Interactive Geometric Sound Propagation for Complex Virtual Environments

Carl Schissler, Dinesh Manocha

UNC Department of Computer Science

http://gamma.cs.unc.edu/HIGHDIFF/

CV²R Consortium for Vision and Virtual Reality

Introduction

The generation of plausible acoustic effects is challenging for interactive applications like virtual reality and games which require low latency and involve complex scenes with many sources.

Our approach is the first to compute sound for dozens of sources in complex environments at interactive rates.

Our Contributions:

1. Backward sound propagation
2. Source clustering
3. Impulse response cache
4. Adaptive impulse response length

Overview

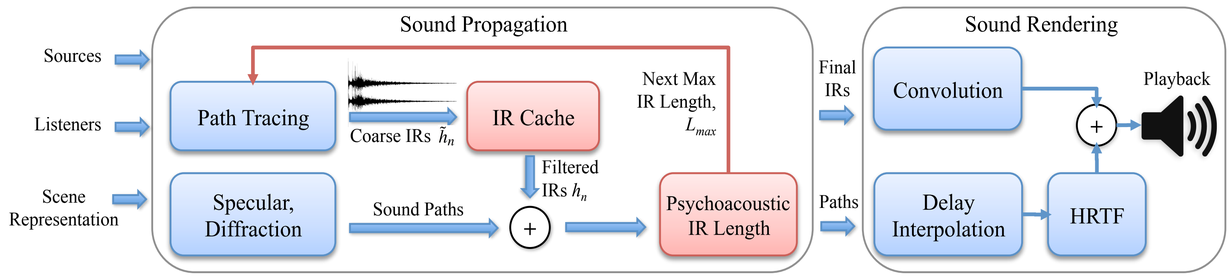
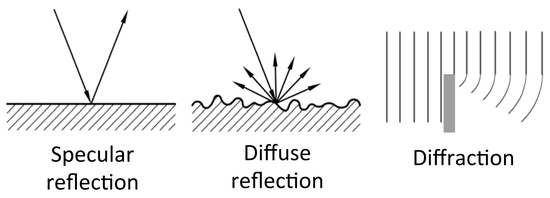


Figure 1: *We use path tracing to compute an IR for each source. Specular early reflections and diffraction are handled separately. After IR processing (IR cache, adaptive IR length), the result is rendered using streaming convolution and played for the user over headphones.*

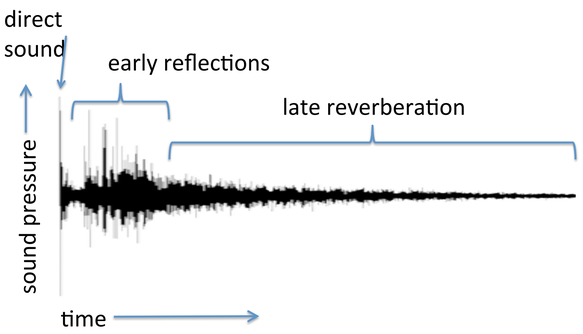
Background

Sound Propagation – given sources, listener, scene geometry, simulate how sound interacts with environment and is heard by the listener.

Acoustic phenomena:



Impulse response (IR) – 1D linear filter, transfer function between source and listener.



Impulse Response Cache

Ray tracing involves a tradeoff between the number of primary rays and the latency/quality of the simulation.

With the IR cache, fewer rays need to be traced because previous simulation results are reused.

The result is lower latency for interactive simulations.

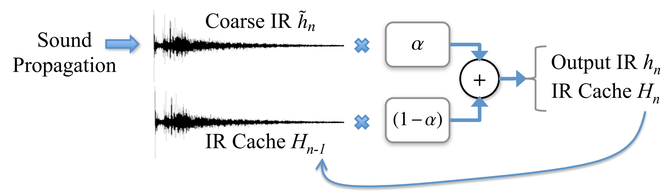


Figure 3: *The output IR on frame n is a linear combination of the IR cache from frame n-1 and a low-quality IR computed on frame n.*

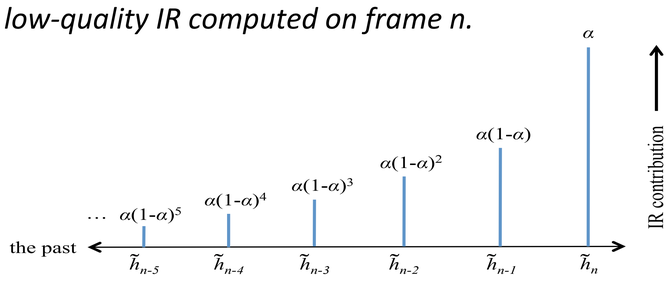


Figure 4: *The contribution of the coarse IRs from previous frames decreases over time.*

Backward Sound Propagation

• Traditional approaches emit rays from each sound source, intersect with listener.
• This is slow for scenes with many sources.

Reciprocity principle – IR is the same if source and listener exchange positions

We trace rays from the listener instead and intersect the rays with sound sources in the scene.

1. Listener-centric sampling produces better IR quality.
2. Faster than tracing rays from each source.

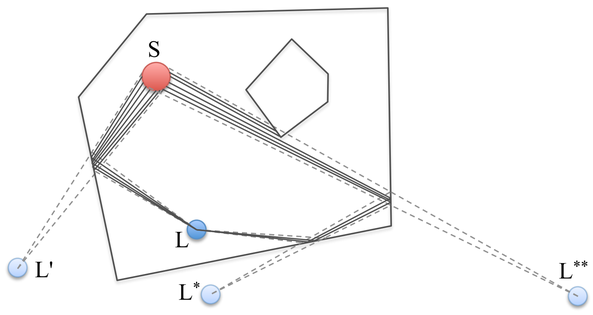


Figure 2: *Backward ray tracing from the listener is used to compute specular early reflections for spherical sound sources.*

Adaptive IR Length

• It is difficult to know the length of the impulse response that must be computed.
• Surface materials, dynamic scene elements influence the IR length.
• A constant IR length can result in wasted computation or truncated IR.

Our method dynamically determines the IR length using a psychoacoustic metric and uses that information to determine how far to trace rays on the next frame..

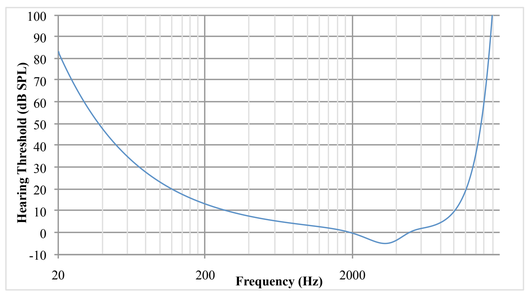


Figure 7: *The threshold of hearing for a human listener varies with frequency.*

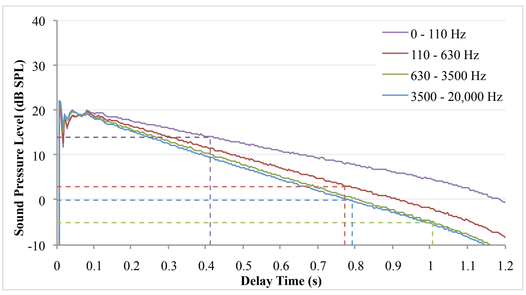


Figure 8: *We use the threshold of hearing to determine the last audible sample in the impulse response for each frequency band.*

Sound Source Clustering

In scenes with many sources, the IRs for nearby sources are frequently similar.

We cluster sources based on their spatial proximity and relation to the listener and simulate each cluster as a single source.

This gives a 2x to 4x reduction in the number of sources and results in a significant speedup.

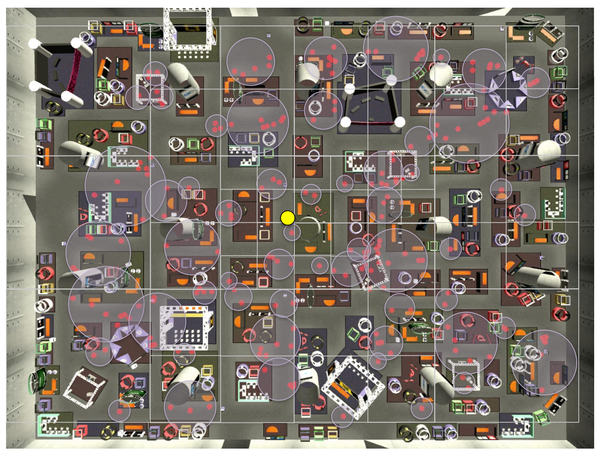


Figure 6: *top-down view of a tradeshow scene, 200 sound sources (red), 50 clusters (purple), listener (yellow).*

Results

Space Station | Office | Hangar

Scene	Scene Complexity #Tris	#Sources	Propagation Avg. #Rays	Avg. #Bounces	IR Length (s)	Memory (MB)	Time (ms) Ray Tracing	IR Cache	Total
Space Station	35,581	21	588	136	0.8 - 1.25	40.8	105.0	3.42	111.57
Office	82,125	24	680	129	0.5 - 1.0	39.3	69.0	2.70	74.08
Hangar	71,461	18	1094	95	1.5 - 3.0	75.3	100.5	6.10	110.91

Figure 5: *The main results of our system for three complex benchmark scenes running on an Intel i7 4770k CPU.*

Our approach can compute plausible acoustic effects for scenes with dozens of sources at interactive rates and is over an order of magnitude faster than previous techniques.

Acknowledgements

We thank the following organizations for their generous support of this research:

1. Army Research Office
2. National Science Foundation
3. Link Foundation Fellowship in Advanced Simulation and Training