Direct and Inverse Sound Propagation Methods

Dinesh Manocha

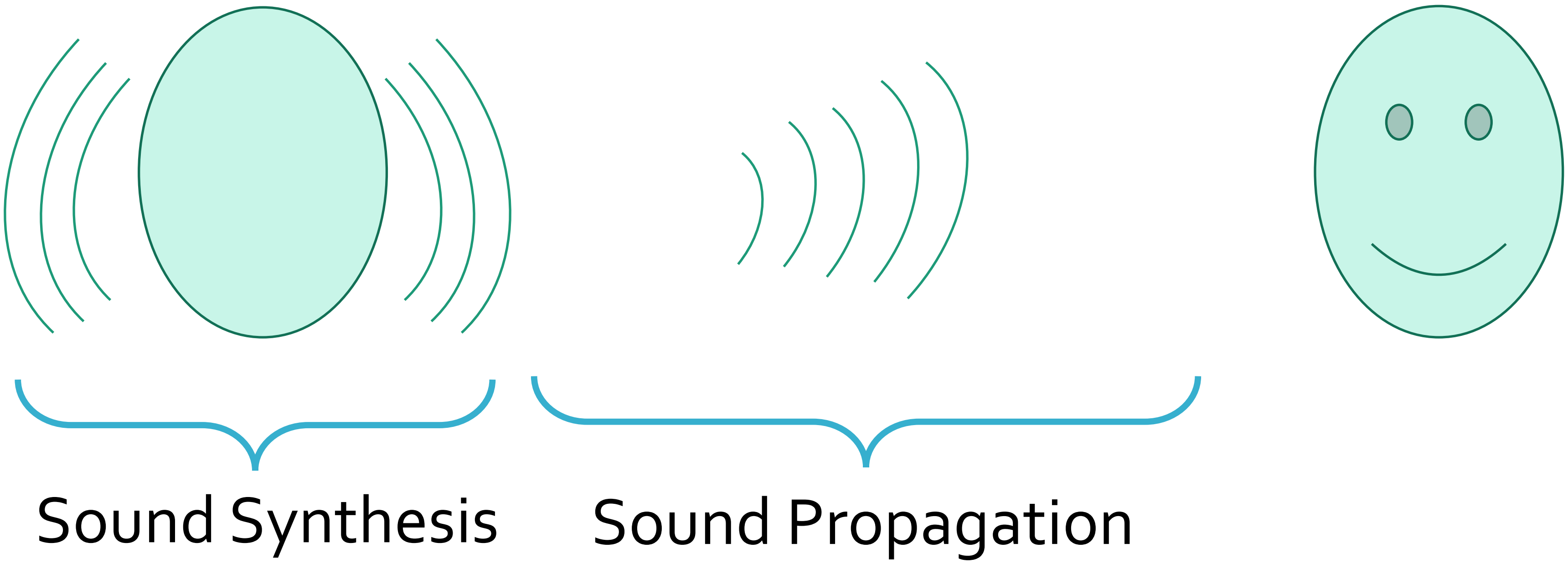
University of Maryland at College Park

dm@cs.umd.edu

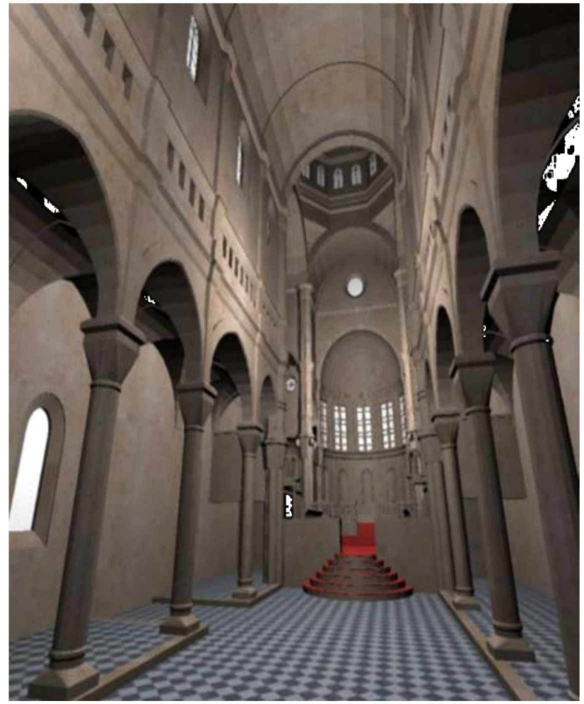
http://gamma.web.unc.edu/research/sound/

Physically-Based Sound Simulation

What is sound?



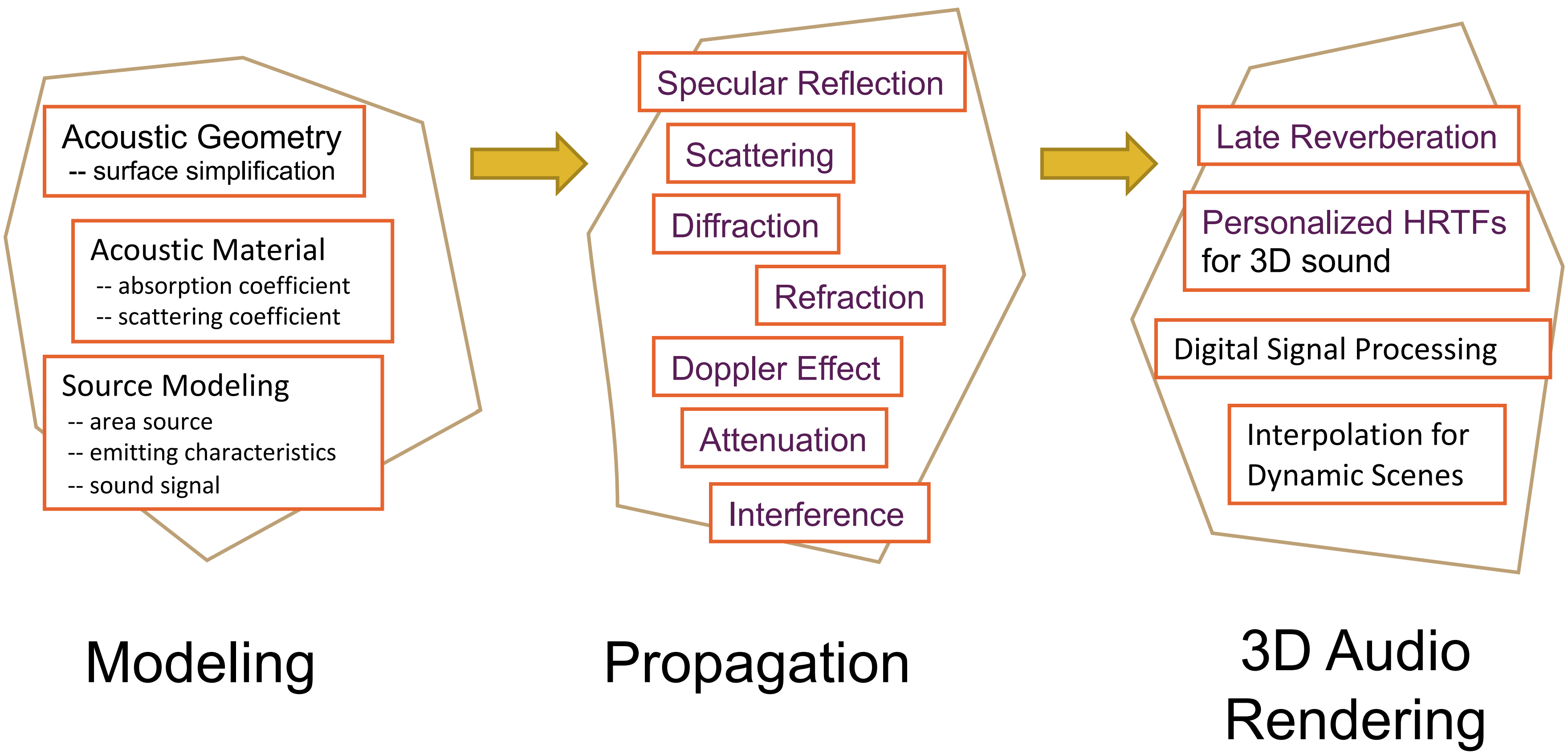
Physically-Based Sound Simulation



Sound Synthesis

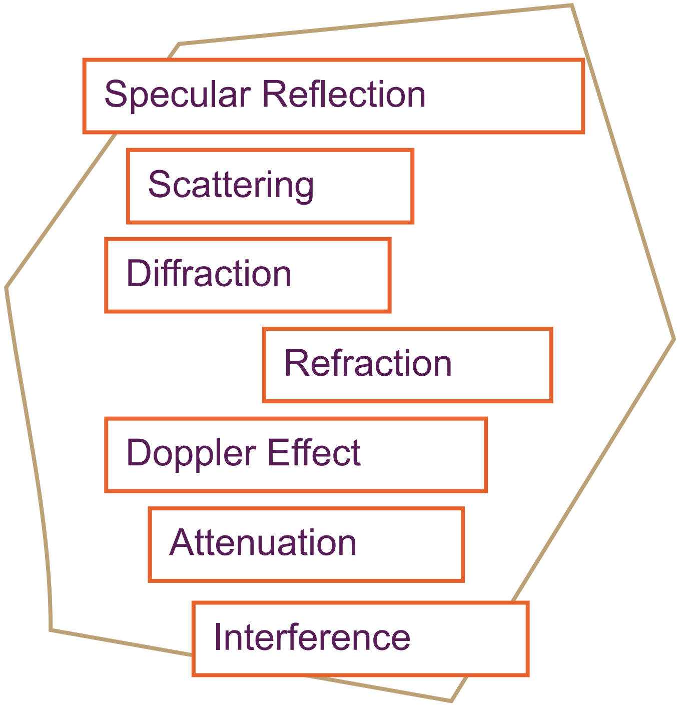
Sound Propagation

Computational Sound Simulation: An Overview



Acoustics vs. Light

- Speed: 340 m/s vs. 300,000,000 m/s
- Frequency: 20 to 20K Hz vs. RGB
- Wavelength: 17m to 17cm vs. 700 to 400 nm

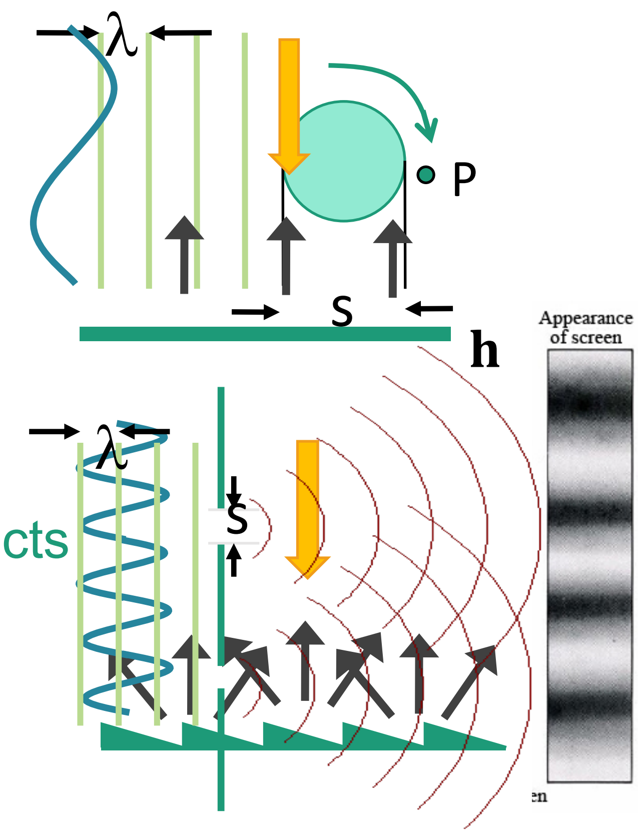


Propagation Effects

Acoustic effects

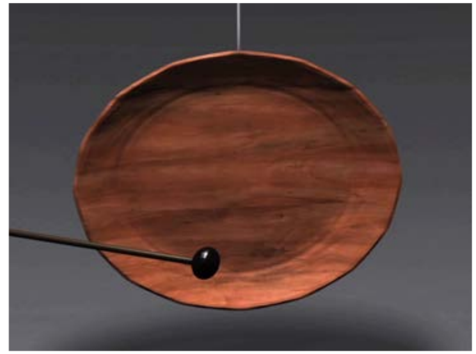
- wavelength << object size
  - specular reflection
  - scattering
- wavelength ~ object size
  - diffraction
  - interference

} wave-effects



Physically-Based Sound Simulation

our work



Interaction

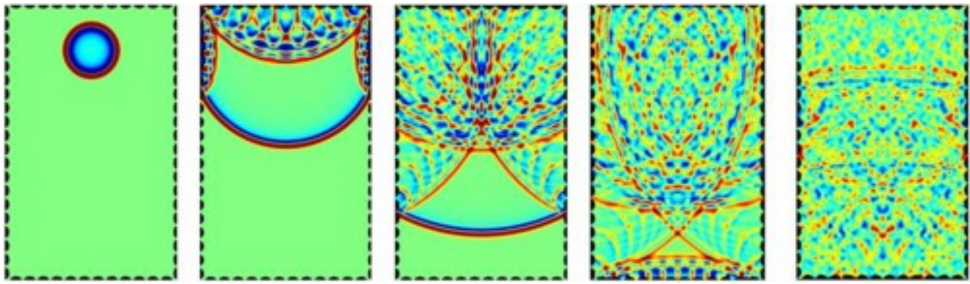
Geometry

Material

Environment

Sound Propagation

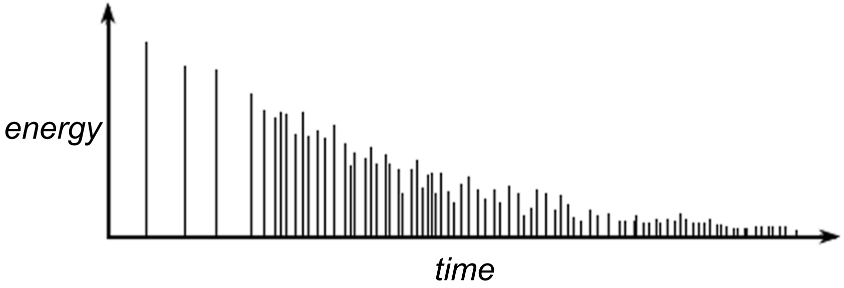
The process by which sound is emitted from a source, interacts with the environment, and is received by a listener.



Perceptual audio cues:

- Size of environment
- Shape of environment
- Locations of sources

Output: Impulse response filter



Sound Propagation – Computational Methods

Numeric:	Geometric:
Sound = wave	Sound ≈ particles, acoustic energy
Practical for low frequencies	Better for high frequencies
Complexity: O(volume), O($freq^4$)	O(log(# primitives)) per ray
Pre-computed	Interactive
Static scenes	Dynamic scenes

Interactive Sound Propagation

- Geometric Propagation
  - Ray-Approximation of Wave Equation
  - High-frequency approximation
  - Getting fast enough for interactive applications
  - Highly dependent on the geometry details
- Methods
  - Ray Tracing [Allen 1958; Krokstad,1968] [Kuttruff,1993]
  - Image Source [Borish,1984] [Dalenback,1992]
  - Beam Tracing [Funkhouser et al.1998] [Funkhouser et al. 1999]
  - Phonon Tracing [Kapralos,2004] [Bertram,2005]
  - Frustum Tracing [Lauterbach et al. 2007] [Chandak,et al. 2008][Taylor et al. 2009]
  - Acoustic Radiance Transfer [Siltanen et al.2007][Antani et al. 2012]
  - Conservative Frusta Tracing (CFT) [Chandak et al. 2009]
  - Diffuse reflections/higher order diffraction [Schissler et al. 2014, 2015]
  - Diffraction kernel and mobile devices [Schissler et al. 2017; Runga et al. 2018]

Interactive Propagation

1. Iterative diffuse + specular reflections

2. Diffraction

   - Approximate geometric models (UTD)

   - Full wave-solvers

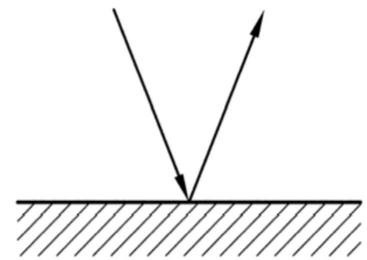
Diffuse Reflections

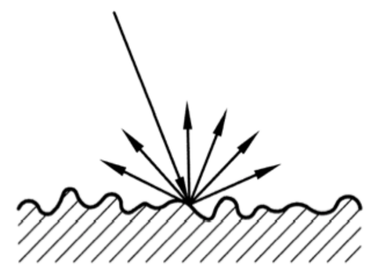
Reflected sound is scattered by rough surfaces

Frequency-dependent scattering

Diffuse sound integral:

$$w(\vec{p},t) = \frac{1}{\pi c} \int \int_S \alpha(\vec{p'}) B\left(\vec{p'}, t - \frac{L'}{c}\right) \frac{\cos\varsigma''}{L'^2} dS' + w_d(\vec{p},t).$$


Specular


Diffuse

Office Scene

6 sources
154k triangles

10^{th} order diffuse
10^{th} order specular
4^{th} order diffraction

Frame time: **33.6 ms**

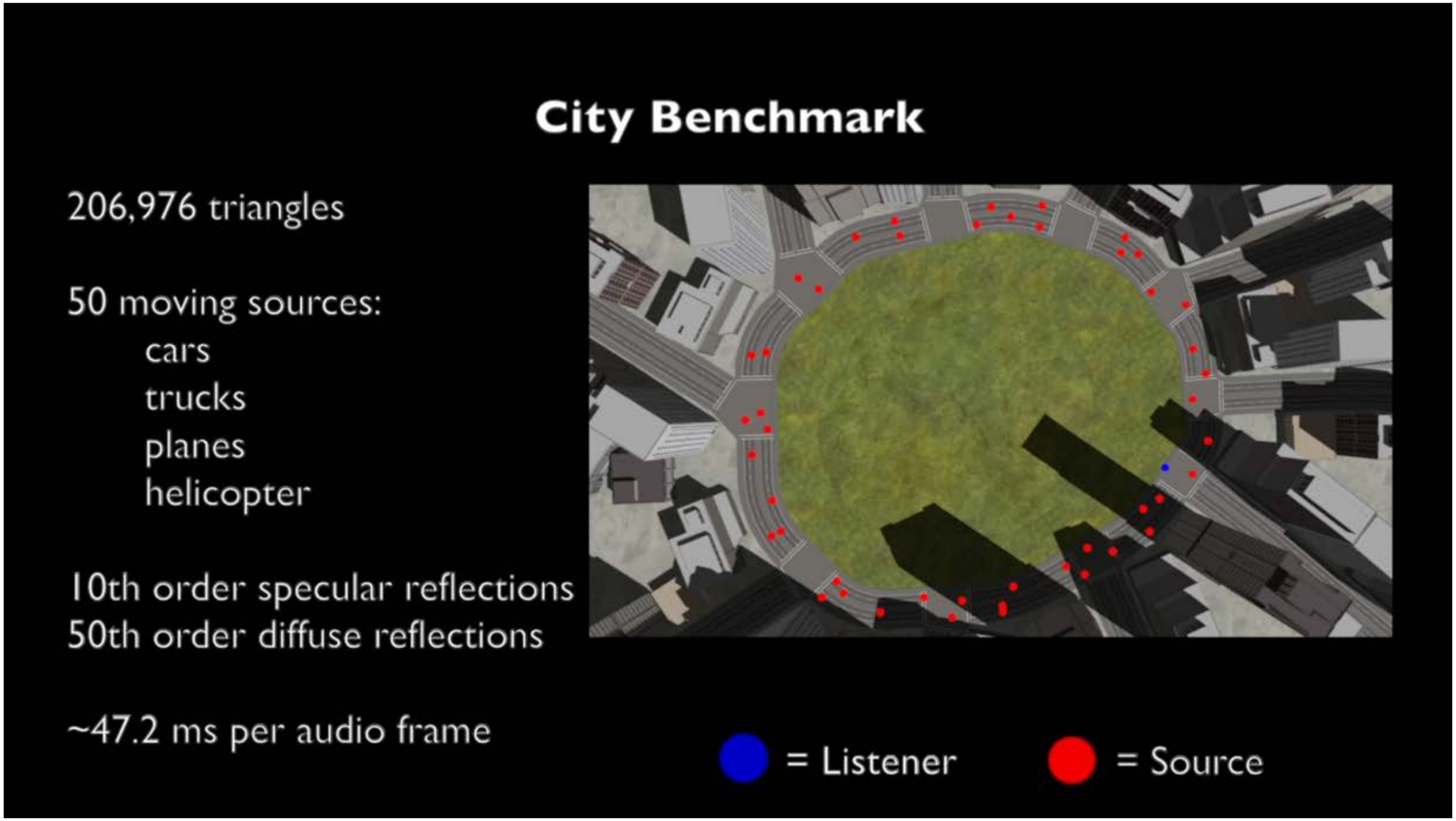
52X improvement

4-core Intel 4770k
CPU

[Chandak et al. 2010; Taylor et al. 2012; Schissler et al. 2014]

Interactive Geometric Propagation

Interactive Propagation
Source Clustering + multiresolution auralization

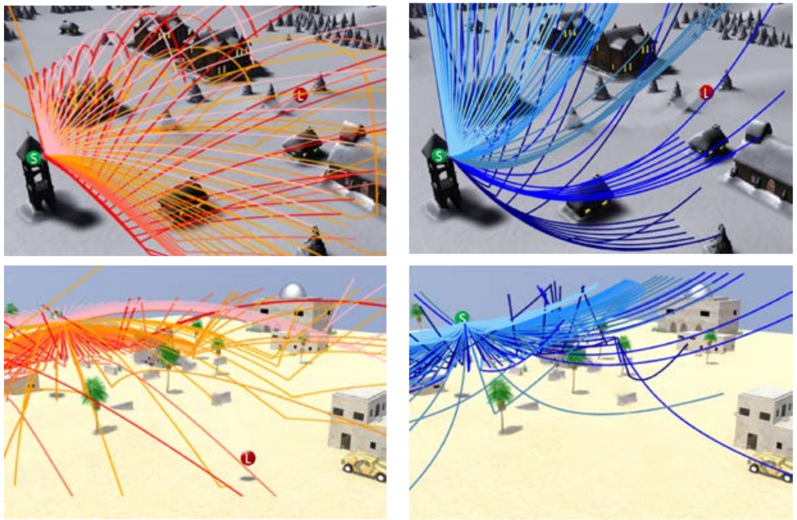


High number of sources: specular + diffuse + diffraction
Source clustering+ rendering on a multi-core PC

Non-Linear Ray Tracing: Participating Media

Tracing 10k rays for 3 orders of reflections

Frame time measured on single thread



Scene	Frame time (ms)	# Surfaces	# medium points	# tetrahedral cells
Desert (median)	***219***	8,000	23,632	144,976
Desert (high)	***369***	16,000	132,742	674,434
Christmas (median)	***259***	8,000	44,862	227,851
Christmas (high)	***443***	16,000	179,382	1,169,353
Reservoir	***233***	4,000	34,771	248,806

Wave-Based Sound Propagation

- Numerical Methods
  - Solve Helmholtz Wave Equation
  - Accurate computation
  - No good computational algorithms for large spaces

$$\frac{\partial^2 p}{\partial t^2} - c^2 \nabla^2 p = F(\mathbf{x}, t)$$

- Methods
  - Finite Element Methods [Otsuru,2004]
  - Boundary Element Methods [Ciskowski,1993]
  - Finite Difference Time Domain [Kunz et al.1993]
  - Digital Waveguide Mesh (DWM) [Savioja et al. 1994]
  - Rectangular Domain Decomposition [Raghuvanshi et al. 2008,2010,2014]
  - Equivalent Source Method [Mehra et al. 2012, 2013,2014]
  - Parallel ARD Solvers on large cluster [Morales et al. 2015, 2017]

Current State of the Art: XAudio2

Distance cue only

Geometric Sound Propagation: Ray Tracing

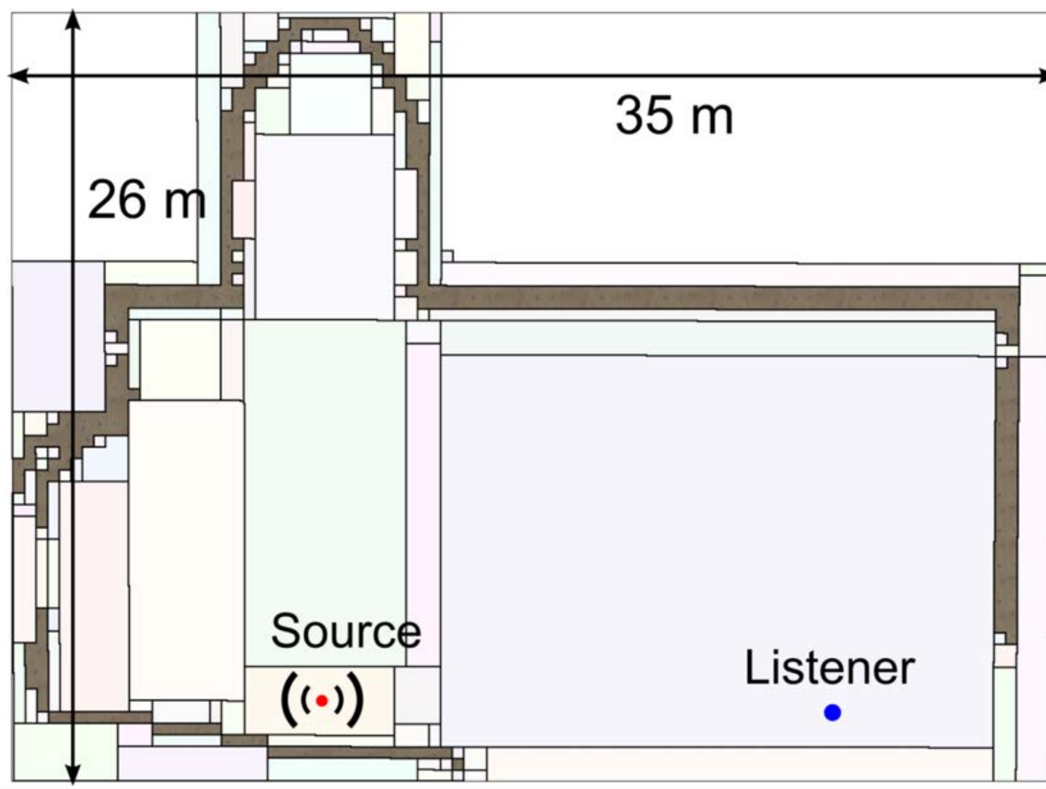
No sound when blocked

Our Wave-Based Solution

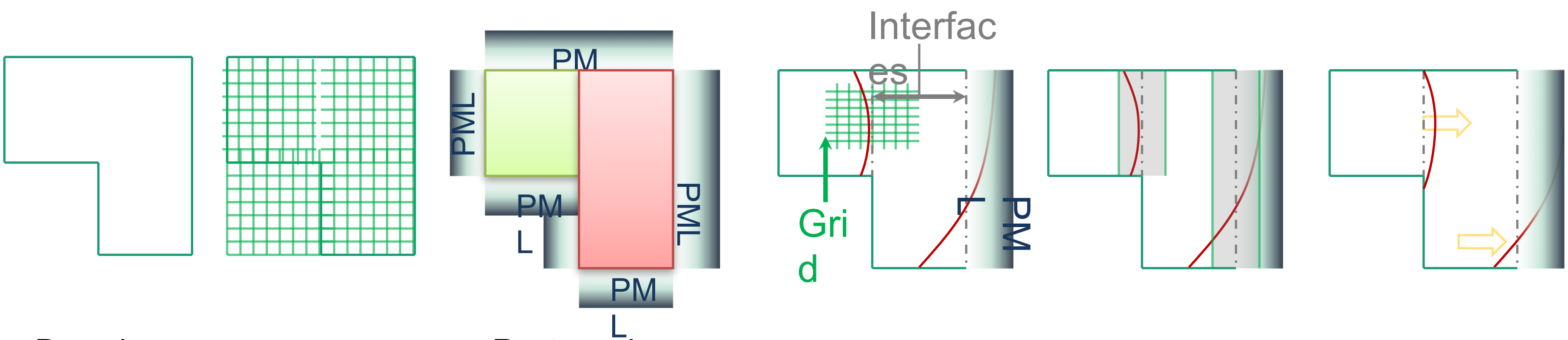
All Wave Effects: Diffraction, Scattering, etc. (our work)

Adaptive Rectangular Decomposition

- Wave equation can be solved very efficiently on a rectangular domain
- Decompose complex domain into rectangles



Adaptive Rectangular Decomposition (ARD)

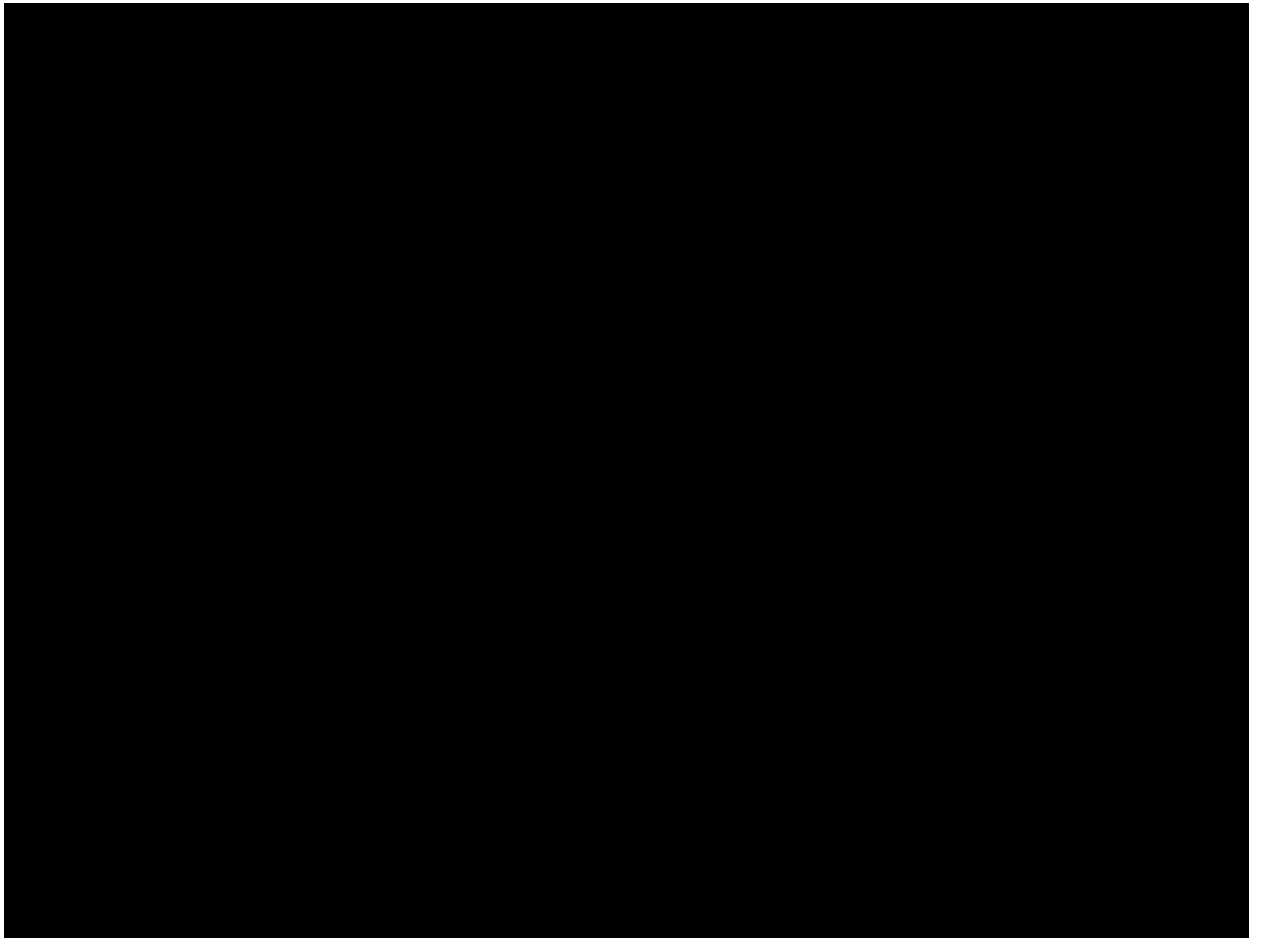


Domain | Voxelization | Rectangular Decomposition | Current Field (t) | Interface handling and DCT | Mode Update and IDCT (t+dt)

Pre-processing

Runtime Simulation

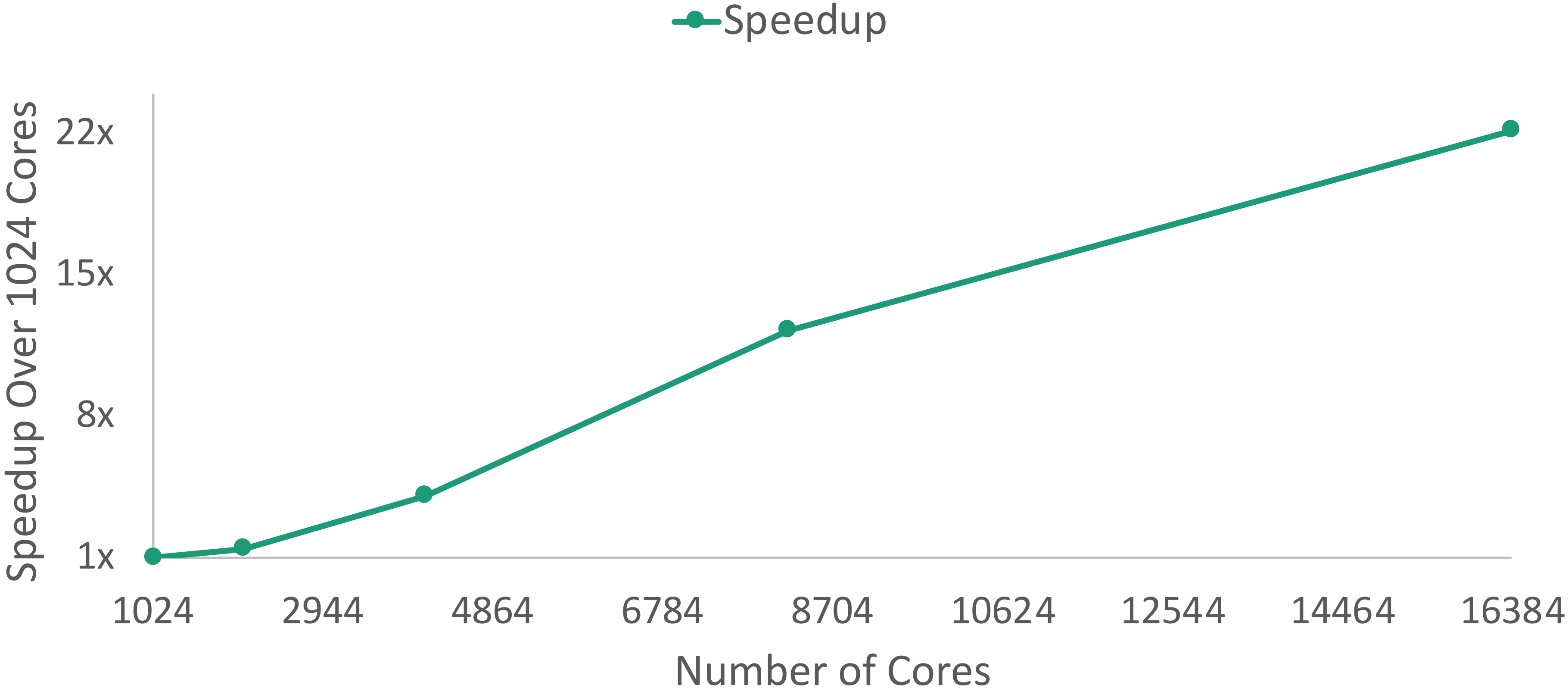
Fast Numeric Propagation



Fast Numeric Propagation

Massive Parallelization: ARD

[Morales et al. 2015]



Equivalent source technique

- Wave-based propagation in large scenes
  - scene spanning hundreds of meters
- Physically accurate propagation
  - all wave-effects e.g. diffraction, interference
- Fast, memory efficient runtime
  - 10s of ms compute, 20-60 MBs

Equivalent source technique

- Helmholtz equation

$$\nabla^2 P + \frac{\omega^2}{c^2} P = 0 \quad + \quad \text{B.C.}$$

$P = P(x, \omega)$ is pressure field in frequency-domain

B.C are boundary conditions

ω is angular frequency

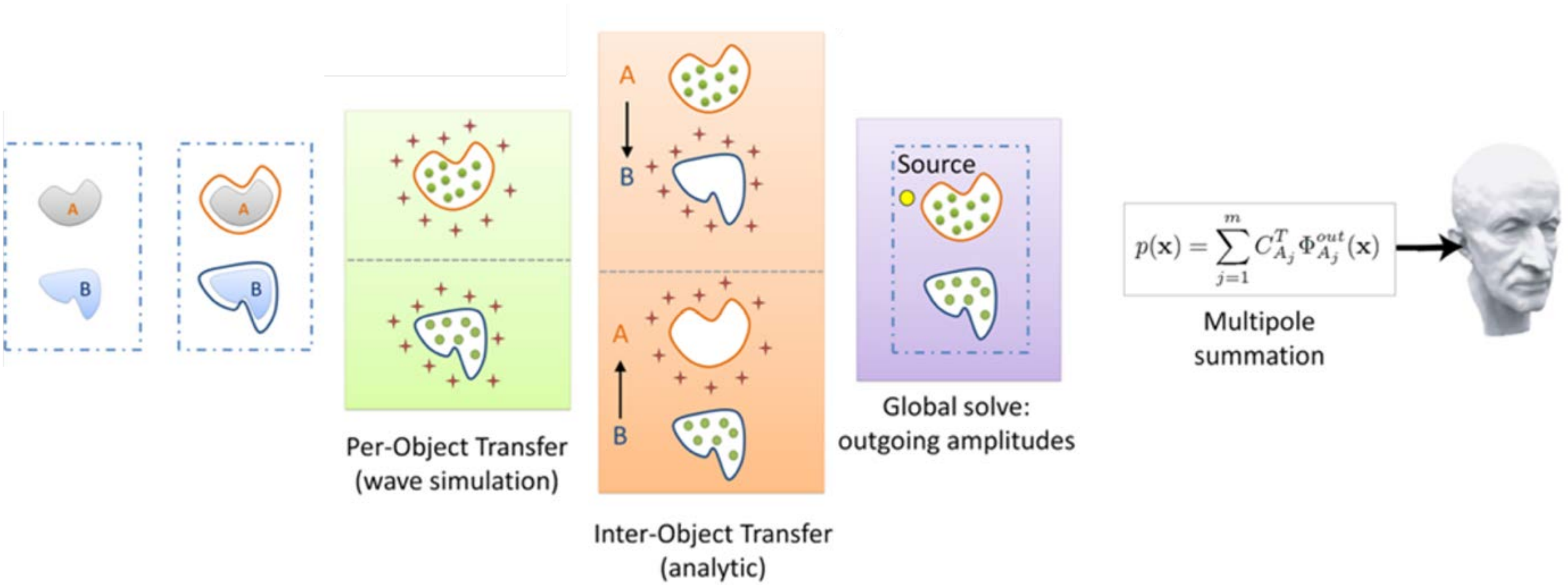
x is point in 3D space

Overall Approach

Object-centric sound propagation (large scenes)

Precomputation

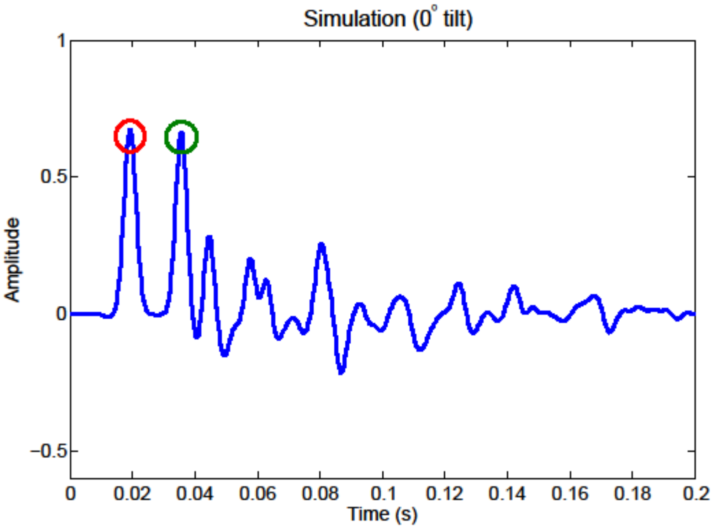
Runtime

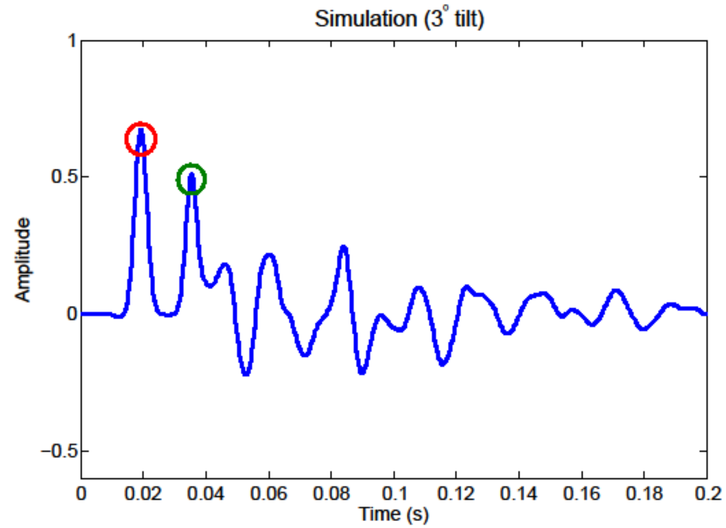


Twilight Epiphany: Validation

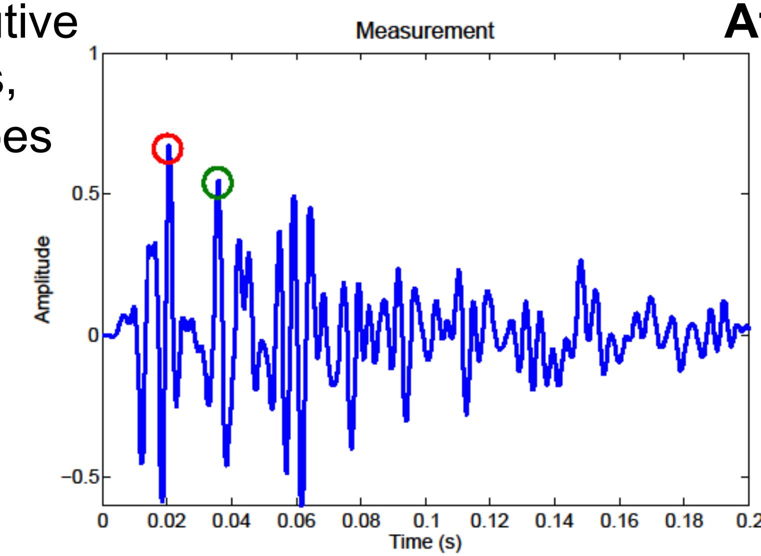
- Twilight Epiphany
  - Art installation at Rice University
- Light show at dusk and sunset
- Used for musical performances, 3D sound art, and a music lab by Music conservatory at Rice University

Twilight Epiphany





Without tilt, consecutive high-amplitude peaks, indicating flutter echoes



After tilting walls, secondary peak amplitude reduced, mitigating flutter echoes

Measured data contains secondary peak with reduced amplitude

Technology Transfer

- http://www.impulsonic.com/
  - Windows PC with headphones
  - Binaural sound with Oculus Rift
  - Ported on Android and other platforms
  - Unity Plugins

  - **Acquired by Valve in Nov 2016**
  - **Our technology is available as Valve Audio SDK**

Sound Simulation: Direct vs. Inverse

- Direct problems: Given the scene description, source and listener location, compute the pressure field

- Inverse: Given the recorded audio signal and partial scene information, use for audio scene understanding + scene reconstruction

Sound Simulation: Direct vs. Inverse

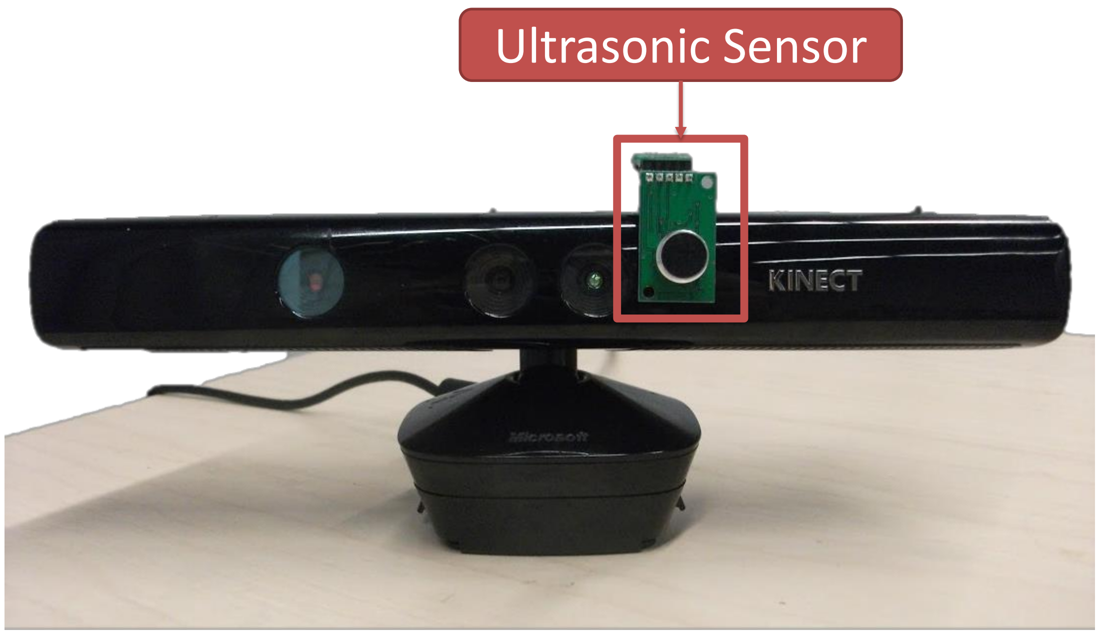
- **Direct problems**: Given the scene description, source and listener location, compute the pressure field
- **Inverse**: Given the recorded audio signal and partial scene information, use for audio scene understanding + scene reconstruction/analysis (<u>*computer vision for sound*</u>)

Sound Simulation: Direct vs. Inverse

- Direct problems: Given the scene description, source and listener location, compute the pressure field
- Inverse: Given the recorded audio signal and partial scene information, use for audio scene understanding + scene reconstruction
- The field of "acoustic scene analysis" is not new
- But we have new tools:
  - Computer vision methods
  - Computing power (GPUs, cloud computing)
  - Machine learning methods
  - Interactive acoustic simulation technology

Multi-modal Cameras

- Combine 2D cameras, depth-sensors + acoustic sensors
- Inverse acoustics: computer vision for audio



3D Reconstruction in the Presence of Glass (CVPR 2015)

Multi-modal Cameras

- Combine 2D cameras, depth-sensors + acoustic sensors
- Inverse acoustics: computer vision for audio

3D Reconstruction in the Presence of Glass (CVPR 2015)

Augmented Reality: Need Sound Effects

Virtual sounding object

Microsoft Hololens™

Acoustic Simulation Prerequisites

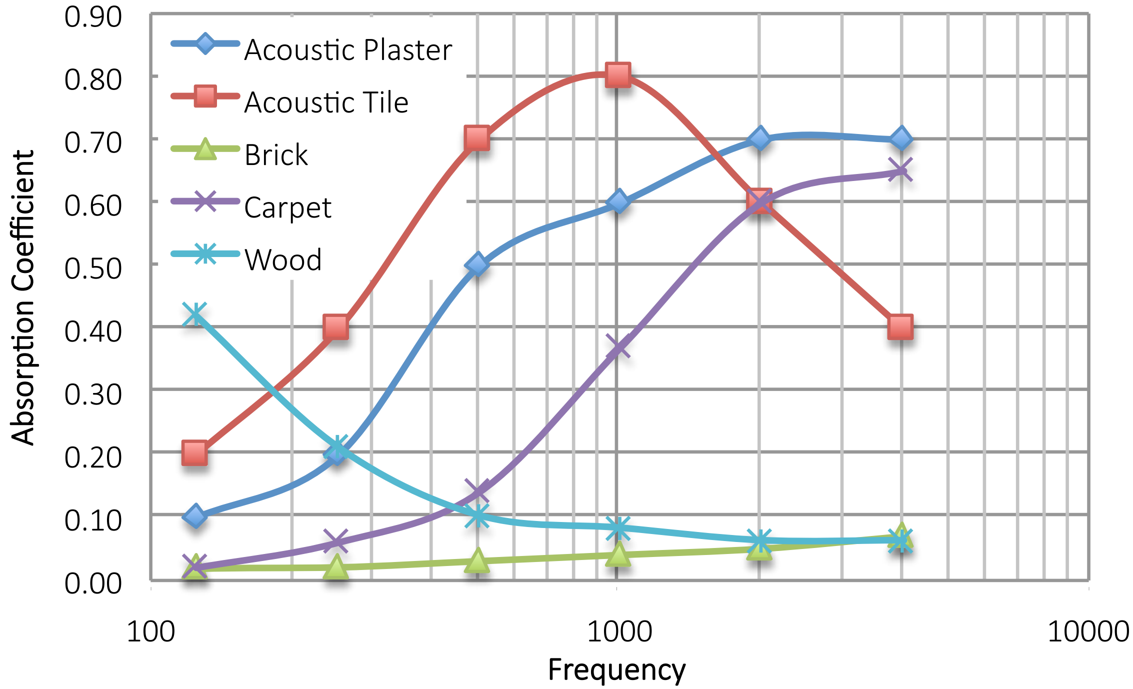
To compute sound propagation, we need at least:

- Source locations
- Listener locations
- **3D model**
- **Acoustic material properties**

} Use computer vision methods

Acoustic Materials

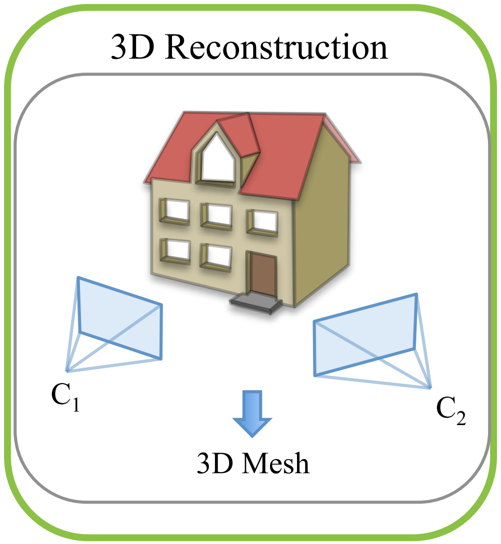
Sound Absorption Coefficient – fraction of incident pressure absorbed



[Egan 1988]

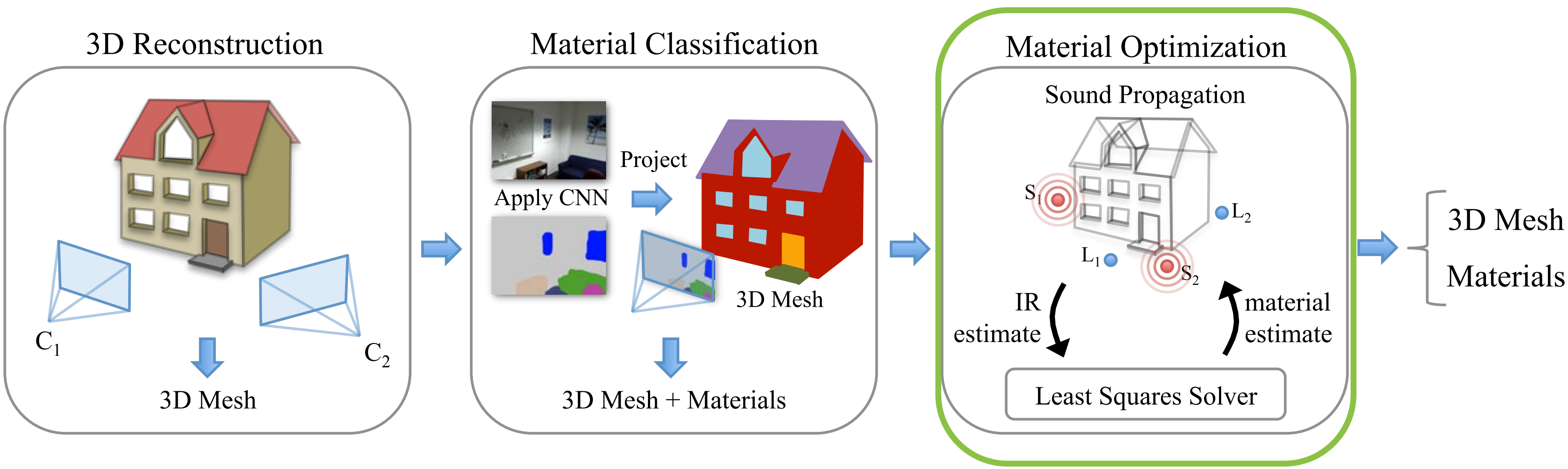
Material Pipeline - Reconstruction

1. Reconstruct 3D mesh from multiple camera viewpoints

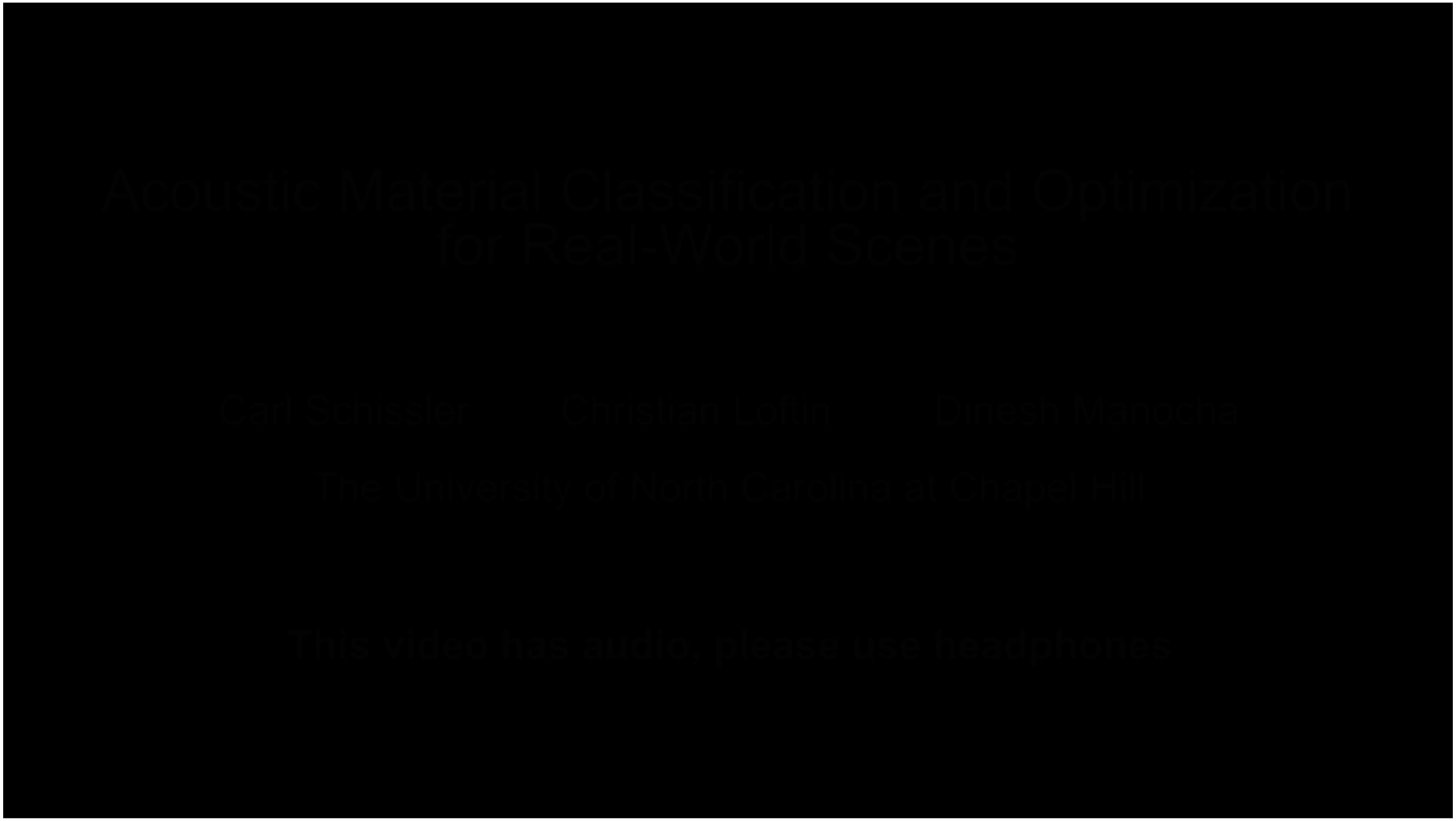


Material Pipeline - Optimization

3. Improve material estimates iteratively using measurements and simulation



Acoustic Material Classification for AR/VR



Noise Reduction for Speech Recognition using Simulation

- Reverberation and noise is a major challenge in Automated Speech Recognition (ASR) algorithms.
- In complicated environments, the placement of sound receivers will affect the intelligibility of speech.
- Given such a scene setup, how to find the best receiver location that maximizes the overall speech intelligibility?

Conclusions

- Sound Simulation and Propagation
  - Considerable Research in the last two decades
  - Good solutions based on geometric and wave-based methods

- Inverse Sound Propagation: Computer vision for sound
  - Scene reconstruction with audi0—visual methods
  - Capturing audio material properties

  - Ongoing Work:
    - Audio-based localization
    - Audio-Visual SLAM

Acknowledgements

- Army Research Office
- DARPA
- Intel
- Microsoft
- National Science Foundation
- NVIDIA
- Oculus
- RDECOM
- Samsung

Questions: Contact at dm@cs.umd.edu