Ambient sound propagation

ZECHEN ZHANG CORNELL UNIVERSITY
NIKUNJ RAGHUVANSHI MICROSOFT RESEARCH

JOHN SNYDER MICROSOFT RESEARCH
STEVE MARSCHNER CORNELL UNIVERSITY

Presenter: Denis Thy

Review: Refractive Radiative Transfer Equation

In area where the refractive index changes, the light path might not be a straight line.

General transporting equation

Based on Hamiltonian optics

Use radiance as fundamental quantity

Contents

Introduction

Previous work

Incoherent wave source

Encoder

Runtime

Introduction

Related works

Computer Graphics

- Geometric acoustics
- Wave model

Wave solvers

- Finite-Difference Time-Domain (FDTD) Simulation

Coherent sources

Overview

Precomputation of FDTD wave simulation [Taflove and Hagness 2005]

Incoherent signal synthesis

Parameters fields encoding

Runtime rendering

Wave Simulation

Finite Difference Time Domain (FDTD) numerical method to simulate wave propagation

- Both space and time are divided into discrete segments. Space is segmented into box-shaped cells. → **Space is divided into voxels**

Simple to implement

Numerical dispersion: Different frequencies travel at different speed

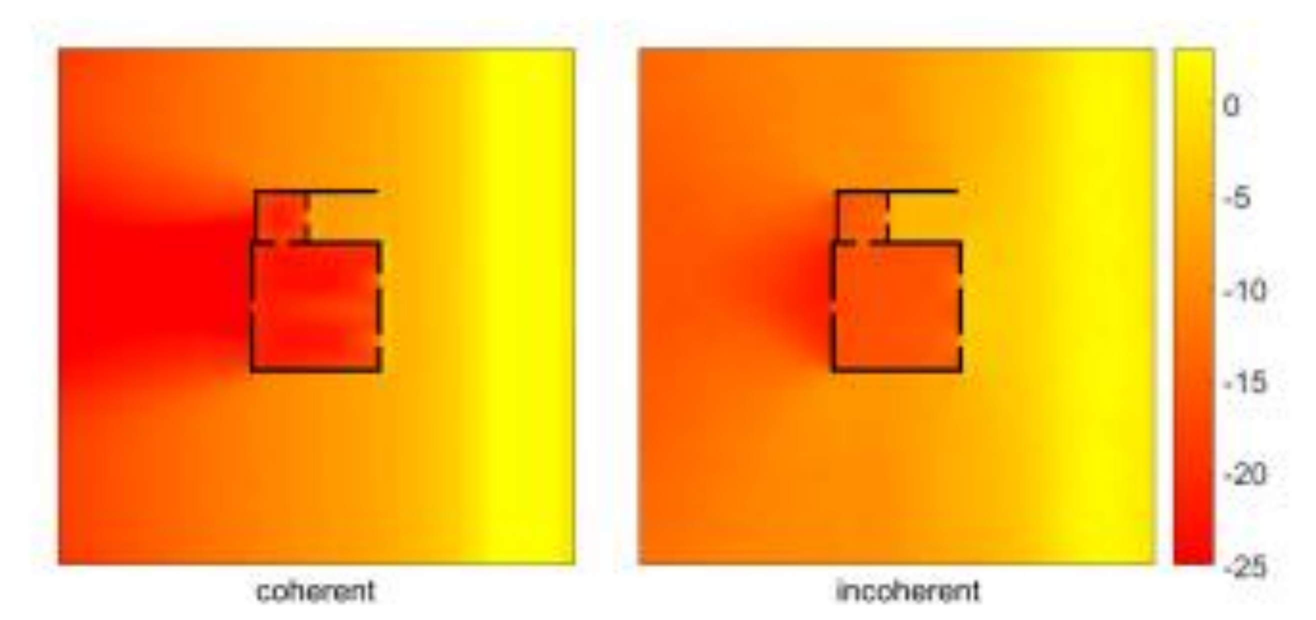
Numerical dissipation: High frequencies are attenuated

Incoherent signal synthesis

Incoherent sound sources

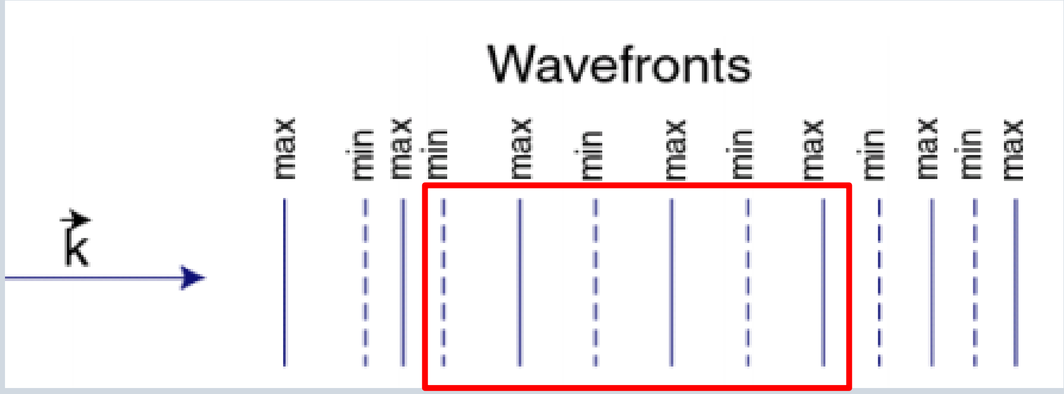
Coherence is the tendency of wave field observations at different times and places to be correlated.

Two wave sources are perfectly coherent if they have a **constant phase difference** and the **same frequency** → **Interferences**



Temporal incoherence

Temporal coherence is a measure of the correlation of wave's phase at different points along the direction of propagation – it tells us how monochromatic a source is.

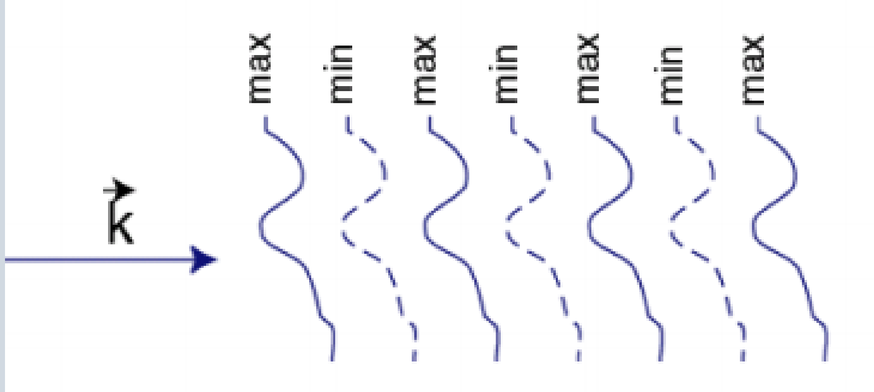


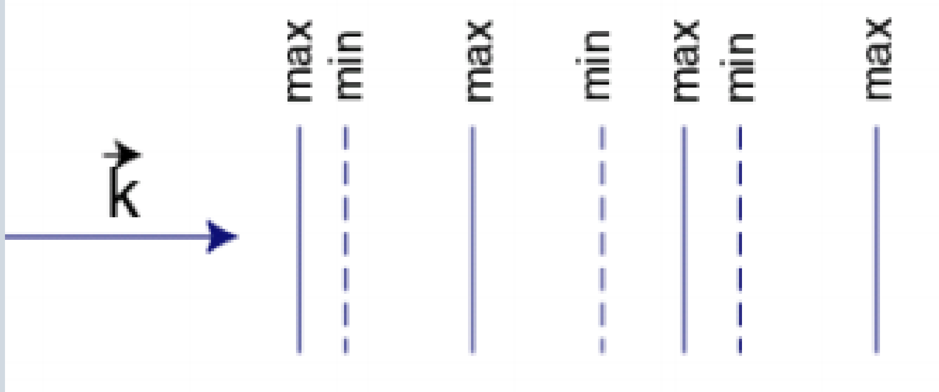
To ensure temporal incoherence they use ideal white noise in discrete Fourier domain → constant amplitude with **independently random phase** at each frequency bins

Image from http://users.ntua.gr/eglytsis/OptEng/Coherence_p.pdf

Spatial incoherence

Spatial coherence is a measure of the correlation of a wave's phase at different points transverse to the direction of propagation - it tells us how uniform the phase of a wavefront is.





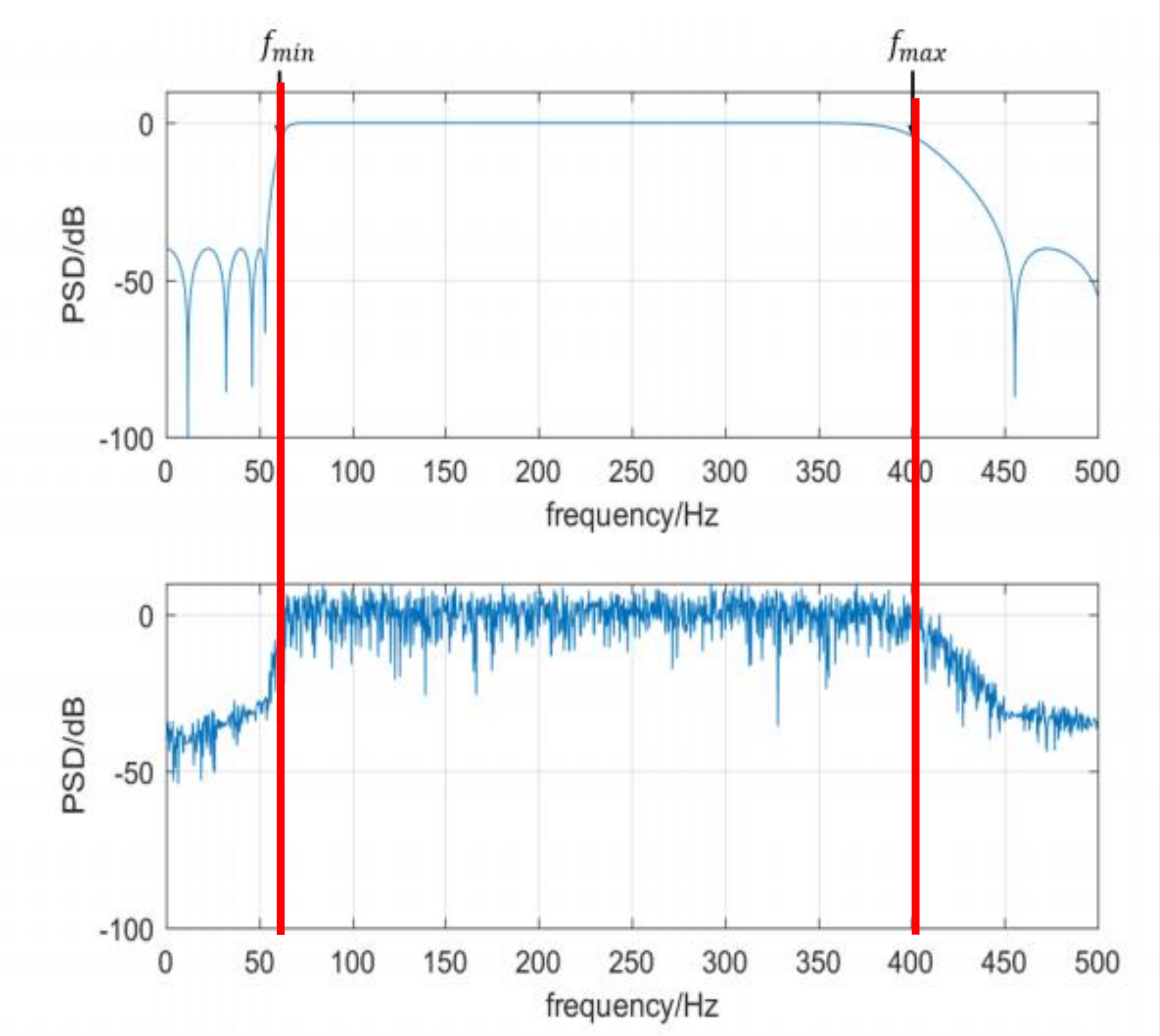
To ensure spatial incoherence, they generate **independent signals** in each FDTD grid cell covered by the source.

Images from http://users.ntua.gr/eglytsis/OptEng/Coherence_p.pdf

Online synthesis

Fourier synthesis: precise but costly in memory

Filtering zero-mean white noise on-the-fly during simulation



Encoder

Encoder

3D grid of low-order spherical harmonics coefficient to represent the directional distribution of acoustic power

Computes flux density at each time step

Accumulates the spherical harmonics coefficients

Compression of the perceptual parameters

Acoustic flux density

Time-averaged directional power distribution for a point

$$E(\Theta) = \frac{1}{T} \int_0^T p^2(t)\delta\left(\Theta - \hat{f}(t)\right) dt$$

- $\hat{f}(t)$ *is a unit vector containing information about the flux direction*
- $p^2(t)$*is the instantaneous power*
- Θ is a direction
- δ *is the dirac function*

Acoustic flux density

Time-averaged directional power distribution for a point

$$E(\Theta) = \underbrace{\frac{1}{T}\int_0^T}_{\text{Average over time}} \underbrace{p^2(t)\,\delta\left(\Theta - \hat{f}(t)\right)}_{\text{Power in one direction}} dt$$

- $\hat{f}(t)$ *is a unit vector containing information about the flux direction*
- $p^2(t)$ *is the instantaneous power*
- Θ is a direction
- δ *is the dirac function*

Spherical Harmonics projection

Encode directional power distribution as a smooth spherical function using spherical harmonics

$$E_{l,m} = \frac{1}{T}\int_0^T p^2(t) Y_{l,m}\left(\hat{f}(t)\right) dt$$

$$Y_{l,m} \text{ are the spherical harmonics}$$
$$E_{l,m} \text{ are the encoded SH coefficients}$$

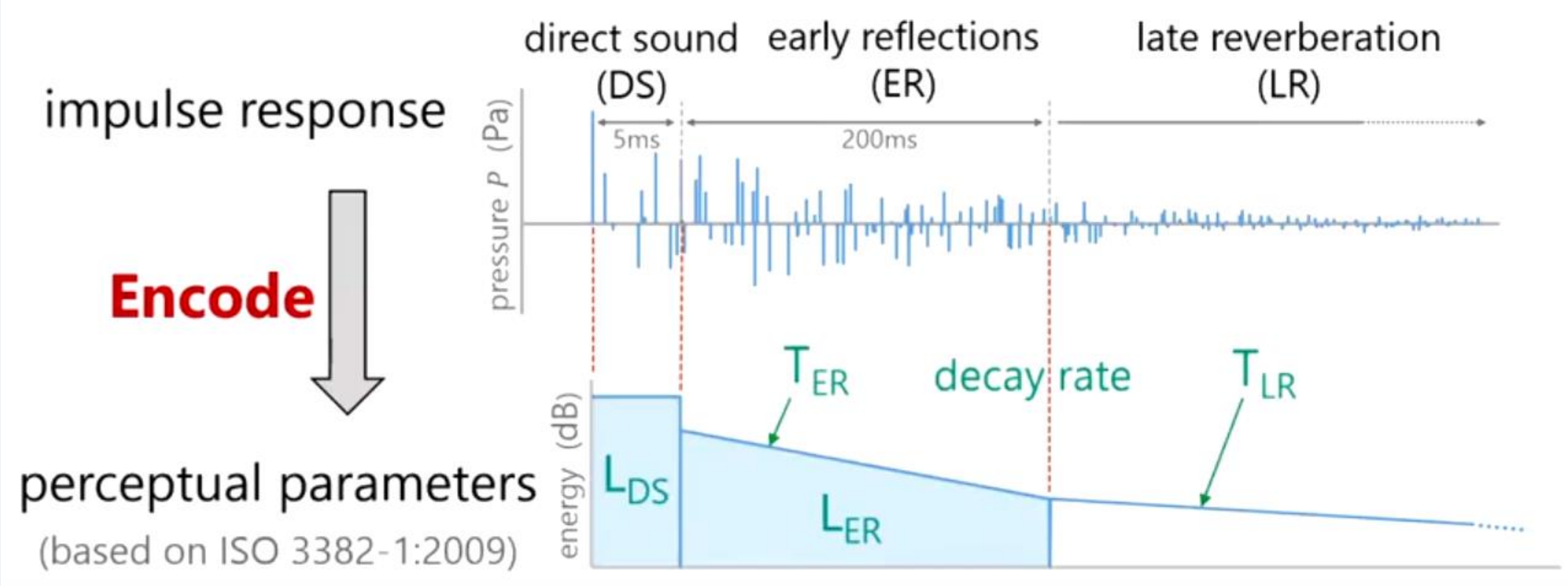
Input power distribution

They are able to reconstruct the input power distribution using the formula:

$$E(\Theta) \approx \sum_{l=0}^{n-1} \sum_{m=-l}^{l} E_{l,m} Y_{l,m}(\Theta)$$

→ Summing all the SH

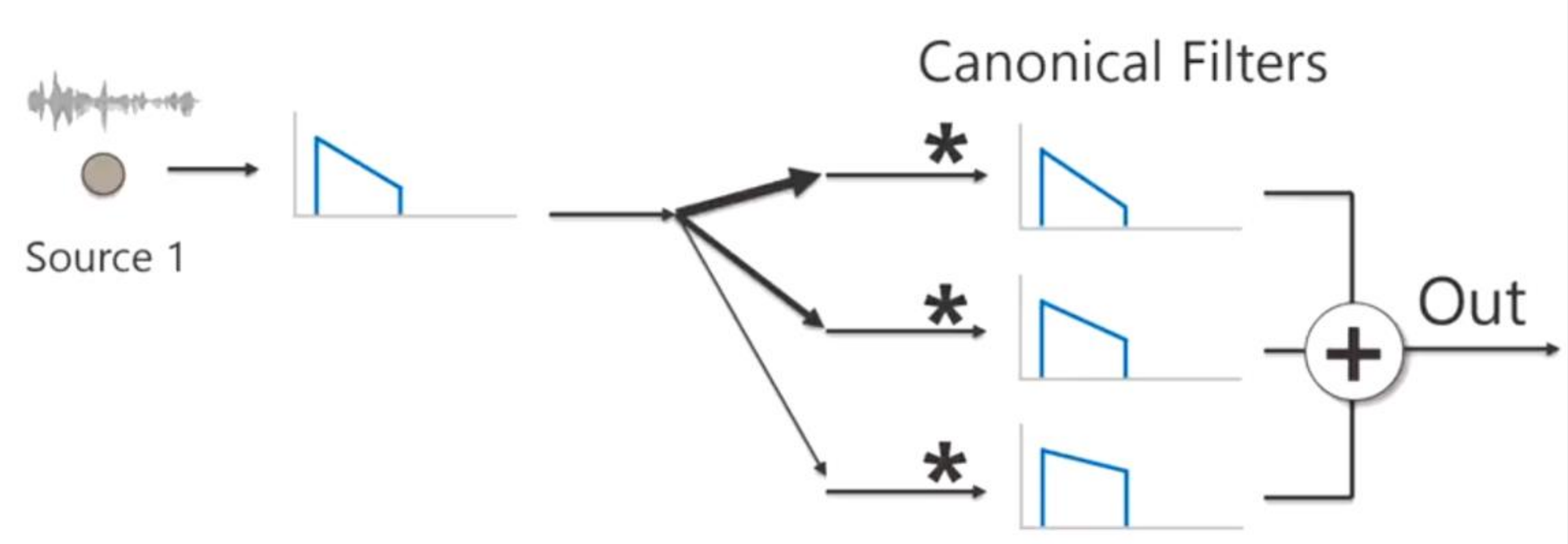
Compression [Raghuvanshi et al. 2014]



Taken from [Raghuvanshi et al. 2014]

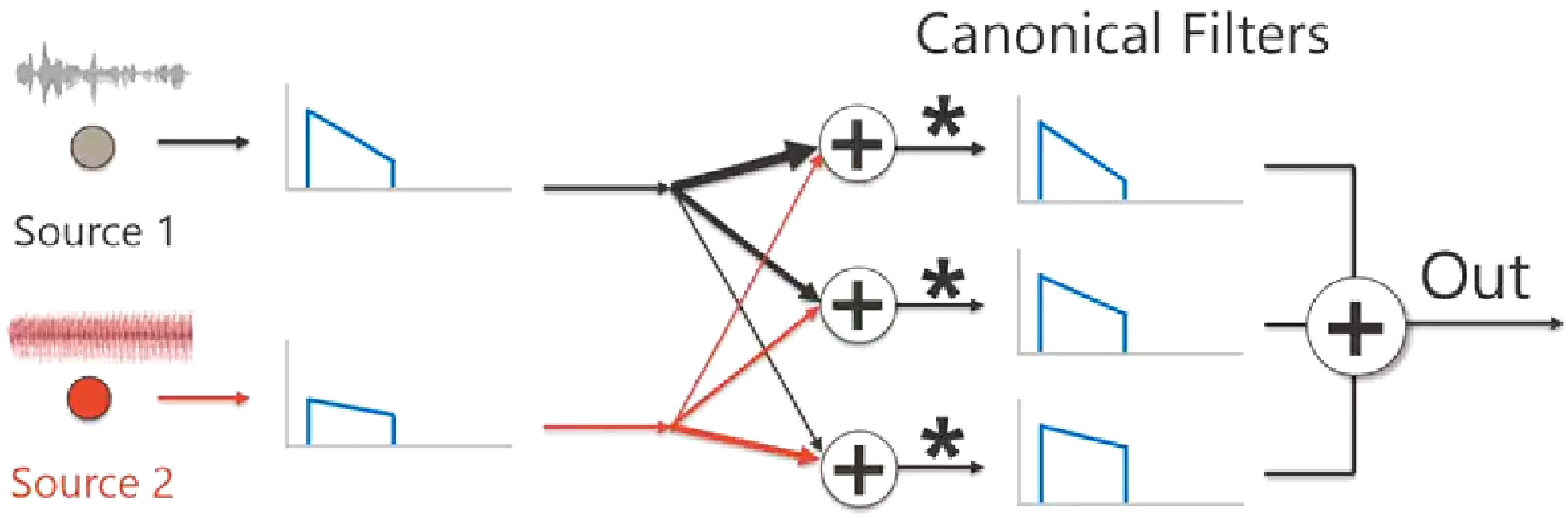
Runtime

Runtime Decoding [Raghuvanshi et al. 2014]



Taken from [Raghuvanshi et al. 2014]

Runtime Decoding [Raghuvanshi et al. 2014]



Taken from [Raghuvanshi et al. 2014]

Spatialization

Human Auditory system

- Interaural Phase Difference (IPD)
- Interaural Loudness Difference (ILD)

Interaural Phase difference

Around 200 frequency bins for each direction

Binaural rendering

- $Source\ * HRTF$ at each frequency bin

Chaotic / Extended sources → IPD cues are less detectable

Interaural Loudness difference

Gain function

$$g(f) = \sqrt{\int_{\Omega} E\left(R^{-1}(\Theta)\right) \|H(\Theta, f)\|^2 d\Theta}$$

$$E \ is \ directional \ power \ distribution$$

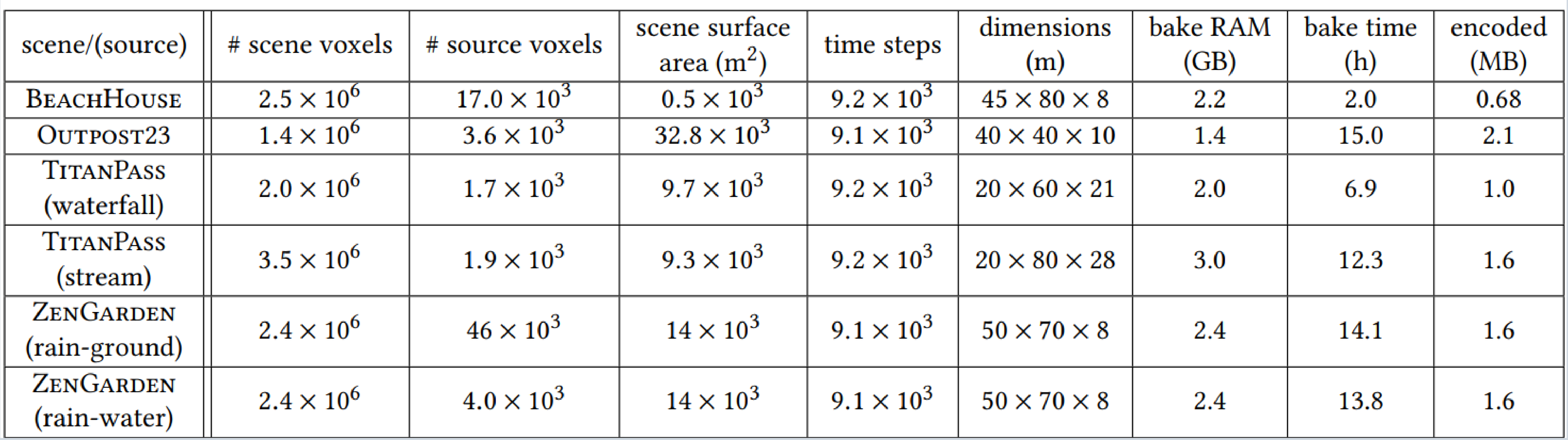
$$\Omega \ is \ direction \ space$$

$$H \ is \ HRTF \ function$$

Results

INTEL I7 CPU @3,70 GHZ AND 32G RAM

Precomputation data



scene/(source)	# scene voxels	# source voxels	scene surface area (m^2)	time steps	dimensions (m)	bake RAM (GB)	bake time (h)	encoded (MB)
BeachHouse	2.5×10^6	17.0×10^3	0.5×10^3	9.2×10^3	$45 \times 80 \times 8$	2.2	2.0	0.68
Outpost23	1.4×10^6	3.6×10^3	32.8×10^3	9.1×10^3	$40 \times 40 \times 10$	1.4	15.0	2.1
TitanPass (waterfall)	2.0×10^6	1.7×10^3	9.7×10^3	9.2×10^3	$20 \times 60 \times 21$	2.0	6.9	1.0
TitanPass (stream)	3.5×10^6	1.9×10^3	9.3×10^3	9.2×10^3	$20 \times 80 \times 28$	3.0	12.3	1.6
ZenGarden (rain-ground)	2.4×10^6	46×10^3	14×10^3	9.1×10^3	$50 \times 70 \times 8$	2.4	14.1	1.6
ZenGarden (rain-water)	2.4×10^6	4.0×10^3	14×10^3	9.1×10^3	$50 \times 70 \times 8$	2.4	13.8	1.6

Conclusion

Ambient sound propagation

- FDTD simulation
- Incoherent source
- Encoder to store perceptual information
- Cheap in memory