Scattering Delay Network: an Interactive Reverberator for Computer Games

Enzo De Sena, Hüseyin Hacıhabiboğlu, Zoran Cvetković

Centre for Telecommunications Research, King's College London, UK
Informatics Institute, METU, Turkey

4 February 2011

Table of contents

Motivation

Reverberators for computer games should:

- provide the best accuracy given...
- ...acceptable computational load (at least real-time)

Computational load

- Priority is always given to video performance
- Not only powerful game workstations, but also slower architectures, e.g. handheld game consoles, ipods, mobile phones, etc..

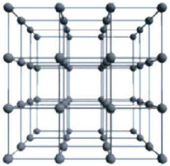
- Provide explicit control of room properties (size, wall absorption...) to streamline game development
- Scalable to different loudspeaker arrangements (2-channels, headphones, 5.1...)

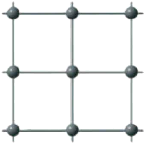
Available reverberators

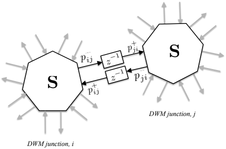
Many possibilities. Trade-off between complexity and accuracy:

- **DSP effects** easy, fast, but inaccurate
- **Statistical** very fast, but not directly related to room properties
- **FDN** fast, but tuning is indirect and trial-and-error
- **Synth. Reverb.** fast, sounds good, but careful tuning of parameters
- **Conv. Reverb.** sounds real, but require actual recordings
- **Ray Tracing** accurate, but heavy load (especially for later parts of RIR, which is also the less important perceptually)
- **DWM** wave equation solution, but heavy load (however, real-time implementation is possible with GPU)

Concepts of Digital Waveguide Mesh



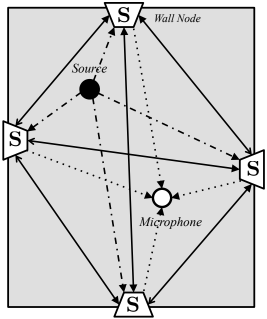




- The simulated space is sampled appropriately, each sampled point is a scattering junction
- Each junction is connected to its geometric neighbors through bidirectional lines
- Execution has two successive stages repeated at each time sample: scattering pass and propagation pass
- The outgoing waves to neighboring junctions $\mathbf{p}_i^-(n) = [p_{i,1}^-(n) \ldots, p_{i,K}^-(n)]^T$ are obtained as: $\mathbf{p}_i^-(n) = \mathbf{S}\mathbf{p}_i^+(n)$ where $\mathbf{p}_i^+(n) = [p_{i,1}^+(n), \ldots, p_{i,K}^+(n)]^T$ are the incoming wave variables
- $\mathbf{S} = \frac{2}{N-1}\mathbf{1}_{(N-1)\times(N-1)} - \mathbf{I}$
- If appropriate spatial sampling, equivalent to FTDT

Table of contents

Design overview



2D room example

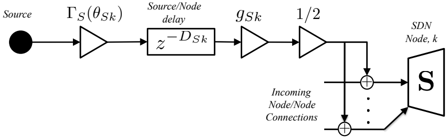
Scattering Delay Network

More importance to the parts of the RIR that are most significant perceptually: render correctly LOS and first-order reflections while approximating higher-order reflections with a single system.

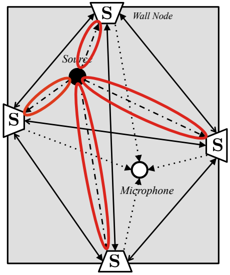
- Position the scattering nodes where the first-order reflections happen (in a ray-tracing sense)
- Mono-directional absorbing delay lines from the source to the scattering node
- Mono-directional absorbing delay lines from the scattering nodes to the microphone
- Bi-directional delay lines between the scattering nodes

Source to SDN connections

For each SDN node, structure as:



- Mono-directional delay line from the source to every SDN node
- Delay $D_{Sk} = \lfloor F_s \|\mathbf{x}_S - \mathbf{x}_k\| / c \rfloor$
- Attenuation $g_{Sk} = \frac{G}{\|\mathbf{x}_S - \mathbf{x}_k\|}$, $G \triangleq c/F_s$
- Weighted by the source directivity pattern



Pressure insertion

The output of the delay line is added to all the incoming lines p^+ of the SDN node

SDN Nodes

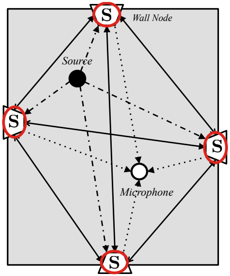
- Substantially a DWM node
- Wave arriving from a node propagates back to all the other nodes
- The outgoing wave variables $\mathbf{p}_i^-(n) = [p_{i,1}^-(n) \ldots, p_{i,K}^-(n)]^T$ are obtained as:

$$\mathbf{p}_i^-(n) = \mathbf{S}\mathbf{p}_i^+(n)$$

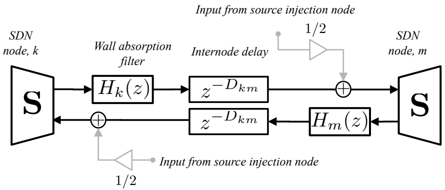
where $\mathbf{p}_i^+(n) = [p_{i,1}^+(n), \ldots, p_{i,K}^+(n)]^T$ are the incoming wave variables

- The scattering matrix $\mathbf{S}$ employed is the DWM lossless matrix:

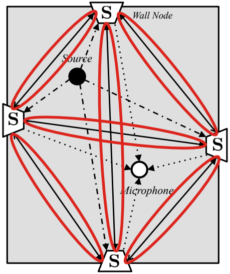
$$\mathbf{S} = \frac{2}{N-1}\mathbf{1}_{(N-1)\times(N-1)} - \mathbf{I}$$



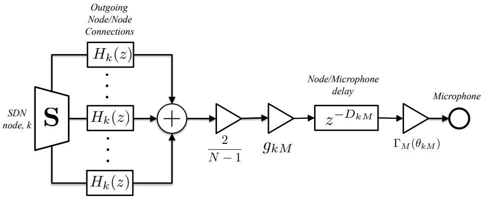
SDN interconnections

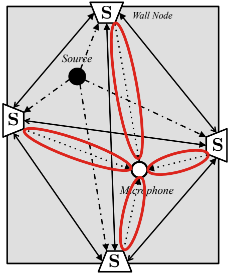


- Bi-directional delay line
- Delay $D_{km} = \lfloor F_s \|\mathbf{x}_k - \mathbf{x}_m\| / c \rfloor$
- No attenuation between SDN nodes
- Filters modeling wall absorption. Will be shown later that these filters can be set explicitly.
- The actual reverberator is formed by these lines + the SDN nodes



SDN node to microphone connections





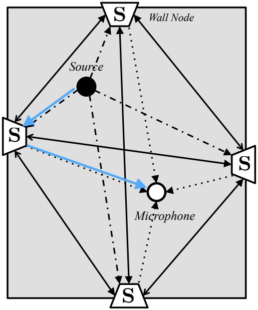
- Attenuation $g_{kM} = 1/\left(1 + \frac{\|\mathbf{x}_k - \mathbf{x}_M\|}{\|\mathbf{x}_S - \mathbf{x}_k\|}\right)$ such that $g_{Sk} g_{kM} = G/[\|\mathbf{x}_S - \mathbf{x}_k\| + \|\mathbf{x}_k - \mathbf{x}_M\|]$
- Weighted by the microphone directivity pattern

Pressure extraction

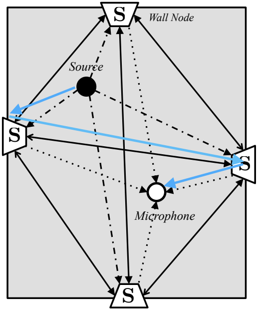
The input of the delay line is taken as the pressure (calculated after the wall filters) at the node as: $p_k(n) = \frac{2}{N-1} \sum_{i=1}^{N-1} \tilde{p}_{ki}^{+}(n)$

Scattering Delay Network

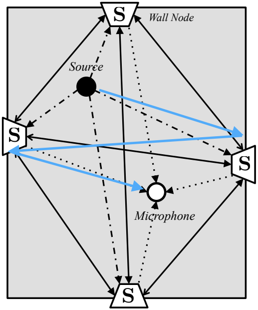
More importance to the parts of the RIR that are most significant perceptually: render correctly LOS and first-order reflections while approximating higher-order reflections with a single system.



I-order reflection



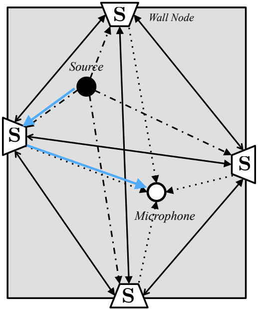
II-order reflection



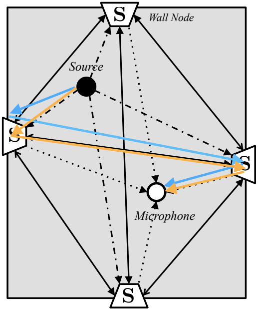
Another II-order reflection

Scattering Delay Network

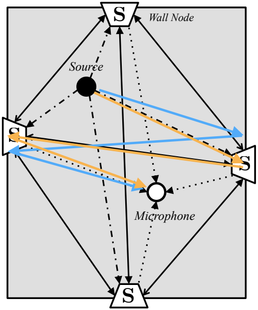
More importance to the parts of the RIR that are most significant perceptually: render correctly LOS and first-order reflections while approximating higher-order reflections with a single system.



I-order reflection

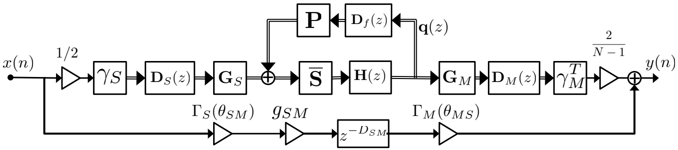


II-order reflection



Another II-order reflection

Transfer function



$$\gamma_S = [\Gamma_S(\theta_{S1}) \cdots \Gamma_S(\theta_{S1}), \Gamma_S(\theta_{S2}) \;\cdots\; \Gamma_S(\theta_{SN})]^T$$

$$\gamma_M = [\Gamma_M(\theta_{1M}) \cdots \Gamma_M(\theta_{1M}), \Gamma_M(\theta_{2M}) \;\cdots\; \Gamma_M(\theta_{NM})]^T$$

$$\mathbf{D}_S(z) = \mathrm{diag}(z^{-D_{S1}} \ldots z^{-D_{S1}}, z^{-D_{S2}} \;\ldots\; z^{-D_{SN}})$$

$$\mathbf{D}_M(z) = \mathrm{diag}(z^{-D_{M1}} \ldots z^{-D_{M1}}, z^{-D_{M2}} \;\ldots\; z^{-D_{MN}})$$

$$\mathbf{G}_S = \mathrm{diag}(g_{S1} \cdots g_{S1}, g_{S2} \cdots \; g_{SN})$$

$$\mathbf{G}_M = \mathrm{diag}(g_{M1} \cdots g_{M1}, g_{M2} \cdots \; g_{MN})$$

$$\overline{\mathbf{S}} = \mathrm{diag}(\mathbf{S}, \; \mathbf{S}, \; \cdots \; , \mathbf{S})$$

$$\mathbf{S} = \frac{2}{N-1}\mathbf{1}_{N-1\times N-1} - \mathbf{I}$$

$$\mathbf{D}_f(z) = \mathrm{diag}(z^{-D_{12}}, \; \ldots \; , \; z^{-D_{NN-1}})$$

$$\mathbf{H}(z) = \mathrm{diag}(H_1(z) \cdots H_1(z), H_2(z) \cdots \; H_N(z))$$

$\mathbf{P}$ is a permutation matrix

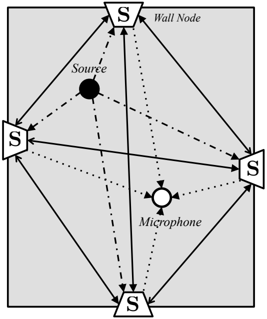
System's transfer function

$$H(z) = \overline{g} z^{-D_{SM}} + \frac{1}{N-1}\mathbf{k}_M(z)\left(\mathbf{H}(z^{-1}) - \overline{\mathbf{S}}\mathbf{P}\mathbf{D}_f(z)\right)^{-1}\overline{\mathbf{S}}\mathbf{k}_S(z),$$

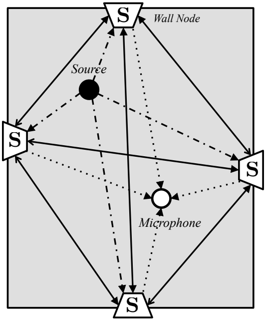
where $\mathbf{k}_M = \gamma_M^T \mathbf{D}_M \mathbf{G}_M$ and $\mathbf{k}_S = \mathbf{G}_S \mathbf{D}_S \gamma_S$.

Interactivity

- Source or microphone movement ⇒ position of SDN nodes is recalculated and delay line lengths are updated (removing/adding samples)
- Source of microphone turns ⇒ update in the directivity function Γ_S and Γ_M only.
- Changes in wall absorption characteristics or room dimensions can also be implemented.

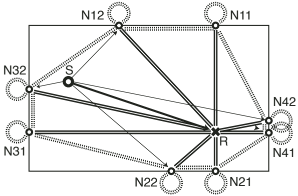


Recording Techniques



- Any microphone arrangement can be simulated. Intensity panning techniques (e.g. VBAP, sine pann. law..) or virtual coincident arrays (e.g. Ambisonics, HOA..) do not require duplication of the reverberator: multiple weighting Γ_M.
- Same for virtual binaural recording: Γ_M for each ear is an HRTF (both amplitude and delays).
- Non-coincident arrays, equivalent to multiple observation point $\Rightarrow$ multiple SDNs. However, for near-coincident arrays, SDN nodes will be very close $\Rightarrow$ the same SDN "backbone" can be shared, and only the microphone delay lines duplicated.

Relation to previous work



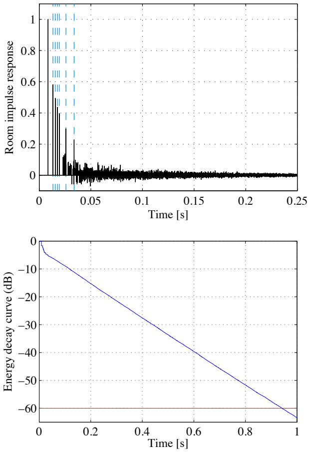
- Originally Digital Waveguide Networks were meant to be used as reverberators, but mainly used as complete room simulators
- One important follow-up:

  M. Karjalainen *et al.* "Digital Waveguide Networks for Room Response Modeling and Synthesis," in *118th Conv. AES*, 2005.

DWN	**SDN**
Microphone is a scattering node	Microphone is passive element
Additional delay lines codirectional with room-axes	Not present (less accurate for low-freq room modes, but faster).
Freq-dependent wall-filter modelled via self-connection at node (tuned heuristically)	Freq-dependent wall-filter modelled via explicit wall filters
Attenuation on source-to-node and source-to-microphone attenuations only. Inaccurate I-order reflections and problems with source close to wall.	First-order reflections are correct by construction.

Outline

Room Impulse Response (RIR) Example



Simulation setup

- Shoebox room 9x7x4m
- Source in the center of the room
- Microphone in $\mathbf{x}_M = [2, 2, 1.5]$
- Frequency-independent absorption
  $\alpha = 0.2 \Rightarrow H_i(z) = \sqrt{1-\alpha} = 0.89$

- Blue lines (theoretical I-order reflections): correct by construction
- Dense reverberation tail
- Energy decay curve, $\left(\int_t^\infty |h(\tau)|^2 \, d\tau\right)$ is exponential (as for most real rooms). Thus scattering and attenuation by walls, microphone, and source lines provide realistic attenuation.
- Reverberation time $RT_{60} \approx 0.94$.

Sabine and Eyring formulas

As a rough metric to assess the performance of the proposed method, comparison of the reverberation time of SDN with the two well-known empiric formulas:

Sabine formula

- First formula from empirical studies of 1910

$$T_{60} = \frac{0.161V}{\sum_i A_i \alpha_i}$$

where V is the room volume, A_i is the area of the i-th face, α_i is the absorption per square meter.

Eyring formula

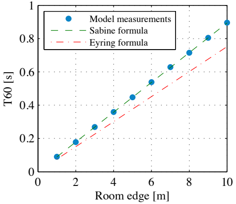
- Eyring modified Sabine formula because it yield $\alpha > 1$ for some acoustically dead rooms

$$T_{60} = -\frac{0.161V}{\left(\sum_i A_i\right)\log_{10}\left(1 - \frac{\sum_i A_i \alpha_i}{\sum_i A_i}\right)}$$

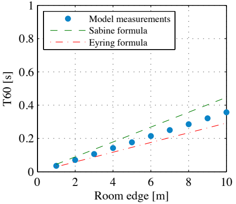
Results

Simulation setup

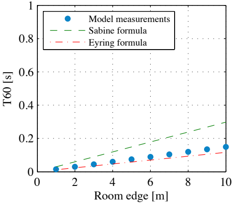
- Cubic room with edge x and uniform absorption $\alpha = 0.3, 0.6, 0.9$
- $\Rightarrow T_{60,Sab} = x\frac{0.161}{6\alpha}$ and $T_{60,Eyr} = x\frac{-0.161}{6\log_{10}(1-\alpha)}$
- Source and microphone in the middle of the room (LOS removed)



$\alpha = 0.3$



$\alpha = 0.6$



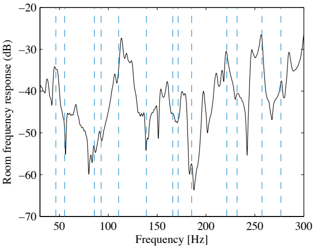
$\alpha = 0.9$

- Model's RT_{60} also increases linearly with x
- Always between the two formulas for all three absorption coefficient
- Model's RT_{60} get closer to Eyiring formula prediction as absorption increases

Low frequency room modes

Room low-frequency modes

One of the main acoustical properties of rooms. The frequencies that result in resonance phenomena: $f_{n_x,n_y,n_z} = \frac{c}{2}\sqrt{\left(\frac{n_x}{l_x}\right)^2 + \left(\frac{n_y}{l_y}\right)^2 + \left(\frac{n_z}{l_z}\right)^2}$. Depending on the position of the source/observer, picks or dips in the frequency response.



Simulation setup

- Shoebox room 2x2.14x3.74m
- Source and microphone 1m away from one another around the centre of the room
- Frequency-independent $\alpha = 0.12$

- Axial modes (i.e. $f_{n_x,0,0}$, $f_{0,n_y,0}$, and $f_{0,0,n_z}$) are drawn in blue
- They correspond to peaks and dips in the Room Frequency Response

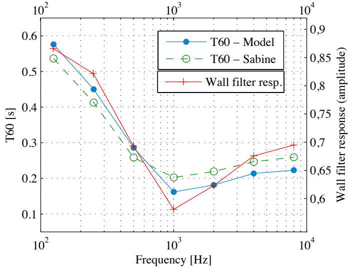
Frequency-dependent reverberation

In FDN, filters are introduced to control frequency-dependent reverberation time. In SDN's the introduced filters model directly wall characteristics:

Simulation setup

- Cubic room 5x5x5m
- Source and microphone $\approx$ 3m away from one another around the centre
- All walls have frequency-dependent absorption (cotton carpet) implemented as a second-order IIR minimum-phase filter.

- Wall reponse as in figure (right axis)
- Reverberation time estimate through Sabine formula:
  $$T_{60,Sab} = \frac{0.161V}{\sum_i A_i \alpha_i(\omega)} = \frac{0.161 \cdot 5}{6(1-|H(e^{j\omega})|^2)}$$
- Very close mach between RT_{60} produced by the model and Sabine formula's estimates $\Rightarrow$ We are able to explicitly control wall properties



Conclusions

- A perceptual room simulator (reverberator) has been presented
- Room modelled by scattering nodes interconnected by bidirectional delay lines
- These SDN nodes are positioned where first-order reflections originate
- $\Rightarrow$ First-order reflections are simulated correctly, while a rich but less accurate reverberation tail is obtained, consistent with the room characteristics.
- Transfer function was derived
- Examples and comparison with reverberation time formulas showed its performance
- Explicit control of wall acoustical properties