了上述声音建模对应方程的解析解[6, 4, 5]，从而得到组成物体声音对应的一系列频率和大小。同时，利用第二部分的工作，我们提出了一些基于听者认知方面的加速方案，通过牺牲一定的声音质量，使得能够更快速地进行声音合成。本节工作主要贡献如下：

1) 该声音合成方法可以用于实时交互性的模拟场景，提供更加真实的碰撞声音以及连续型的滚动声音。

2) 本声音生成方法基于离散化的刚体几何信息物理表示(利用了四面体来作为物体网格信息的基本元进行声音震动模型建模)，因此简化了对应公式且易于实现。

1.2 心理认知实验

本节工作的目的主要是建立从人体感知层面对上述声音合成模拟工作提供加速方案的理论支持，因此简化了碰撞声音的场景，限制在小球碰撞场景中。本节利用声音合成方案是对已记录的声音片段进行二次处理来适配碰撞的场景变化，二次处理包括声音频率，大小，音效的添加等。现考虑以下的视觉场景：一个2D的红球沿直线向一个2D蓝球移动直到这两个小球的边缘相互接触。在这两个小球接触后，红球停下运动，蓝球向前以直线运动远离红球。尽管在这样的场景中没有直接的信号提示表明这两个小球之间的运动存在某种因果关系，但观察者仍然会很可能认为是红球造成了蓝球向前运动[7]。这就是一个关于运动的因果认知的例子，这种立刻地，自动产生且不可磨灭地因果认知来源于低层次地认知输入[8]。它区别于高层次的认知干扰，这种低层次的认知描述了我们对真实世界的理解是来源于对逻辑规则和自身认知知识的运用，而不是高层次知识比如物理定律，相关背景知识或者经验等[9]，这也是因果认知对印象建立进行阐述的重要特点[8,9,10]。因果认知对于动态时间的时空性质是非常敏感的[8,11]。比如，在上述的视频播放进行实验的条件下：(1) 如果在红色小球停下和蓝色小球运动之间有一个暂时的延迟；(2) 如果两个小球的边缘在碰撞的时刻是分开或者重叠；(3)又或者如果蓝色的小球开始运动的方向和之前红色小球运动的方向不一致，这种针对于红色小球造成蓝色小球向前运动的认知就会减少[12,13,14]。对于这样的触发事件也能够被周围形状的运动所影响[14]，这表明人体的认知系统能够快速处理时空信息和实觉内容



**图1. (a): 在2D((b))和3D((d)) 测试环境下虚拟现实相关设备的使用。(c): 在2D和3D测试环境下红球向碰撞点移动，停下，在一定延迟后，蓝球向前移动。(e): 两球碰撞会发出碰撞声音，同时这个声音的位置可以被改变。**

信息来从感知输入中立刻建立因果印象[9,15]。

迄今为止，研究人员已经大量的利用了2D的利用视频播放的方式来检验因果感知事件[8,12,13,14]。这主要取决于在真实世界中改变运动物体时空性质的困难性。具体而言，想要暂停真实世界中的碰撞事件且没有外界工具的介入是非常困难的：比如使用磁铁或者数字遥控器来来在透明轨道上以既定速度移动金属物体。然而，虚拟现实技术提供了方法在模拟的3D环境中来处理这样的难题。而且虚拟现实可以在3D环境中轻易改变声音的位置。前述工作已经表明当一个碰撞事件如果伴随一个表明碰撞物体之间关联的声音信息(如碰撞声音)，观察者会留下出比无声碰撞事件更强的因果印象(即认为红色物体造成蓝色物体向前运动的可能性更大了)[16]。然后对于人体感知系统是否能对形成上述因果印象时，同时对声音位置信息进行正确解析仍然是未知的。比如如腹语的情形，人们会认为声音是从表演者手里的娃娃发出的[17]。因此本工作想要解决如下问题：

1) 上述的在2D环境下进行经典小球碰撞实验所得出的结论能否延展到虚拟3D环境中。

2) 碰撞声音的位置信息能否影响对小球碰撞因果关系的认知。

因此本工作进行了三个对应的实验来解决上述问题。在每一个实验中，一个一开始运动的红色小球与一个一开始静止的蓝色小球碰撞，在0-400毫秒的延迟过后，蓝色小球开始向前运动 (见图1(b,d))。在实验一中，参与者能够在2D的环境下建立起该事件的因果联系，印象，无论是有没有碰撞声音的发出。第一个实验是Guski and Troje [16]之前工作的再现，是被设计来确保以前他们工作所发现的结论可以被延申到虚拟环境(2D)中。实验2，参与者完成同样的观察任务但是实在3D虚拟现实环境中。实验二中的碰撞声音的位置仍然是在碰撞点的标准位置上。实验二的目的是确保以前工作的结论能够延展到3D的虚拟现实环境中，实验三与实验二实验条件相同，除了改变了声音的位置信息 (见图1(e))。总的来说，本节工作具体如下贡献：

1) 在虚拟现实环境中复刻了之前因果印象实验的工作证明了虚拟技术可以用作对检测人认知和心理具体价值

2) 检验了声音位置被改变的碰撞声音对人感知这样的碰撞事件因果性的影响。

3) 测验了人在虚拟环境中对声音位置的估计情况。

第二章 相关工作

2.1 声音合成

对于利用计算机进行声音合成的方法比较广：如今范围用的最广就是利用采样的方法，将真实环境中采样的声音放入仿真环境中 (广泛见于电影声音制作中 – Foley[18])。该方法充分利用采样，实时性较高，且能够对收集的声音进行再处理 (如改变频率，大小，添加环境音效等)。但对于大型复杂环境而言，当场景发生物体太多且复杂，模拟环境中需要存储的声音也会比较多，不仅使得采样工作耗时耗力，且高质量的采样声音存储占用内存也比较高[6]。因此，Perry Cook, D.Brandon Lloyd等提出了在基于傅里叶合成基础上利用分析声音频谱，提取合理正弦波进行更底层的声音合成的方法来应用于交互环境中[1,19]。该方法利用了频谱建模的方法，认为声音由正弦波和噪音组成，且正弦波的叠加能够表示出所合成声音绝大多数的性质。该方法快速且与声音设计师的开发思路相接近，问题在于这是一个基于数据收集的方法，它对声音采样的数量要求也不低。

为了合成声音更具有真实性：能够根据敲击位置的不同有不同的反馈，O’Brien等人提出了从输入物体网格信息和材质信息来提取振动模式(振动模式被认为以正弦波的形式存在)从而生成声音的方法[21]，成功建立起声音物理性质和模拟之间的联系，使得声音的模拟能够细化到每一种材质，每一个物体。为了实现更快的声音合成速度，Raghuvanshi和Lin等提出了一系列的加速技术来实现可以用于实时场景，同时利用弹簧质点模型对物体表面建模来近似振动的情形，可以得到更易实现的解析解[6]。另一种加速手段，由James，Dingzeyu和Changxi等提出[4,5]：利用从频率域对振动模式进行处理，再利用傅里叶逆变换/拉普拉斯逆变换回到时间域。本工作同样利用弹簧质点模型对振动进行建模处理，但使用了更精确的方法-有限元分析[21,22]-来计算弹簧质点模型的相应参数，如胡克常数和质点质量等。利用Eftychios D. Sifakis等提出的对刚体进行有限元分析离散化建模的方法[2]，结合物体四面体网格信息，我们可以建立对刚体表面振动更准确的模型来合成更自然的声音。对于实时场景而言，James和Lin等都提出需要将弹簧振动的偏微分方程转换为常微分方程，从而使得每个质点的每一个自由度对应的振动模式能够独立地被计算[4,5,6]。

最后需要将声音合成系统放入模拟场景中，需要合适的物理引擎来处理对物体作用的实时碰撞力，本工作利用了Unity游戏引擎[20]进行处理，并利用其进行简单的场景测试。

2.2 感知研究与测试环境模拟

对于因果关系感知的联系是人体对动态世界进行感知的基本。它广泛地存在于日常生活中，比如6到10个月的婴儿就能够对视觉场景中的物体运动因果关系有认知[23,24]。正如Michotte的经典工作[11]，物体间的碰撞或者说发射事件被认为是非常有价值的系统来检测因果联系的认知(详见于[8,12,13,14,16,25])。因此，这样的因果联系的推断可以在认知层面上进行建模，2.21和2.22节叙述了最近利用虚拟现实技术进行相应研究的方法和例子。

2.21 因果感知

近来，许多研究者已经提出了因果关系可以通过推断在视觉场景中观察到的物体之间联系的边缘概率来得到对应的感知建立，就是所谓的噪音牛顿框架[13,26]。因为即使具有基本的物理知识，人对简单的可感知的变量的估计(如:距离，速度和时间等)是本能不确定的，对于他们的准确估计必须根据预估和提前观察。一些简单的运动例子：1.人是会认为物体是缓慢运动的，或者越动越慢(而不是按照运动学的一直运动)。2.对于一个碰撞事件而言，人会认为一个静止的物体在被撞后应该立即运动，这样这个碰撞才被认为是发生了[13]。因此当通过视觉在观察这样的碰撞事件时，对于速度，物体位置的估计，时间延迟等都是基于人自己之前有一定的预估和观察。

噪音牛顿框架更进一步的认为人体是有自己在神经回路中建立一套属于自己的物理模型[27]来对真实的物理规则进行近似和估计，从而生成一套随着时间变化的感知。所以对于基于感知和物理建模的碰撞事件，物体应该按照动量守恒定律进行运动，而对于没有感知与物理的建模，物体将随机运动。所以两个物体在碰撞前后的速度，位置，以及运动的时间延迟等都被用作重要因素来对这两个物体之间的因果联系进行判断。这一套理论提供了人体对运动事件因果感知的量化估计，它是随时间和空间的变化而变化的。比如对于上述碰撞事件而言，当两个小球之间的运动延迟和位置拉大时，对于他们之间存在碰撞的因果联系的可能性估计就会降低[13]。

尽管噪音牛顿框架已经从视觉层面揭示了人类对物理世界的感知情况[26]。但对于同时包含视觉和听觉信息的情况还需要进行建模及测试。对于听觉信息如声音位置信息等能否在建立上述因果联系时同样被人体感知仍然是不清晰的，因此本工作目的就是做出对物理世界进行感知层面建模复杂度与深度分析的第一步。

2.22 虚拟现实

虚拟现实技术已经展现了作为低成本高效的手段，可充分应用于模拟极端环境来测试和训练人类。比如，虚拟现实结合游戏系统被广泛应用来训练人进行在灾难逃生[29,39,31,32]，医疗急救[33,34]，灭火训练[35,36,37]，航空安全[38]，交通和火灾[39,40,41]等。这些针对训练设计的系统都非常有用，因为他们利用模拟的方法将本很危险的环境放置在了安全可控的环境中。

然而，以前的虚拟现实设备受限于计算能力和较差的声音与视频质量，造成上述的研究只能集中于简单的重复性过程训练。传统的虚拟现实系统也受限于需要大型测试环境和大量的动作捕捉设备，使得利用虚拟现实技术进行心理学方面的研究不够顺利。因为参与者需要经过大量训练来适应复杂的虚拟现实设备。因此，以前的心理学实验都仅仅基于视频播放和问卷调查等，虚拟现实技术仅仅是一种想要被用于进行更好实验的设想[42,43,44]。

近来由于虚拟现实工业的大发展，许多消费级的虚拟设备比如简化了环境的搭建，同时GPU计算能力的大量提升使得听觉和视觉体验得到极大提升。在过去的两年时间里，虚拟现实已经被广泛应用于实验性研究(模拟复杂环境)。比如地震的模拟[45]，视觉导航[46],语义分析[47]，和机器人控制[48,49]。各种虚拟现实平台被用于人机交互场景设计[50]以及自动驾驶模拟[51]。在认知研究方面，研究人员已经利用虚拟现实设备来研究人类的欺骗行为[52]，重力感知[53]，视觉认知调节能力[54]，触觉重定向能力[55]，以及位置感知能力[56]。本工作目的是建立起在3D虚拟环境中完整的心理学实验流程，证明虚拟现实在心理学实验的可行性。

第三章 技术方法

本方法中，声音在刚性物体受外界力冲击下刚体物体的表面弹性振动产生的。本方法将这些振动都近似为正弦波的简谐振动形式[6]。这些振动能够搅乱物体周围空气从而进行波形的传递从而发出声音。如果声音的频率在20-2200Hz之间，将可以被人耳听到。当给予这些振动合理的衰减模式(本方法采用理想的指数型衰减进行近似)，结合傅里叶叠加原理[1]，将随时间衰减的正弦波进行叠加，将得到与物体本身网格形状和材质适配的声音(如金属球声，砖块声等)。

对于物体表面振动波形的提取，本方法利用物体网格信息建立对应的弹簧质点模型，利用基本的弹簧振动来近似物体表面因碰撞产生的振动波形。因此，对于一个输入的由顶点和边组成的物体网格，我们将质点代替顶点，带阻尼的弹簧代替边从而建立起基于物体形状的弹簧质点系统[5,6]。下面，我将从单个质点的振动模型开始进行整个声音建模的方法解析。

3.1基本质点振动模型

对于每一个质点而言，它的运动是一个基本的弹簧阻尼模型(见图1),这里认为质点x，只有一个自由度(即在x一个方向运动)，但实际上三维物体中每一个质点应有三个自由度(x,y,z方向)。因此一个具有n个质点的物体，对应的震动模型的质点运动应用一个3n\*1的向量进行描述。

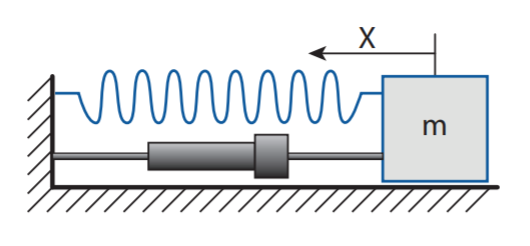


图1 弹簧阻尼模型

从而可以利用下述二次常微分方程，获得质点简谐运动时合理的正弦波:

(1)

x描述质点位置，m为物体质量，k表示弹簧的刚度(stiffness)，也是这里的胡克常数；f(t)描述了跟随时间变化的外力；则表示了x的二次倒数，也就是加速度; 则表示了x的一次导数，也就是速度；c表示衰减模型，本工作中都需要设置合理的衰减系数，使得能够产生震动且最后停止(即欠阻尼情形)，即c应该小于临界阻尼cc=。为了得到公式(1)的解析解，我们利用拉普拉斯逆变换，并认为f(t)为狄拉克函数形式的单位冲击，得到x的解析解。

(2)

(3)

(4)

(5)

w是没有衰减的振动频率，是衰减因子，因此是有阻尼下，衰减的振动频率。从公式中也可以得到，衰减频率必须在的情况下才有意义，同时由这里也可以证明c<cc。利用傅里叶叠加的原理[1],对公式(2)中得到的每一个质点对应的正弦波进行叠加，就得到了物体产生的声音，从中我们可以看出：我们在这里是将力以狄拉克函数(Dirac Delta Function)形式，作为单位冲击来表示的。因此对于多个不同强度的任意外力，我们需要将他们进行叠加从而得到声音的强度信息，这也是在声音合成中唯一需要实时计算的步骤，利用的是如下时间积分。

(6)

3.2 三维物体振动模型

对于一个三维物体而言，利用矩阵思维将公式(1)进行拓展，得到如下方程:

(7)

其中表示质点位移，物体由n个质点组成那么 ,为质点质量组成的质量矩阵(Mass Matrix,是一个对角矩阵)，为胡克常数组成的刚度矩阵(Stiff Matrix,是一个对称矩阵)，为外力。很明显，这个方程是一个二次型偏微分方程，是非线性的，想要求解来得到质点位移是非常复杂的。因此我们想要将其转换为常微分方程，从而得到每一个质点位移的解析解[2,3]。核心思想是利用对K和M进行广义特征值分解：

(8)

通过公式(8)的广义特征值分解得到了由特征向量组成的震动模型形状描述的矩阵和由特征值组成的对角矩阵。U决定了每一个质点得到的正弦波对最终合成声音的贡献程度，的对角元素即是。根据特征值分解的性质，是满秩矩阵。所以可以用线性合成的方式表示公式(7)中的质点位移，。然后把它带入公式(7)中，可以得到公式(9)：

(6)

再根据特征值分解的性质：和，同时利用Rayleigh衰减模型 公式(5)的偏微分方程形式最终可以变为如下常微分方程形式:

(10)

所以到公式(10)为止我们得到了关于质点位移的常微分方程，而且其中向量中的每一个元素都是相互独立的，所以可以解出公式(2)得到的每一个质点的位置偏离方程，且通过矩阵，我们在预处理阶段便得到每一个正弦波的频率。需要合成声音的部分只剩下通过碰撞力得到每一个正弦波对应的振幅，从而合成声音最终的大小。

3.4 有限元分析

通过输入的物体模型得到的质量矩阵和刚性矩阵决定了极大程度上决定了最终声音合成的质量。因此本方法将利用有限元分析的方法来获得更精确的质量矩阵和刚性矩阵。对角阵和对称阵!

3.5 声音输出 (OpenAL)

当得到物体所有质点的振动正弦波(数量为3N)后，利用傅里叶叠加的原理，将声音所有波形叠加在一起即可以的到最终的声音(见代码说话)。

3.6 声音合成加速处理

本节工作利用了Lin等人提出的模式截断的方法，对单个物体的声音合成进行了加速处理以达到实时的要求。

第四章 测试及实验

模拟出的声音push 到虚拟环境中去

就可以把自己那一套套进去，具体写出声音的过程，为什么没有做之前的东西，因为见到那球体碰撞暂时不需要，**也不需要考虑物体材质的信息**。且我们重点是提供声音位置信息，所以直接用了且加环境音，

第五章 结果与讨论

5.1 场景中实时声音合成

第六章 总结

人对声音位置有感知，然后所以基于位置信息进行加速方案处理(质量分级)

参考文献

[1] Cook, P. R. (2002). *Real sound synthesis for interactive applications*. AK Peters/CRC Press.

[2] Sifakis, E., & Barbic, J. (2012, August). FEM simulation of 3D deformable solids: a practitioner's guide to theory, discretization and model reduction. In *Acm Siggraph 2012 courses* (p. 20). ACM.

[3] Liu, L., Wang, R., Xu, S., & Bai, Y. (2018, August). Development of AR/VR Project Based on Unity3D Engine. In *Journal of Physics: Conference Series* (Vol. 1069, No. 1, p. 012033). IOP Publishing.

[4] Li, D., Fei, Y., & Zheng, C. (2015). Interactive acoustic transfer approximation for modal sound. *ACM Transactions on Graphics (TOG)*, *35*(1), 2.

[5] James, D. L. (2016, July). Physically based sound for computer animation and virtual environments. In *ACM SIGGRAPH 2016 Courses* (p. 22). ACM.

[6] 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.

[7] Hume, D. (2006). A Treatise of Human Nature: Volume 1: Texts.

[8] Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in cognitive sciences*, *4*(8), 299-309.

[9] Roser, M. E., Fugelsang, J. A., Dunbar, K. N., Corballis, P. M., & Gazzaniga, M. S. (2005). Dissociating processes supporting causal perception and causal inference in the brain. *Neuropsychology*, *19*(5), 591.

[10] Schlottmann, A., & Shanks, D. R. (1992). Evidence for a distinction between judged and perceived causality. *The Quarterly Journal of Experimental Psychology*, *44*(2), 321-342.

[11] Natsoulas, T. (1961). Principles of momentum and kinetic energy in the perception of causality. *The American Journal of Psychology*, *74*(3), 394-402.

[12] Mayrhofer, R., & Waldmann, M. R. (2014). Indicators of causal agency in physical interactions: The role of the prior context. *Cognition*, *132*(3), 485-490.

[13] Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological review*, *120*(2), 411.

[14] Scholl, B. J., & Nakayama, K. (2002). Causal capture: Contextual effects on the perception of collision events. *Psychological Science*, *13*(6), 493-498.

[15] Roser, M. E., Fugelsang, J. A., Dunbar, K. N., Corballis, P. M., & Gazzaniga, M. S. (2005). Dissociating processes supporting causal perception and causal inference in the brain. *Neuropsychology*, *19*(5), 591.

[16] Guski, R., & Troje, N. F. (2003). Audiovisual phenomenal causality. *Perception & psychophysics*, *65*(5), 789-800.

[17] Alais, D., & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology*, *14*(3), 257-262.

[18] Ament, V. T. (2014). *The Foley grail: The art of performing sound for film, games, and animation*. Focal Press.

[19] 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.

[20] Engine, U. G. (2008). Unity game engine-official site. *Online][Cited: October 9, 2008.] http://unity3d. com*, 1534-4320.

[21] Director-O'Brien, J. F. (2001, August). Synthesizing sounds from physically based motion. In *ACM SIGGRAPH 2001 video review on Animation theater program* (p. 59). ACM.

[22] O'Brien, J. F., Shen, C., & Gatchalian, C. M. (2002, July). Synthesizing sounds from rigid-body simulations. In *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation* (pp. 175-181). ACM.

[23] Cohen, L.B., & Oakes, L.M. (1993).How infants perceive a simple causal event. *Develo-pmental Psychology*, *29* (3), 421.

[24] Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cog-nition*, *25*(3), 265-288.

[25] Ye, T., Qi, S., Kubricht, J., Zhu, Y., Lu, H., & Zhu, S. C. (2017). The martian: Examining hu-man physical judgments across virtual gravity fields. *IEEE transactions on visualization and computer graphics*, *23*(4), 1399-1408.

[26] Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological review*, *116*(4), 661.

[27] Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neu-roanatomy of intuitive physical inference. *Proceedings of the national academy of sciences*, *113*(34), E5072-E5081.

[28] Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in cognitive sciences*, *21*(10), 749-759.

[29] Mól, A. C. A., Jorge, C. A. F., & Couto, P. M. (2008). Using a game engine for VR simulations in evacuation planning. *IEEE computer graphics and applications*, *28*(3), 6-12.

[30] Reznek, M., Harter, P., & Krummel, T. (2002). Virtual reality and simulation: training the future emergency physician. *Academic Emergency Medicine*, *9*(1), 78-87.

[31] Tarnanas, I., & Manos, G. C. (2001). Using virtual reality to teach special populations how to cope in crisis: the case of a virtual earthquake. *Studies in health technology and informatics*, *81*, 495-501.

[32] Xi, M., & Smith, S. P. (2014, March). Simulating cooperative fire evacuation training in a virtual environment using gaming technology. In *2014 IEEE Virtual Reality (VR)*(pp. 139-140).

[33] Stansfield, S., Shawver, D., Sobel, A., Prasad, M., & Tapia, L. (2000). Design and imp-lementation of a virtual reality system and its application to training medical first responders. *Presence: Teleoperators & Virtual Environments*, *9*(6), 524-556.

[34] Andreatta, P. B., Maslowski, E., Petty, S., Shim, W., Marsh, M., Hall, T., ... & Frankel, J. (2010). Virtual reality triage training provides a viable solution for disaster‐preparedness. *Academic emergency medicine*, *17*(8), 870-876.

[35] Backlund, P., Engstrom, H., Hammar, C., Johannesson, M., & Lebram, M. (2007, July). Sidh-a game based firefighter training simulation. In *2007 11th International Conference Information Visualization (IV'07)* (pp. 899-907). IEEE.

[36] Cha, M., Han, S., Lee, J., & Choi, B. (2012). A virtual reality based fire training simulator integrated with fire dynamics data. *Fire Safety Journal*, *50*, 12-24.

[37] Tate, D. L., Sibert, L., & King, T. (1997, March). Virtual environments for shipboard fire-fighting training. In *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality* (pp. 61-68). IEEE.

[38] Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE transactions on visualization and computer graphics*, *21*(4), 529-538.

[39] Backlund, P., Engstrom, H., Johannesson, M., & Lebram, M. (2007, July). Games and traffic safety-an experimental study in a game-based simulation environment. In *2007 11th International Conference Information Visualization (IV'07)* (pp. 908-916). IEEE.

[40] McComas, J., MacKay, M., & Pivik, J. (2002). Effectiveness of virtual reality for teaching pedestrian safety. *CyberPsychology & Behavior*, *5*(3), 185-190.

[41] Padgett, L. S., Strickland, D., & Coles, C. D. (2005). Case study: using a virtual reality computer game to teach fire safety skills to children diagnosed with fetal alcohol syndrome. *Journal of pediatric psychology*, *31*(1), 65-70.

[42] Kozlov, M. D., & Johansen, M. K. (2010). Real behavior in virtual environments: Psychology experiments in a simple virtual-reality paradigm using video games. *Cyberpsychology, behavior, and social networking*, *13*(6), 711-714.

[43] Olivier, A. H., Bruneau, J., Cirio, G., & Pettré, J. (2014). A virtual reality platform to study crowd behaviors. *Transportation Research Procedia*, *2*, 114-122.

[44] Rovira, A., Swapp, D., Spanlang, B., & Slater, M. (2009). The use of virtual reality in the study of people's responses to violent incidents. *Frontiers in behavioral neuroscience*, *3*, 59.

[45] Li, C., Liang, W., Quigley, C., Zhao, Y., & Yu, L. F. (2017). Earthquake safety training through virtual drills. *IEEE transactions on visualization and computer graphics*, *23*(4), 1275-1284.

[46] Zhu, Y., Mottaghi, R., Kolve, E., Lim, J. J., Gupta, A., Fei-Fei, L., & Farhadi, A. (2017, May). Target-driven visual navigation in indoor scenes using deep reinforcement learning. In *2017 IEEE international conference on robotics and automation (ICRA)* (pp. 3357-3364). IEEE.

[47] Zhu, Y., Gordon, D., Kolve, E., Fox, D., Fei-Fei, L., Gupta, A., ... & Farhadi, A. (2017). Visual semantic planning using deep successor representations. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 483-492).

[48] Hertkorn, K., Roa, M. A., Brucker, M., Kremer, P., & Borst, C. (2013, November). Virtual reality support for teleoperation using online grasp planning. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2074-2074). IEEE.

[49] Yan, X., Khansari, M., Bai, Y., Hsu, J., Pathak, A., Gupta, A., ... & Lee, H. (2017). Learning grasping interaction with geometry-aware 3d representations. *arXiv preprint arXiv:1708.07303*, *1*.

[50] Lin, J., Guo, X., Shao, J., Jiang, C., Zhu, Y., & Zhu, S. C. (2016, November). A virtual reality platform for dynamic human-scene interaction. In *SIGGRAPH ASIA 2016 virtual reality meets physical reality: Modelling and simulating virtual humans and environments* (p. 11). ACM.

[51] Shah, S., Dey, D., Lovett, C., & Kapoor, A. (2018). Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and service robotics* (pp. 621-635). Springer, Cham.

[52] Aravena, C., Vo, M., Gao, T., Shiratori, T., Yu, L. F., & Contributors, E. (2017). Perception Meets Examination: Studying Deceptive Behaviors in VR. In *CogSci*.

[53] Ye, T., Qi, S., Kubricht, J., Zhu, Y., Lu, H., & Zhu, S. C. (2017). The martian: Examining human physical judgments across virtual gravity fields. *IEEE transactions on visualization and computer graphics*, *23*(4), 1399-1408.

[54] Lin, J., Zhu, Y., Kubricht, J., Zhu, S. C., & Lu, H. (2017). Visuomotor Adaptation and Sensory Recalibration in Reversed Hand Movement Task. In *CogSci*.

[55] Azmandian, M., Hancock, M., Benko, H., Ofek, E., & Wilson, A. D. (2016, May). Haptic retargeting: Dynamic repurposing of passive haptics for enhanced virtual reality experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 1968-1979). ACM.

[56] Bruder, G., & Steinicke, F. (2014, November). Threefolded motion perception during immersive walkthroughs. In *Proceedings of the 20th ACM symposium on virtual reality software and technology* (pp. 177-185). ACM.

[49]

[50]