Interactive Sound Propagation with Bidirectional Path Tracing

Author: Chunxiao Cao Zhong Ren Carl Schissler Dinesh Manochał Kun Zhou
Speaker: Jongwon Jang

Review of last presentation

Pseudo-marginal Metropolis Light Transport

- Naive Monte Carlo path tracing has limit.
  - Only few paths can reach a light
- Metropolis Light Transport is light path sampling algorithm.
  - Focus on successful eye-to-light ray
- Evaluating MLT is to hard, so they estimate it
  - Using Markov Chain Monte Carlo (MCMC) method, specifically MH algorithm.
- Needs unbiased estimator
  - Build unbiased estimator using unbiased transmittance estimator by ratio tracking

Introduction

Introduction

- Bidirectional Soud Transport (BST)
- Sound and visual rendering have correlation
- Sound rendering
  - Wave-based
    - Accurate but expensive and limited to static scenes
  - Geometric Acoustic (GA)
    - Ray Tracing
    - Can handle complex environments
    - An insufficient number of rays can lead to aliasing artifacts in sound rendering
    - A larger ray budget slows down the performance

There novel contributions include...

- New metric for assessing the quality of stochastic sound propagation
- New algorithm that simulates the sound propagation by bidirectional path tracing (BDPT) with multiple importance sampling (MIS)
- Using signal-to-noise ratio (SNR) criterion to analyze different sound propagation

Bidirectional Sound Transport

1. Background

Acoustic Transport Equation

$$L(\mathbf{x}' \to \mathbf{x}, t) = L_0(\mathbf{x}' \to \mathbf{x}, t) + \int_\Omega L_s(\mathbf{x}' \to \mathbf{x}, t) * M(\mathbf{x}'' \leftrightarrow \mathbf{x}', t) dA_{\mathbf{x}''},$$

[Siltanen et al. 2007]

L0: Radiance emitted by x’

$$L_s(\mathbf{x}' \to \mathbf{x}, t) = L(\mathbf{x}'' \to \mathbf{x}', t) G(\mathbf{x}'' \leftrightarrow \mathbf{x}') \rho(\mathbf{x}'' \to \mathbf{x}' \to \mathbf{x}).$$

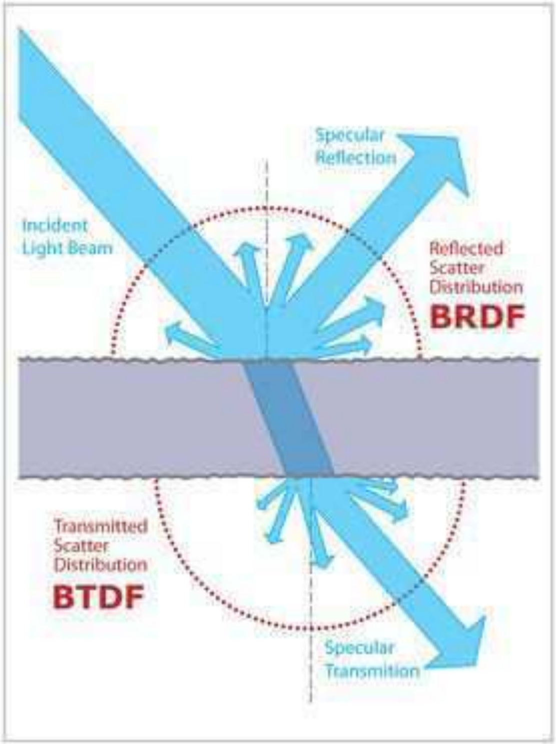
G: Geometry Term (Energy dispersion and occlusion during propagation)

ρ: .Bidirectional Scattering Distribution Function (BSDF)

M: Media Term (Energy Absorption and time delay caused by propagation media)

$$M(\mathbf{x}'' \leftrightarrow \mathbf{x}', t) = e^{-\alpha|\mathbf{x}'' - \mathbf{x}'|} \delta(t - \frac{|\mathbf{x}'' - \mathbf{x}'|}{c})$$

Acoustic Transport Equation



Acoustic Transport Equation

$$L = L_0 + TL$$

$$L = (\mathrm{Id} - T)^{-1} L_0 = \sum_{i=0}^{\infty} T^i L_0$$

Inverse of (Id-T) can be expressed with Neumann Series.

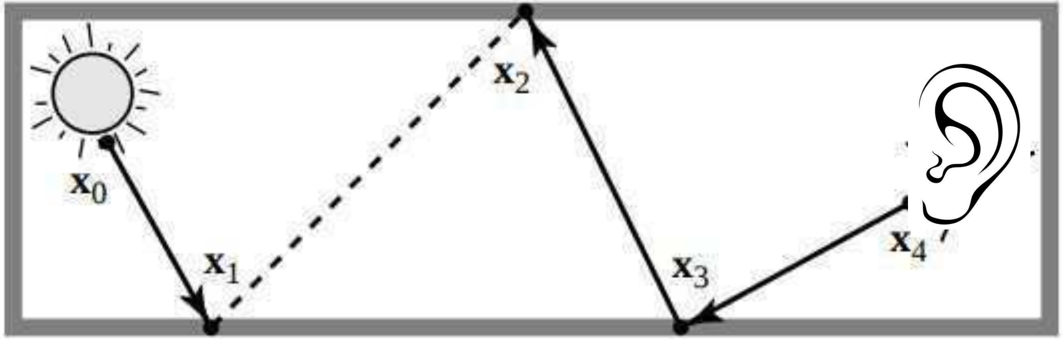
$$\frac{1}{1-x} = 1 + x + x^2 + \cdots$$

Evaluating $T^i L_0$ with Monte-Carlo Path Tracing

$$E[\frac{f(X)}{p(X)}] = \int_\Sigma f(x)\frac{d\mu_1}{d\mu_2}d\mu_2 = \int_\Sigma f(x)d\mu_1$$

- μ1: Path Space of bounce i
- μ2: Probability Measure of Path Generation
- f(X): Energy Impulse generated by sound propagation along path
- p(X): Probability of Generation

Bidirectional Path Tracing



$$p(\mathbf{x}_0 \cdots \mathbf{x}_k) = p_g(\mathbf{x}_0 \to \mathbf{x}_1) G(\mathbf{x}_0 \leftrightarrow \mathbf{x}_1) \cdot \prod_{i=1}^{k-1} p_f(\mathbf{x}_{i-1} \to \mathbf{x}_i \to \mathbf{x}_{i+1}) G(\mathbf{x}_i \leftrightarrow \mathbf{x}_{i+1})$$

pg: the probability density for the source to generate a path in the direction x0 → x1

pf: the probability density for the outgoing path to be in the direction xi → xi+1 when given the incident path direction xi-1 → xi

Bidirectional Path Tracing

$$p(\mathbf{x}_0 \cdots \mathbf{x}_k) = \prod_{i=0}^{k-1} V_i G_{i,i+1}$$

(s,t)-connection

$$\begin{aligned} p_{s,t}(\mathbf{x}_0 \cdots \mathbf{x}_{s+t+1}) &= \prod_{k=0}^{s} V_k G_{k,k+1} \prod_{k=s+1}^{s+t+1} G_{k-1,k} V_k \\ &= \frac{\prod_{k=0}^{s+t+1} V_k \prod_{k=0}^{s+t} G_{k,k+1}}{V_s G_{s,s+1} V_{s+1}}. \end{aligned}$$

Multiple Importance Sampling

x: forward subpath

y: backward subpath

$$T^i L_0 = E\left[\frac{1}{N}\sum_{j\geq 0, i-j\geq 0}\frac{1}{c_j}\sum_{k=0}^{n_j}\frac{w_{j,i-j}(X_{j,k})f(X_{j,k})}{p_{j,i-j}(X_{j,k})}\right]$$

$$w_{k,i-k}(X) = \frac{c_k p_{k,i-k}(X)}{\sum_{j\geq 0, i-j\geq 0} c_j p_{j,i-j}(X)}$$

Sound vs Visual Rendering

- The spatial resolution of the human auditory system is far lower than that of the visual system.
- Temporal information is essential for acoustic simulation.
- Controlling the quality in the temporal dimension is an important task for sound propagation.
    - need an indirect method to control the temporal distribution of samples

2. SNR Metric for Stochastic Sound Propagation

Which metric they use?

- There is no-well accepted standard for evaluating the quality of acoustic rendering.
- Metric should be easy to valuate and correlate well with perceptual evaluation.
- Using energy response as the quality metric
  - Sound simulation algorithm deals with energy response
  - Uses acoustic transport equation
- So they use the **signal-to-noise ratio(SNR)** of the energy response!

Signal-to-noise Ratio

$$\mathrm{SNR}[X] = \frac{E[X]}{\sigma[X]}, \mathrm{SNR}_{\mathrm{dB}}[X] = 10\log_{10}\left(\frac{E[X]}{\sigma[X]}\right)$$

Simple propagation

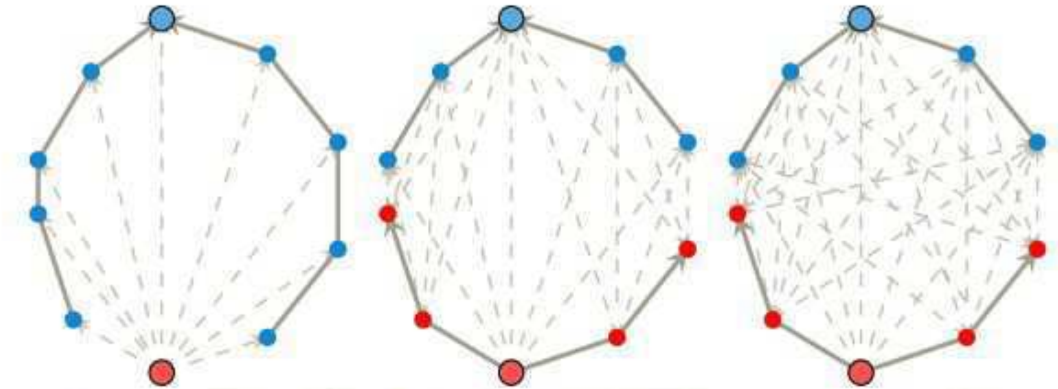
Easy to handle cases that differentiate between early and late reflections

3. Bidirectional Sound Transport Algorithm

Existing GA algorithms

- Generate paths from the source or listener to evaluate Monte Carlo Integration
- To improve the validity of generated paths, diffuse rain is used to build connections between the source/listener with the hit points on the emitted paths.
    - one of the BDPT with zero bounces on either source or listener side
    - One estimator for each $T^i L_0$
- Where different paths have a radiance with a large difference, diffuse rain may generates energy responses with low SNR.

Path Generation



(a) backward PT+DR (b) traditional BDPT (c) our method

Red: Backward Subpaths

Blue: Forward Subpaths

Sample Allocation among Integrals

$$x_{n,i+1} = \frac{\alpha}{M} \sum_{m=1}^{M} \frac{\frac{\sigma_{mn}^2}{x_{n,i}}}{\sum_{k=1}^{N} \frac{\sigma_{mk}^2}{x_{k,i}}} + (1-\alpha) x_{n,i}.$$

$$\sigma_i^2 = \begin{cases} \sigma_0^2, & Q_0 > Q^* \\ \dfrac{Q_{i-1}\sigma_i^2 + Q_0\sigma_0^2}{Q_{i-1} + Q_0}, & Q_{i-1} + Q_0 < Q^* \\ \gamma\sigma_i^2 + (1-\gamma)\sigma_0^2, & \text{otherwise.} \end{cases}$$

$$Q_i = \begin{cases} Q_0, & Q_0 > Q^* \\ Q_0 + Q_{i-1}, & Q_{i-1} + Q_0 < Q^* \\ Q^*, & \text{otherwise.} \end{cases}$$

where Q^* is a predefined quality standard, and γ is given as

$$\gamma = \frac{Q_{i-1} - \sqrt{Q_{i-1}Q_0(\frac{Q_{i-1}+Q_0}{Q^*} - 1)}}{Q_{i-1} + Q_0}.$$

Algorithm Summary

Algorithm 1 Bidirectional Sound Transport

Input:
Sound environment. subpath budget M, sample budget N, Maximum path bounce K.
Initialize:
Sample probability x_n of bounce n as $1/K$,
variance estimation $\sigma^2_{m,n}$ as 0.
Output: Energy response $E(\mathbf{x}, \omega, t)$ of each frame.

for each frame **do**

    // Preparation
    Allocate samples among bounces according to x_n;
    for $i = 0$ to K **do**
        Allocate samples among different estimators of $T^i L_0$;
    end for

    // Trace step
    Trace new subpaths;

    // Connect step
    Connect between the subpaths to generate samples;
    Evaluate $f(\cdot)$, $p(\cdot)$ and MIS weight $w(\cdot)$ for each sample;
    Evaluate ER$(\mathbf{x}, \omega, t)$ according to Eq. (7) and Eq. (12);

    // Optimize step
    Update $\sigma^2_{m,n}$ according to Eq. (18);
    Update x_n according to Eq. (17).
end for

Interactive Sound Propagation

Diffuse Cache

- Diffuse Cache
  - exploits the temporal coherence of sound propagation to improve the render quality
  - maintains a moving average of certain intermediate variables used in sound propagation to quickly update the contributions
- Consists of a table of path node information and a queue of cache entries
- The information of a path node includes its position, normal and material index.
- stores the value f(X)/p(X), the index of nodes, and the direction (if valid)
- Only G (geometry term) requires recomputation

Path Cache

- Generating a new node on a path is always more costlier than connecting two nodes
- In scenes with static objects, one could reuse the subpaths.
- This only works with BDPT-based algorithms as they generate subpaths from both sides.
- Can be biased
- “Static” → Reuse

Results and Analysis

The benchmarks

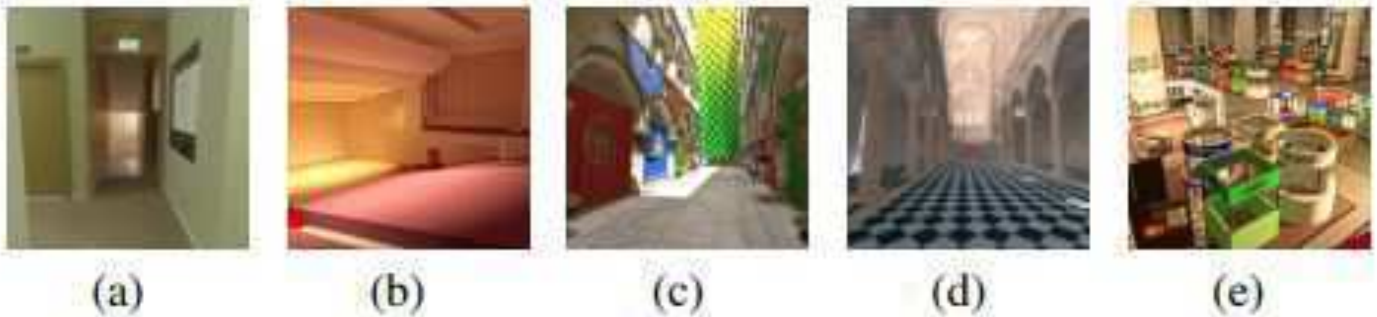
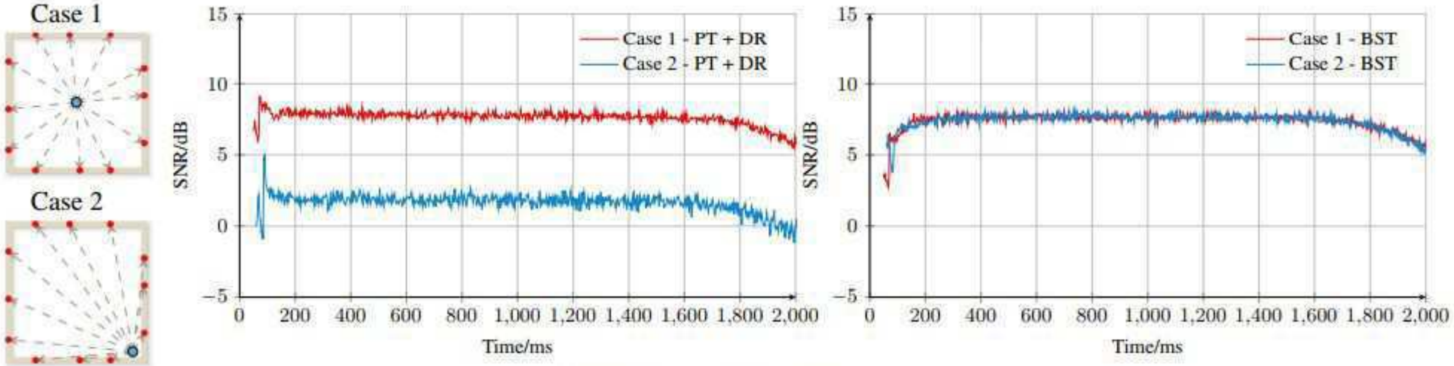


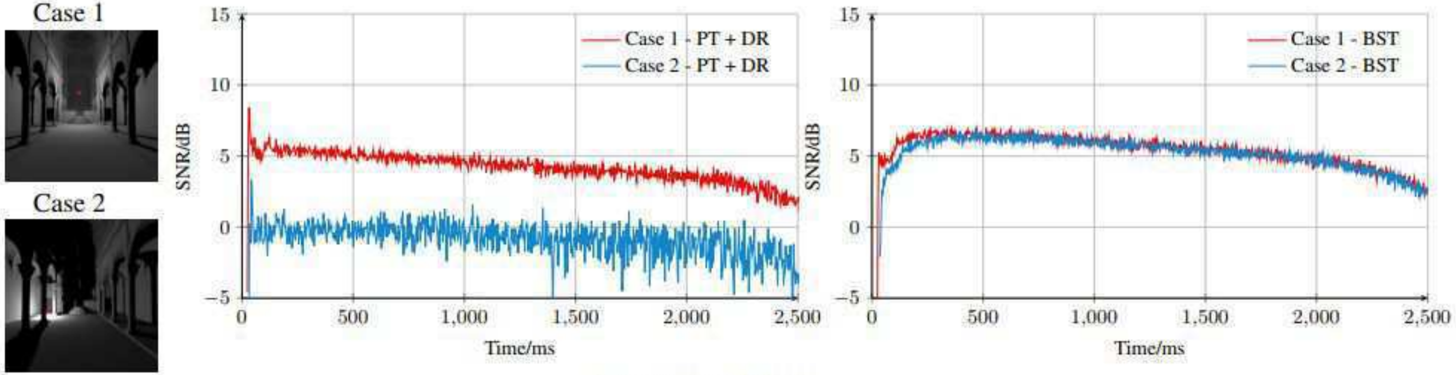
Figure 6: *Sound environments used in our experiments: (a) Roomset, 19266 triangles; (b) Elmia Round Robin, 1047 triangles; (c) Crytek Sponza, 279163 triangles; (d) Sibenik Cathedral, 75155 triangles; (e) Tradeshow, 177405 triangles.*

Intel i7 3.50Ghz CPU and 32GB Memory

Single Source Benchmarks



(a) Cube (12m × 12m × 12m)

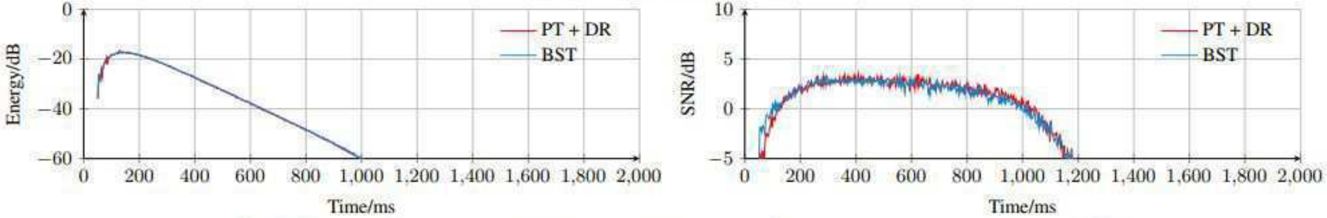


(b) Sibenik Cathedral

Single Source Benchmarks (no cache used)

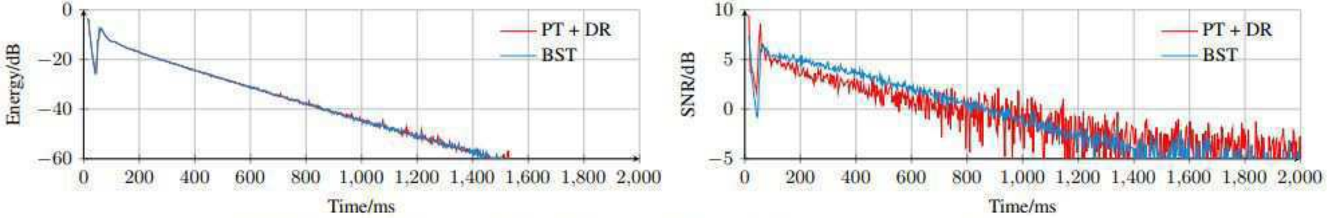
Scene Complexity			Computation Budget		Time Cost (ms)					
#Tris	#Bounces	Method	#Subpaths	#Connections	Trace	Connect	Optimize	Total	ASNR (dB)	Range (ms)

Roomset, 22m×11m×6m



19266	120	PT+DR	620	74400	30.8	4.46		35.3	1.92	54-990
		BST	100	48000	4.76	7.71	1.31	13.9	1.93	

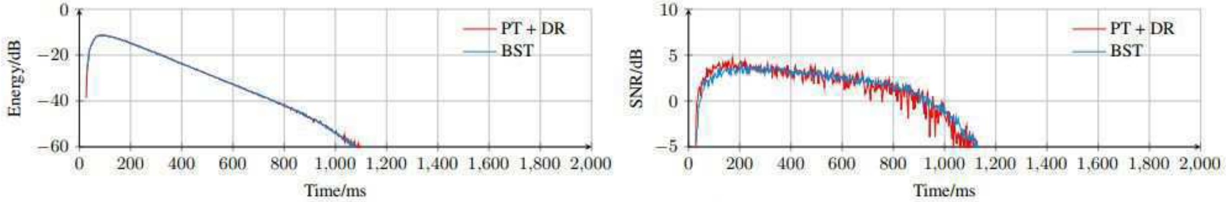
Elmia Round Robin, 42m×39m×16m



1047	100	PT+DR	1900	190000	62.1	41.1		103	-0.69	18-1917
		BST	100	40000	3.50	12.9	1.81	18.3	-0.65	

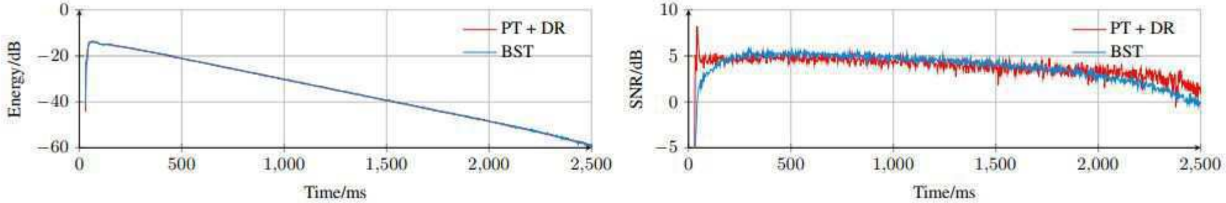
Crytek Sponza, 23m×37m×15m

Single Source Benchmarks (no cache used)



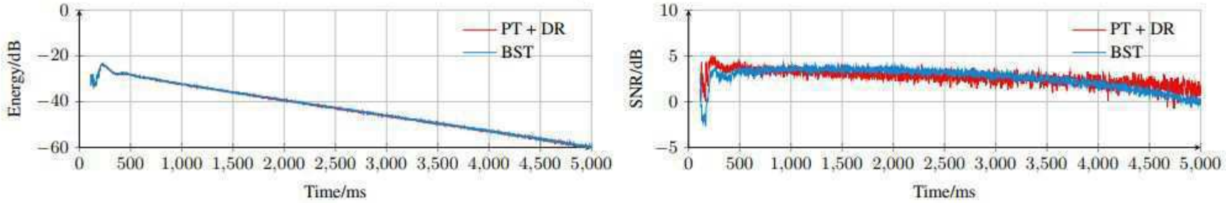
279163	80	PT+DR	2250	180000	90.3	41.7		132	2.36	33-972
		BST	100	32000	4.24	8.84	1.17	14.3	2.37	

Sibenik Cathedral, 19m×46m×35m



75155	90	PT+DR	6500	585000	241	109		349	4.24	33-2193
		BST	100	36000	4.63	11.6	1.18	17.4	4.29	

Tradeshow, 105m×84m×29m

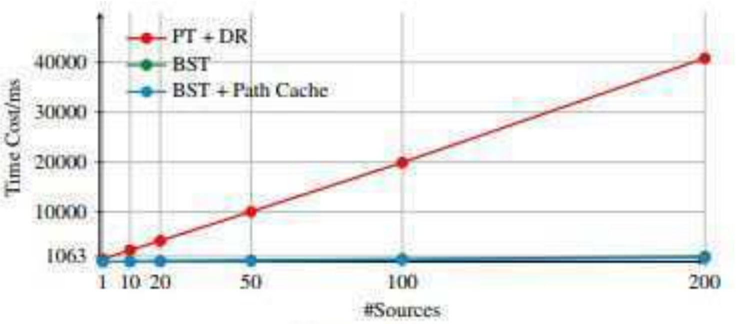


177405	70	PT+DR	28000	1960000	476	240		716	2.88	117-4392
		BST	100	28000	2.25	8.97	1.11	12.4	2.94	

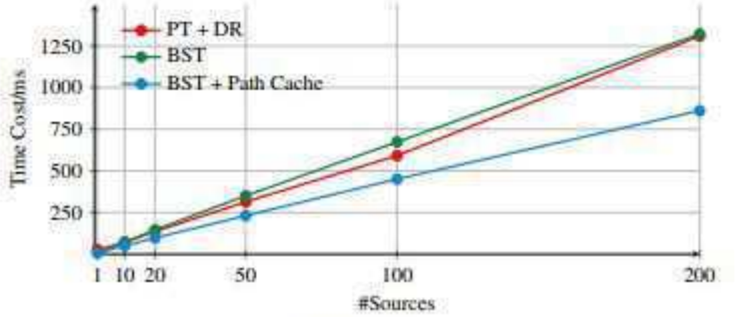
Single Source Benchmarks

- BST achieves higher rendering quality (in terms of the SNR metric) than PT+DR
- In scenes containing moving sources, BST produces more stable results than PT+DR.
- The result shows that our new algorithm outperforms PT+DR in all 5 scenes with a 2.5-58× speedup.

Multiple Sources



(a) Tradeshow



(b) Roomset

- PT+DR
  - the computation cost of the trace step scales sub-linearly with the source number
- BST
  - needs to compute the forward subpaths for every sound source
  - almost linearly with the source number
  - Path Caching can improve the scalability

Limitations

Limitations

- Using GA Methods
  - Inaccuracy at low frequency
  - Inability to simulate all wave-based sound effects
- Using BDPT
  - Difficulty in processing materials with ideal specular reflection
- Many features of the resulting energy response, like the smoothness, cannot be controlled.

End