Perception-Based Methods for Spatial Audio

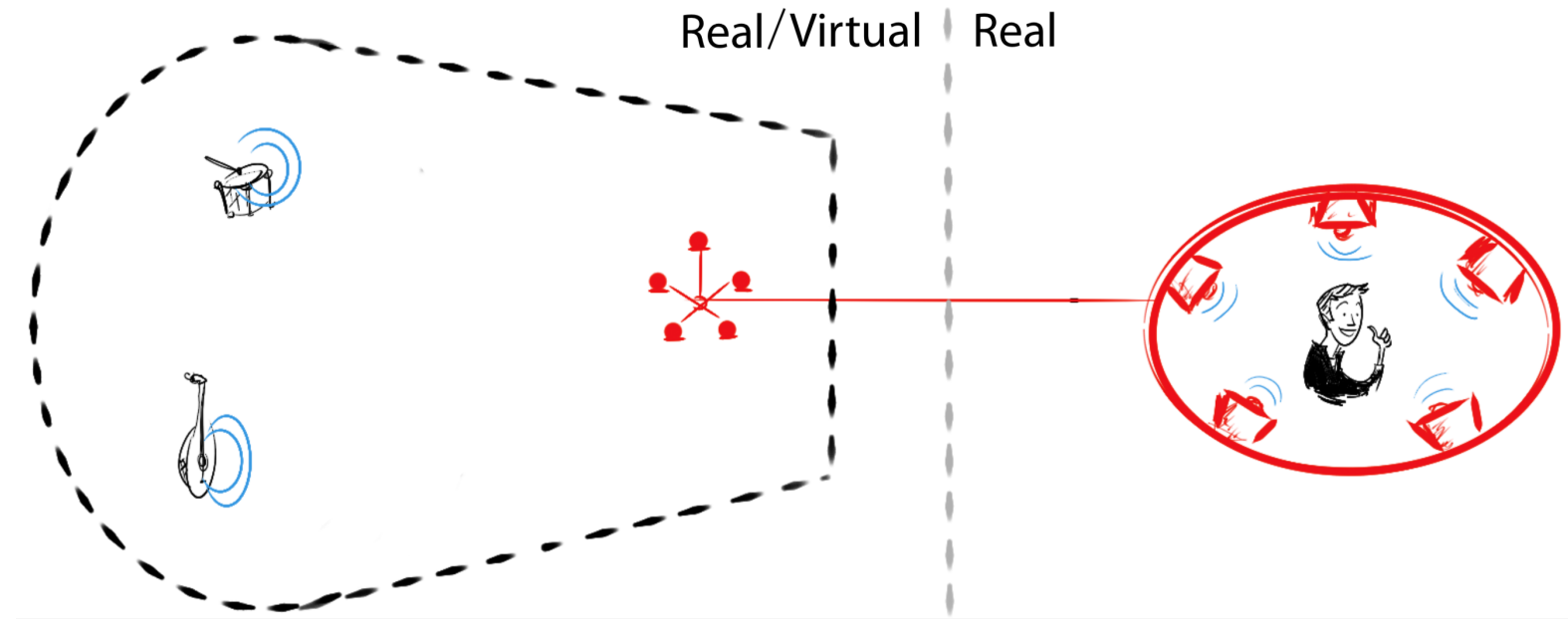
Enzo De Sena

Institute of Sound Recording, University of Surrey, UK

UKAN Seminar, 28 October 2020

Objective

- Making listener feel transported to a different auditory scene, which can be
  - a real recorded one (live music performance, sporting event..)
  - a virtual one (video games, VR/AR, architectural acoustics..)



Physical and cross-talk cancellation methods

	SFR	**Multichannel**	**2-Channel**
Channel count	50+	< 10	2
Equipment Load	High	Commercially viable	Low
Psychoacoustics	None	Required	Critical
Sweet Spot	Large	Medium, small group	Small, individual

- Sound Field Reconstruction (SFR) provide mathematically elegant solution (e.g. HOA, WFS)...
  - but large number of loudspeakers: $r = \frac{c}{f}\frac{N}{2e\pi}$, e.g. $f = 10$ kHz, $r = 0.1$ m $\Rightarrow N = 56$
- 2-channel (cross-talk cancellation) binaural methods, only two channels...
  - but small sweet spot (e.g. [Rose et al., 2002] report ≈ 3 cm)
- We'll focus on multichannel systems with limited equipment load, which need to leverage somehow psychoacoustics

Outline

Acknowledgements

Joint work with:

Prof Zoran Cvetkovic̀ (King's College London)
Prof Huseyin Hacihabiboglu (METU)
Prof Julius O. Smith (Stanford University)
Prof Toon van Waterschoot (KU Leuven)
Prof Marc Moonen (KU Leuven)
Dr Niccoló Antonello (KU Leuven)
Stojan Djordjevic (University of Surrey)
Ashley Andrew-Jones (University of Surrey)
Ege Erdem (METU)

Funding from:

EPSRC, EU Commission, FWO, KU Leuven research funds

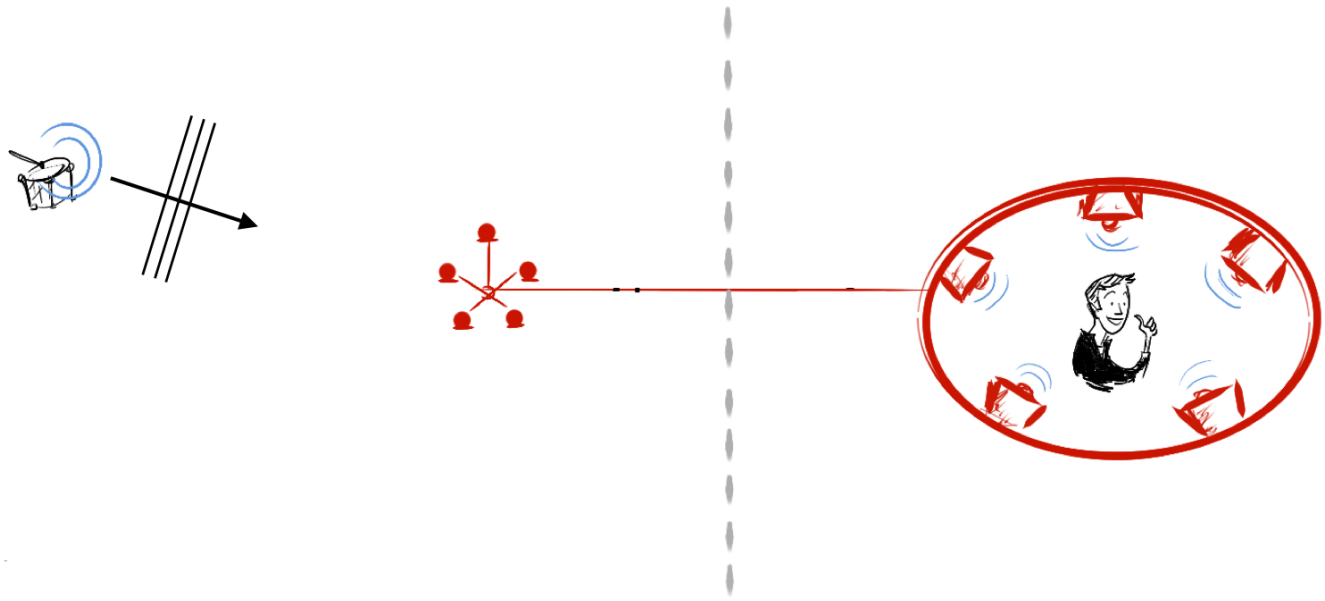
About this talk

- Interrupt me
- Just want to give main ideas
- Details and maths left to references (at the end)
- Will share slides (or find them later today at desena.org)

Outline

Reproduction of a single plane-wave

- Let's start from a simplified case
- Rendering a single plane-wave source
- Plane wave could represent e.g. single source or reflection
- (Multiple plane waves will use superposition)



Reproduction of plane waves

- Assume for now that plane wave direction, θ_s, is known
- Relevant case for computer games, music post-production, spatial audio objects (MPEG-H)

Objective–Reproduced plane wave should be:

1. perceived in correct direction (low localization error)
2. easy to localize (low localization uncertainty)

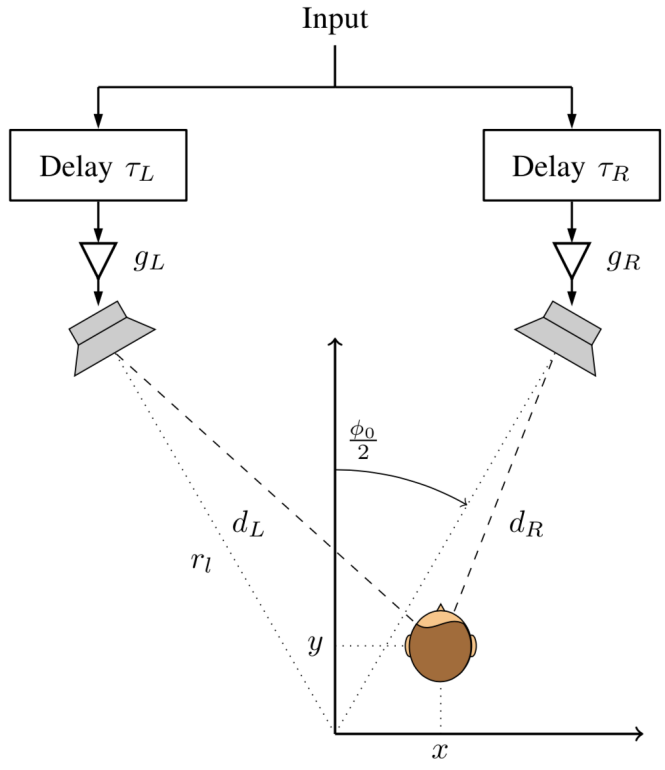
- in the largest possible area (large sweet spot)

How many loudspeakers to use to reproduce plane wave?

- **Question**: how many loudspeakers should we use for a single plane wave?
- Objective analysis in [De Sena et al., 2013] based on active intensity vector field (direction of energy propagation)
- Spatial fluctuations increase with angle between loudspeaker pairs
- **Answer**: use only the two loudspeakers closest to direction of plane wave
- This reduces problem to good ol' stereophonic reproduction

Frequency-independent inter-channel differences

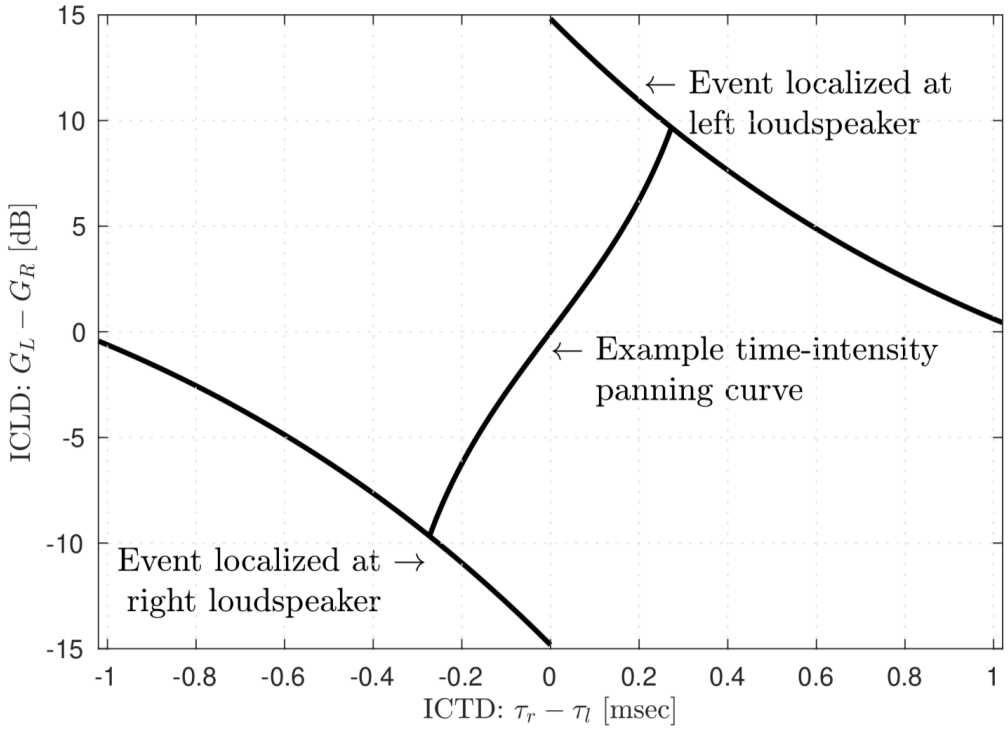
- What should we do with those two loudspeakers?
- Consider frequency independent inter-channel time differences (ICTD) and level differences (ICLD)
- ICTD/ICLDs lead to low coloration [Spors et al., 2013], which is most important attribute for sound quality [Rumsey et al., 2005]



- As long as ICTD below echo threshold, listeners will perceive a fused "phantom source" (summing localization effect)

Position of phantom source

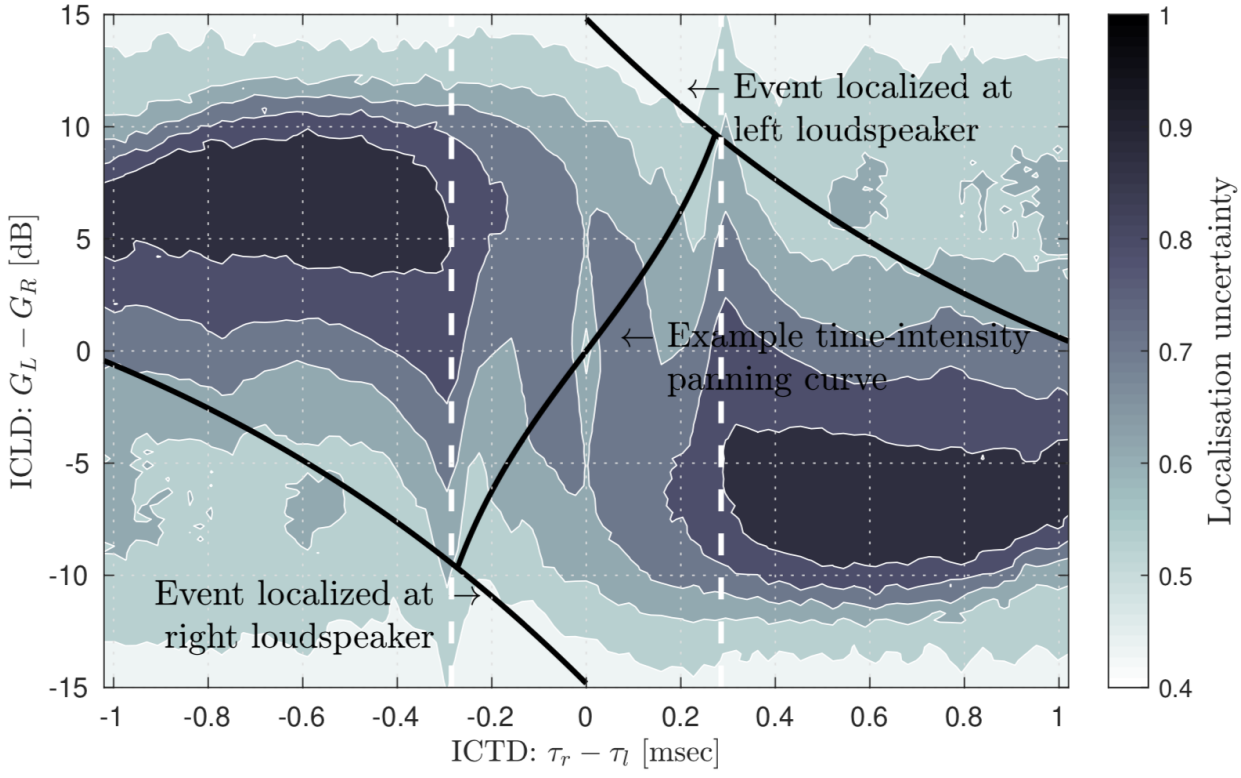
- Position of phantom source depends on ICTD/ICLD pair
- Same position can be achieved with different ICTD/ICLD pair
- One can use e.g. intensity only (most commercial sound recordings), time only, or time-intensity



Adapted from [Williams, 2004]

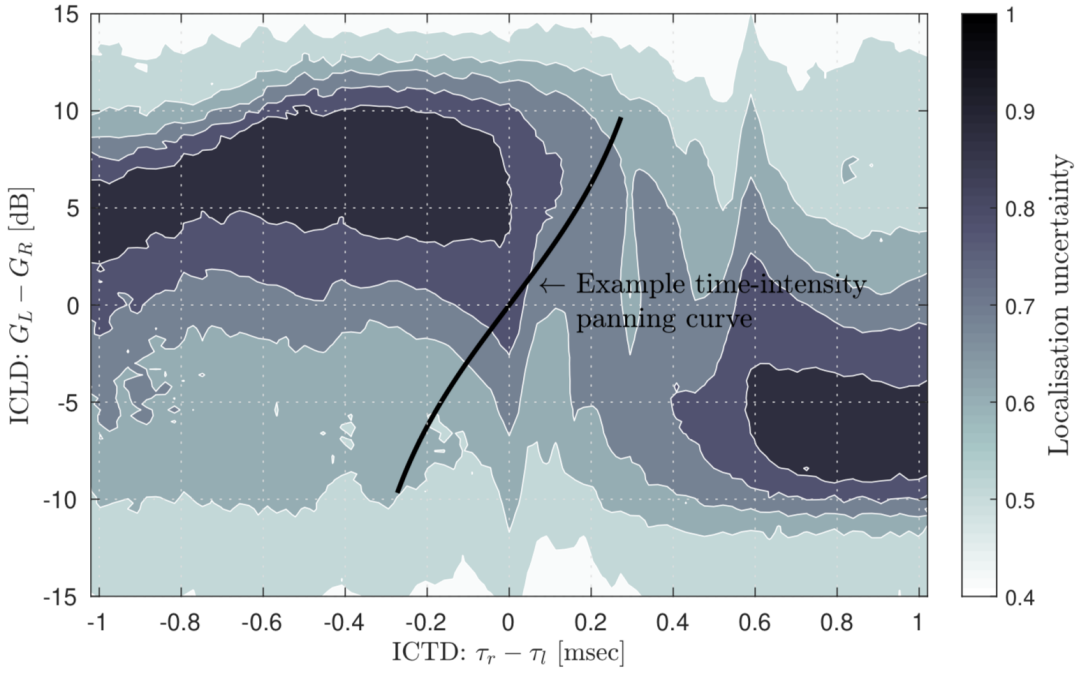
Not all ICTD/ICLD pairs are created equal

- ICTD/ICLD pairs lead to different localization uncertainty
- Computational model in [De Sena et al., 2020]:



- Inconsistent ICTD/ICLD lead to high uncertainty
- The vertical bands correspond to cases where 2 replicates at one ear, but only 1 at the other

Localization uncertainty in off-center positions



- Listener moves 10 cm to the right, then entire plot moves (approximately) to the right
- Now intensity methods lie in area with high uncertainty!
- Time-intensity largely avoids this area

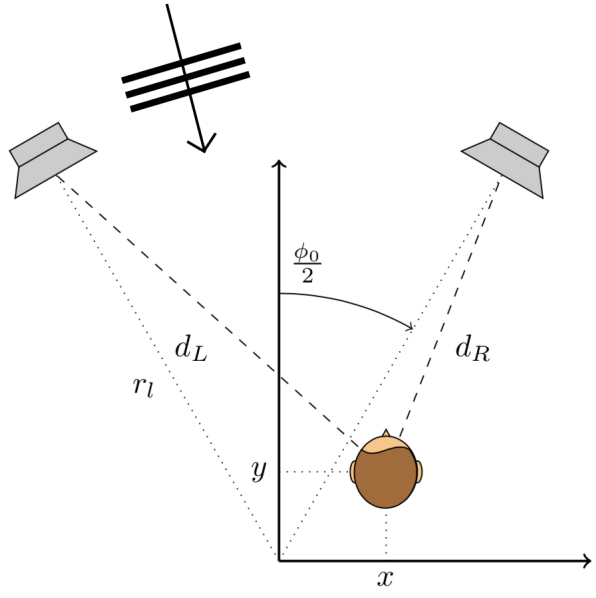
What is happening?

- Useful to define "relative" ICTD/ICLD as observed by the listener:

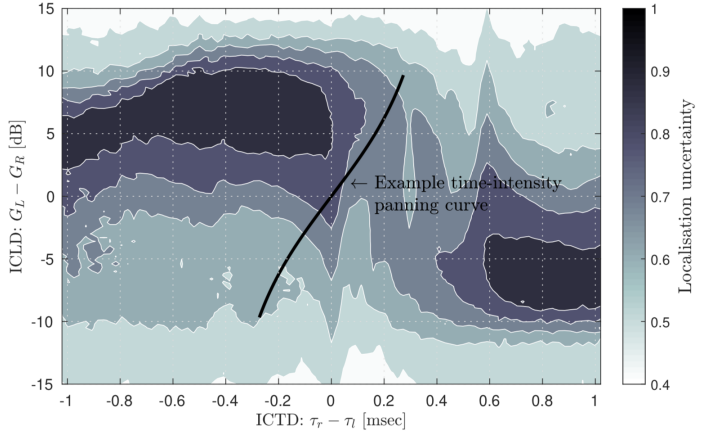
$$\text{RICLD} \approx \text{ICLD} - \frac{x}{r_l} \frac{20 \sin\left(\frac{\phi_0}{2}\right)}{\log_e(10)} ,$$

$$\text{RICTD} \approx \text{ICTD} - x \frac{2}{c} \sin\left(\frac{\phi_0}{2}\right) .$$

where c speed of sound



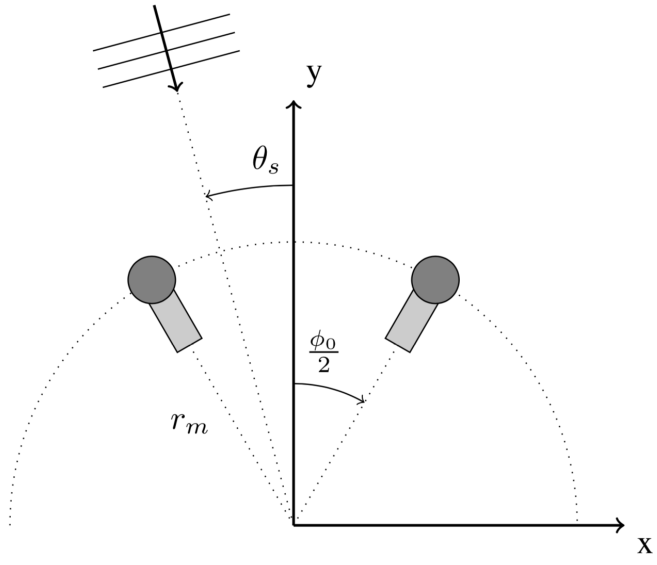
- E.g. consider ICTD $= 0$ ms and ICLD $= 5$ dB (left leading)
- RICTD $= -0.29$ ms and RICLD $= 4.78$ dB, i.t. contradicting
- Adding a small ICTD will delay the onset of contradictory RICTD/RICLD



Parametrization of ICTD (time-delay microphone array)

- Convenient now to specify ICTD and ICLD functions of θ_s, including a parameter taking into account how much we rely on ICLD compared to ICLD (time-intensity trade-off)

- Let the ICTD be defined according to the delay that would be observed on two spatially separated microphones as in figure:

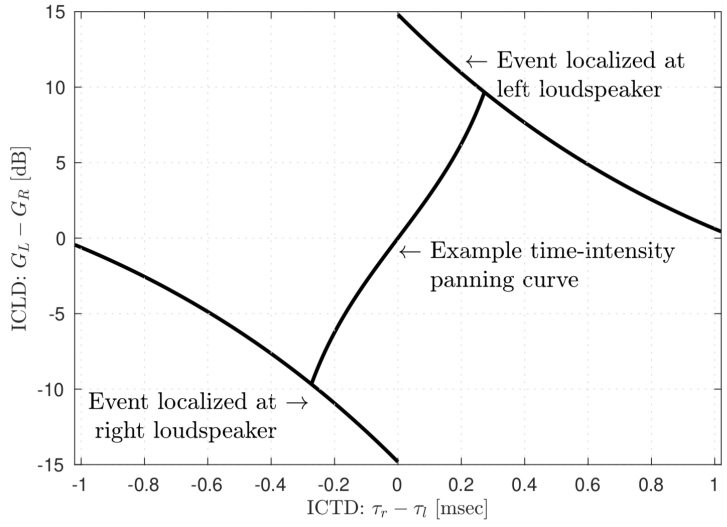


$$\mathrm{ICTD}(\theta_s, r_m) = 2\frac{r_m}{c}\sin\left(\frac{\phi_0}{2}\right)\sin\theta_s$$

where r_m is the array radius

- This parametrization is convenient since it allows to easily extend to the case of recording with circular arrays

Parametrization of ICLDs



- Psychoacoustic curves give only extreme positions
- Could use different curves, for instance [De Sena et al., 2013]:

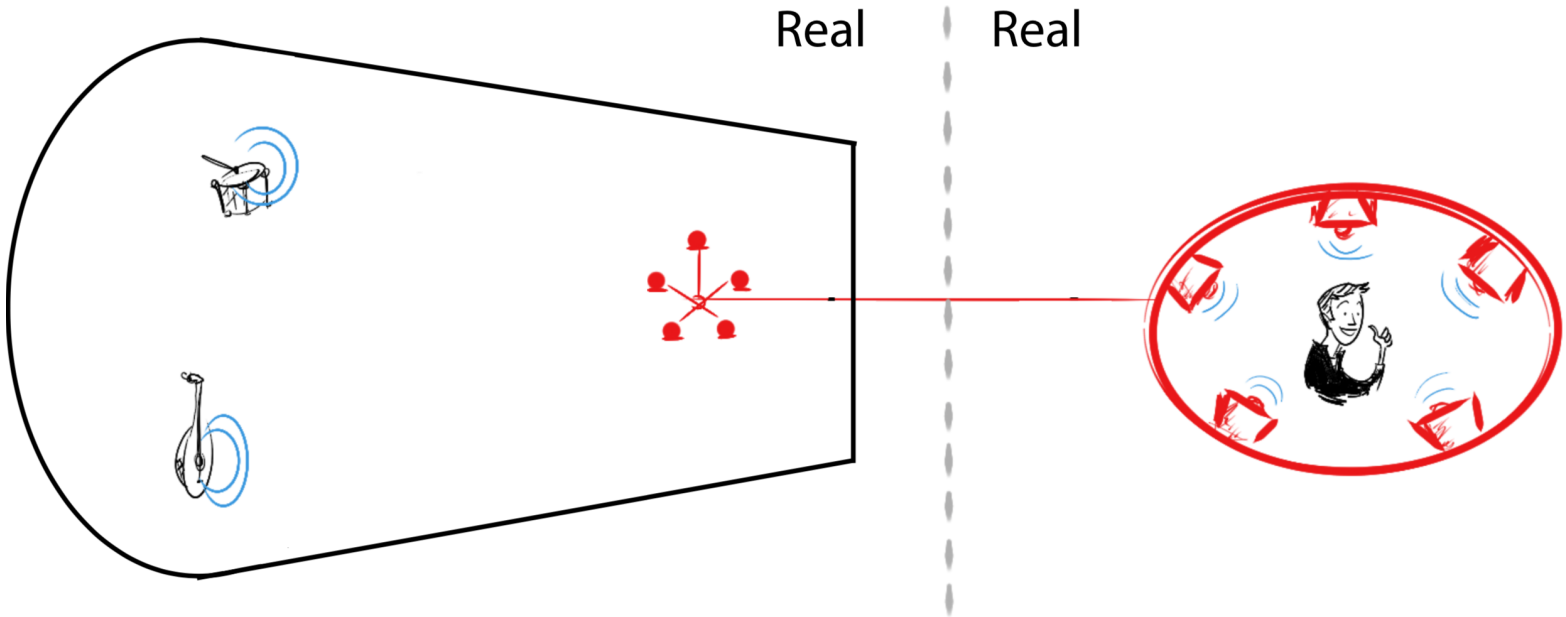
$$\mathrm{ICLD}(\theta_s, r_m) = 20 \log_{10} \frac{\sin\left(\frac{\phi_0}{2} + \beta(r_m) + \theta_s\right)}{\sin\left(\frac{\phi_0}{2} + \beta(r_m) - \theta_s\right)}$$

  where $\beta(r_m)$ is a parameter used to fit the extrema
- With this parametrization, a higher r_m leads to more reliance on ICTDs and lower ICLDs

Outline

Perceptual recording of a real sound scene

- Let's move now to the case of recording a real acoustic scene with multiple sources, reflections etc
- Now direction of sources/reflections is unknown



To DOA or not to DOA

- Possible approach involves two steps:
  - direction of arrival (DOA) estimation (its own field of research)
  - take the signal and artificially add ICTD/ICLD
  - This is the approach followed by many methods, as e.g. Dirac/SDM/SIRR
- DOA estimation is not trivial and carries errors, especially for short time windows
- Let's try to use a different approach...

How to bypass DOA

- Connect each microphone to a corresponding loudspeaker
- We already chose ICTDs parametrisation to be the delay between two microphones
- The pair of microphones will have correct ICTDs by construction (without need to know source direction)

- Microphone directivity pattern $\Gamma(\theta)$: sensitivity of microphone as function of direction θ
- Use the directivity pattern to obtain desired ICLD(θ_s, r_m)
- This process makes DOA estimation unnecessary!

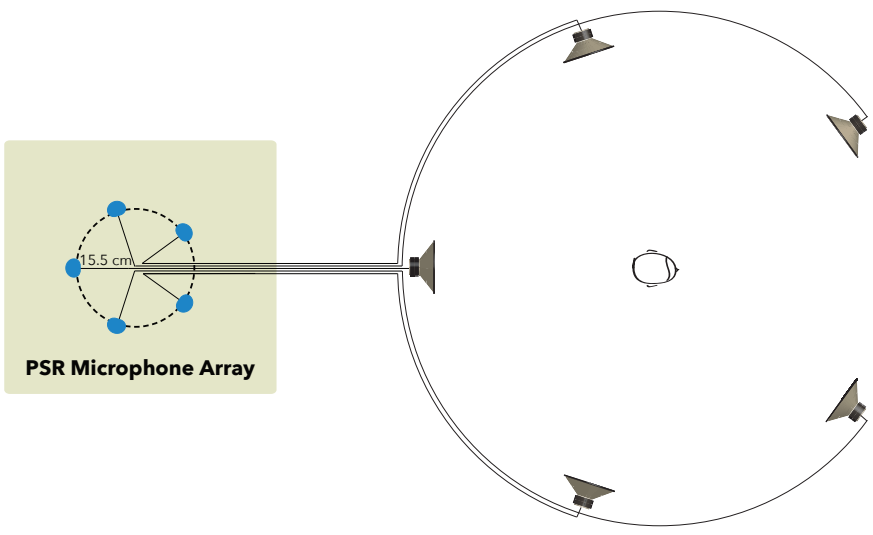
- Rest of this section will describe this design process in more in detail [De Sena et al., 2013]

How many loudspeakers in total?

- 5 channels: minimum for 360° perspective [Fletcher 1953]
- 5 channels: minimum for envelopment [Ando et al. 1986]

Perceptual Soundfield Reconstruction (PSR) Array [Johnston et al., 2000]

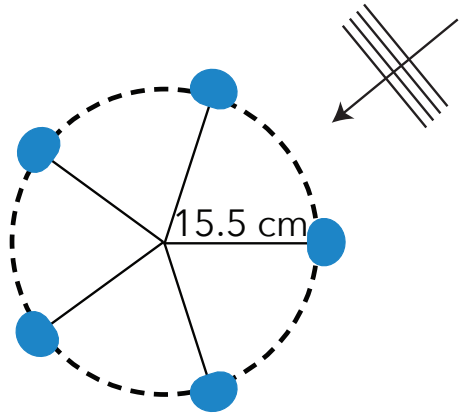
- 5 channels, uniformly distributed



First design step – # constrain only 2 loudspeakers active

How to have only 2 loudspeakers active for each wave?

- Design microphone directivity pattern $\Gamma(\theta)$ such that it does not pick up directions beyond neighbouring microphone



$$\Gamma(\theta) = \begin{cases} 0 & \theta \notin [-\phi_0, \phi_0] \\ ? & \theta \in [-\phi_0, \phi_0] \end{cases}$$

where ϕ_0 is angle between microphones

- From here on we consider microphone pairs
- Design $\Gamma(\theta)$ for $\theta \in [-\phi_0, \phi_0]$ to capture required (ICTD,ICLD) pairs

Second design step – polar pattern

- We know from earlier closed form expression for ICLD(θ, r_m)
- The relation that connects ICLD and directivity patterns is

$$\text{ICLD}(\theta, r_m) = 20 \log_{10} \frac{\Gamma_{l+1}(\theta)}{\Gamma_l(\theta)}$$

  with $\Gamma_{l+1}(\theta)$ and $\Gamma_l(\theta)$ polar patterns of adjacent mics
- We have 1 equation but 2 unknowns
- Add constraint $\Gamma^2_{l+1}(\theta) + \Gamma^2_l(\theta) = 1$ for equal loudness

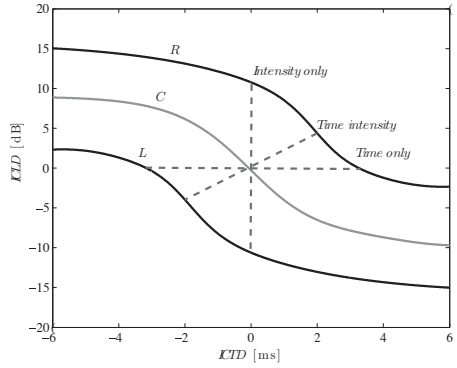
Second design step – polar pattern (cont'd)

Summary of first and second design steps

- Putting together $\mathrm{ICLD}(\theta, r_m) = 20 \log_{10} \frac{\Gamma_{l+1}(\theta)}{\Gamma_l(\theta)}$ (TID) and $\Gamma^2_{l+1}(\theta) + \Gamma^2_l(\theta) = 1$, and symmetry across microphones:

$$\Gamma(\theta) = \begin{cases} \left[1 + \dfrac{\sin^2(\theta + \beta(r_m))}{\sin^2((\phi_0 + \beta(r_m)) - \theta)}\right]^{-1/2} & \theta \in [0, \phi_0] \\ \left[1 + \dfrac{\sin^2(\beta(r_m) - \theta)}{\sin^2(\theta + (\phi_0 + \beta(r_m)))}\right]^{-1/2} & \theta \in [-\phi_0, 0] \\ 0 & \text{elsewhere} \end{cases}$$

- We can still choose r_m: higher r_m leads to larger ICTDs

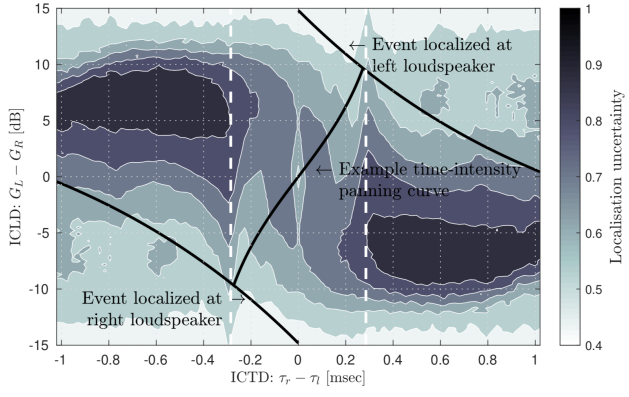


Third design step – array radius

- We saw earlier: larger ICTDs lead to larger sweet spot
- Where do we stop? Two vertical bands in figure
- Array radius such that panning curve end-point touches vertical band [De Sena et al., 2020]:

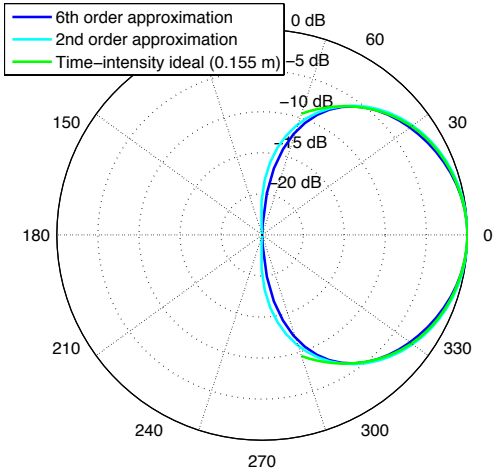
$$r_m \approx r_h \frac{\cos\left(\theta_e - \frac{\phi_0}{2}\right) + \frac{\phi_0}{2} + \theta_e - \frac{\pi}{2}}{2\sin^2\left(\frac{\phi_0}{2}\right)}$$

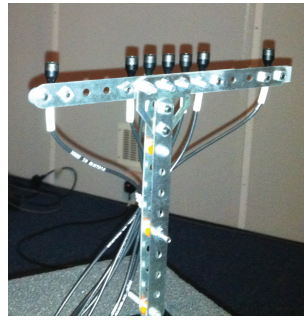
- For $r_h = 9$ cm and 5 loudspeakers, optimal radius $r_m \approx 15$ cm.



Microphone directivity that approximates ICLD(θ_s, r_m)

- Now we have the entire design: microphone positions, orientations and directivity pattern $\Gamma(\theta)$
- Are we done? No... there is no microphone that can implement exactly $\Gamma(\theta)$, but we can get close enough
- Second and higher-order microphones, e.g. differential mics [De Sena et al., 2012], Eigenmike [Elko and Meyer], or filter-and-sum beamformers can be used for this purpose





Subjective listening tests and extensions

Subjective listening tests [De Sena et al., 2013]

- Similar performance to second-order Ambisonics and VBAP in sweet-spot centre
- Lower uncertainty in off-centre positions

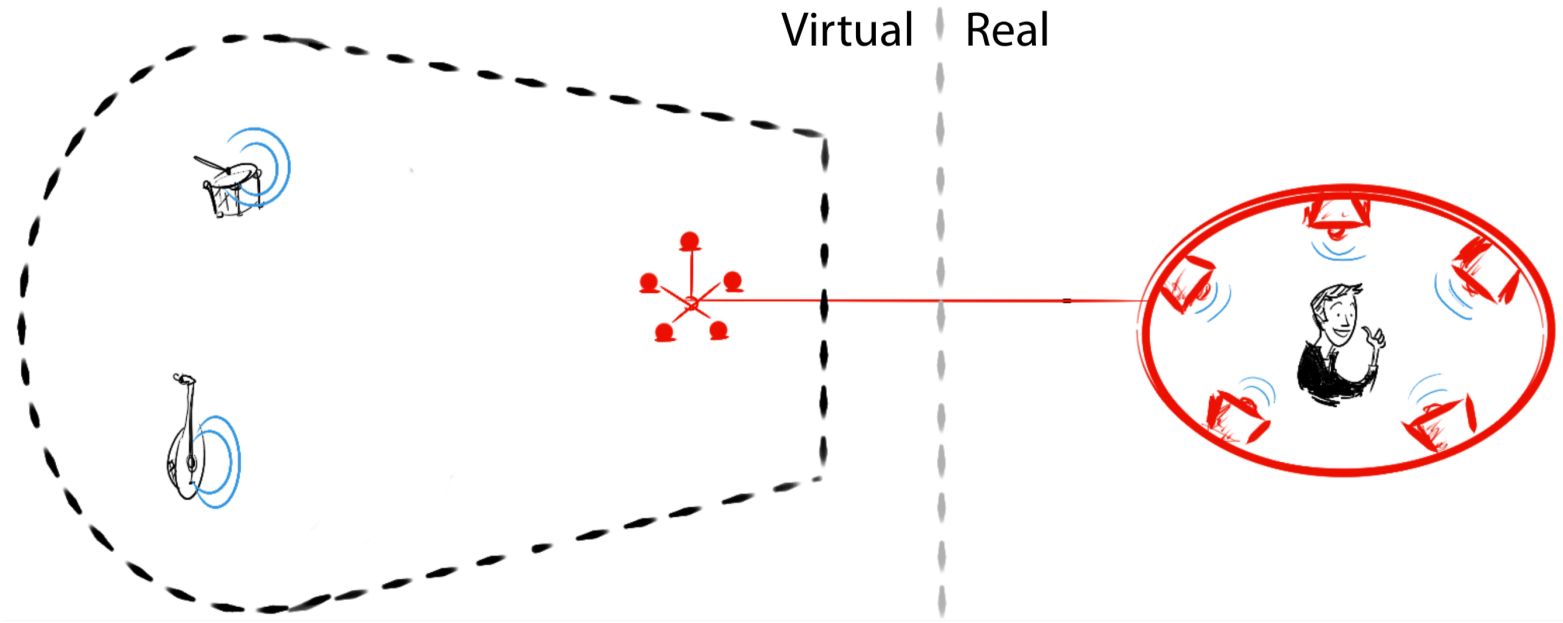
Current work and extensions

- PSR recently extended to third dimension using extrapolation from Eigenmike [Erdem et al., 2019]
- Time-intensity in the vertical dimension leads to a perceived improvement in stability of sweet spot [Andrew-Jones, 2019]

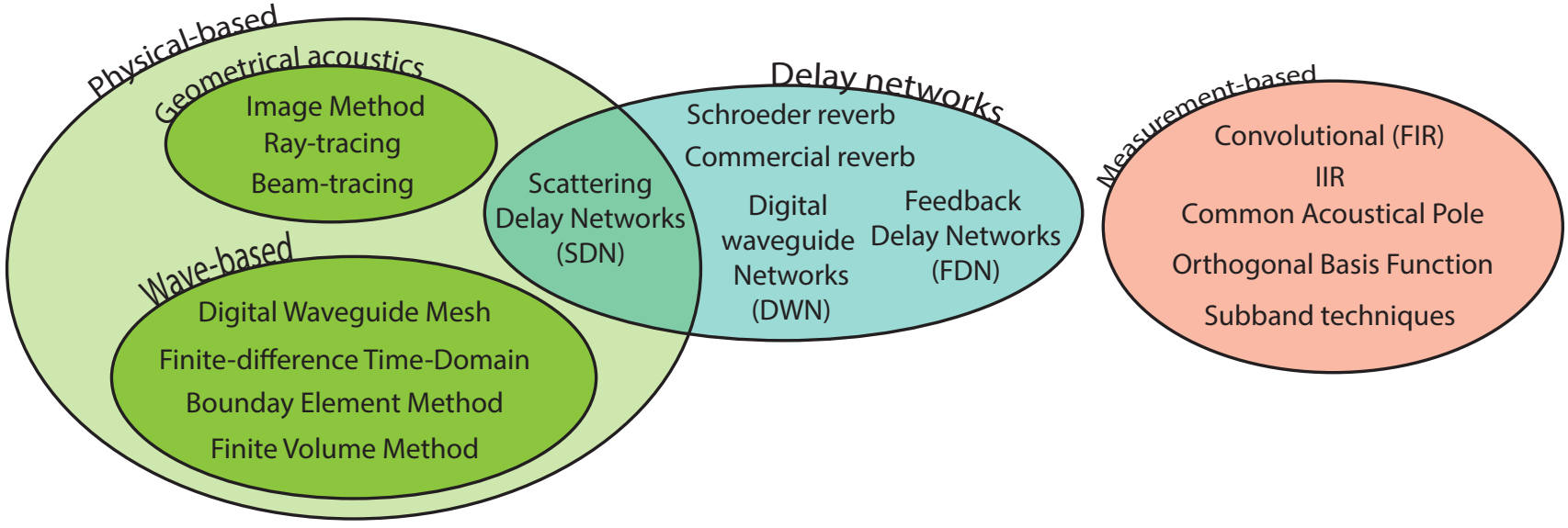
Outline

Perceptual Simulation of Room Acoustics

1. Simulate virtual room acoustics
2. *Virtual* recording and *real* reproduction (simulate microphone array as described in first part of talk)



Overview



- Overview of more than 50 years of room acoustic simulation in [Välimäki *et al.*, 2012], [Välimäki *et al.*, 2016] and [Hacıhabiboğlu *et al.*, 2017]
- Wave-based models are the most accurate ones

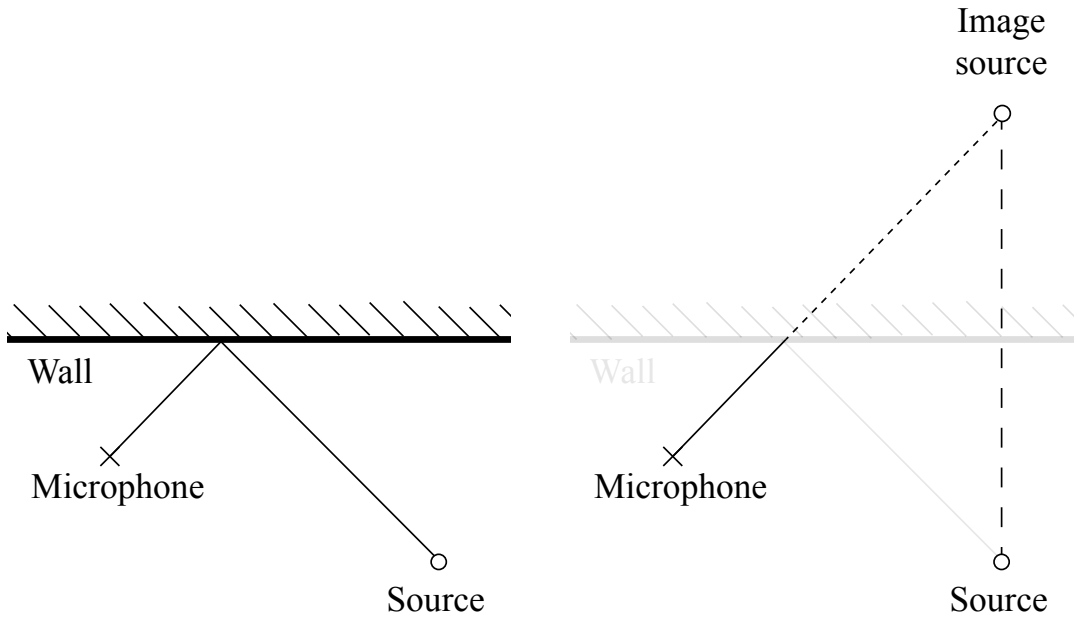
Rendering of dynamic scenes with wave models

- In a complete wave model of a room:
  - sources and listeners can be moved
  - *spatialized* using microphone arrays or "virtual dummy head"

Example: How expensive is a wave-based model?

- Audio bandwidth $= 20$ kHz ≈ 1.27 cm wavelength
- Spatial samples every 0.63 cm or less
- $3.65 \times 5.8 \times 2.4$ m room requires > 200 million grid points
- 3D finite difference model requires one multiply and 6 additions per grid point $\Rightarrow$ 70 billion FLOPS at $F_s = 50$ kHz
- $30 \times 15 \times 6$ m concert hall requires $>$ *3 quadrillion FLOPS*

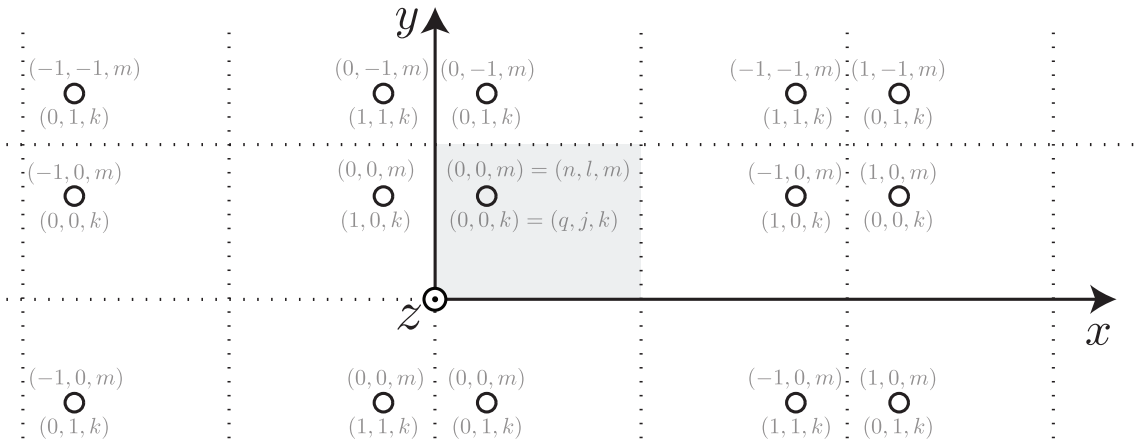
Image method for single reflector



- Wave propagation in half space is equivalent for:
  1. source and wall
  2. source and image source **(no wall)**
- Exact for rigid wall ($\nabla p \cdot \mathrm{n} = 0$)
- Approximation for non-rigid wall

Image method for rectangular room

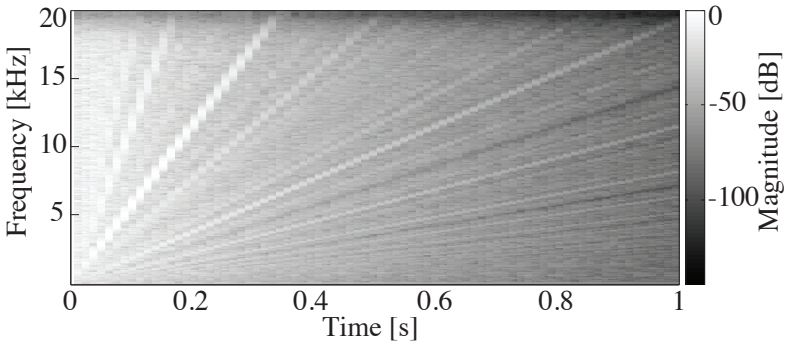
- With multiple reflectors: remove wall, mirror source *and* opposite wall



- Rectangular room: proven to be solution of wave equation for rigid walls [Allen and Berkley, 1979]
- Non-rectangular rooms also possible, but need expensive computations of image source visibility [Borish 1984]

Sweeping echoes in image method [De Sena *et al.*, 2015]

- Perfectly rectangular rooms cause so-called *sweeping echoes*
- Regular simulation setups yield stronger sweeping echoes



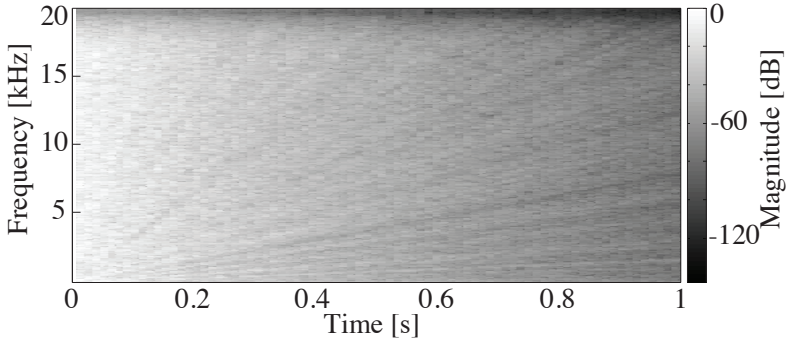
RIR: play

Anechoic: play REV: play

Room dimension: $4 \times 4 \times 4$

Mic. position $[1, 2, 2]$

Source position $[2, 1.5, 1]$



RIR: play

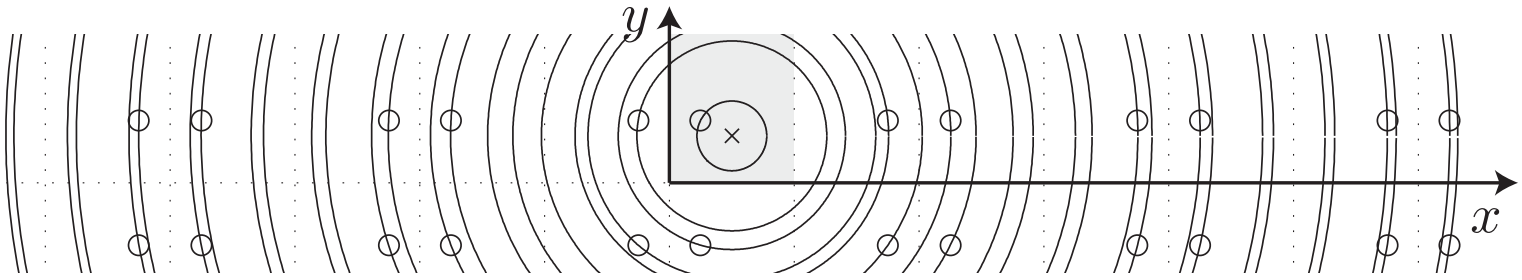
Anechoic: play REV: play

Room dimension: $4.1 \times 4.2 \times 4.3$

Mic. position $[1.4, 2.5, 2.6]$

Source position $[2.7, 1.8, 1.9]$

Cause of sweeping echoes



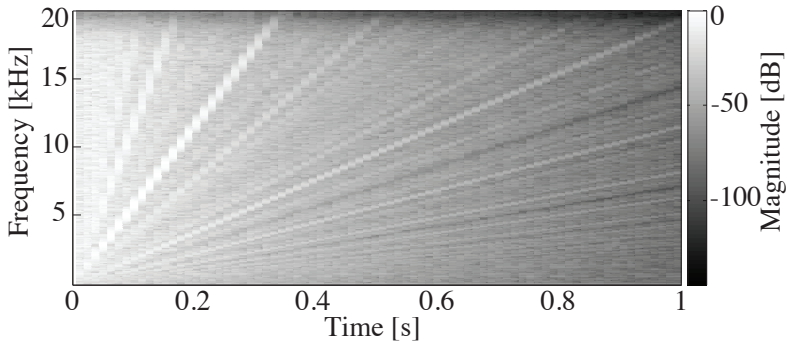
- Due to orderly alignment of images along 3 axes
- Pairs get closer like $1/t$, thus freq. resp. stretches with t
- Does not happen in most real rooms due to small imperfections

How to avoid sweeping echoes

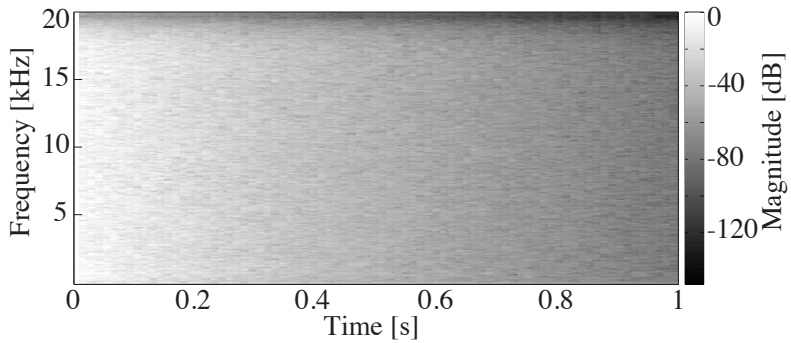
- Choose non-regular setup, e.g. random, or...

Randomized Image Method

- Small random displacement of images
- Uniform in ± 8 cm sufficient to remove completely
- Same complexity of rectangular image method



RIR: play REV: play

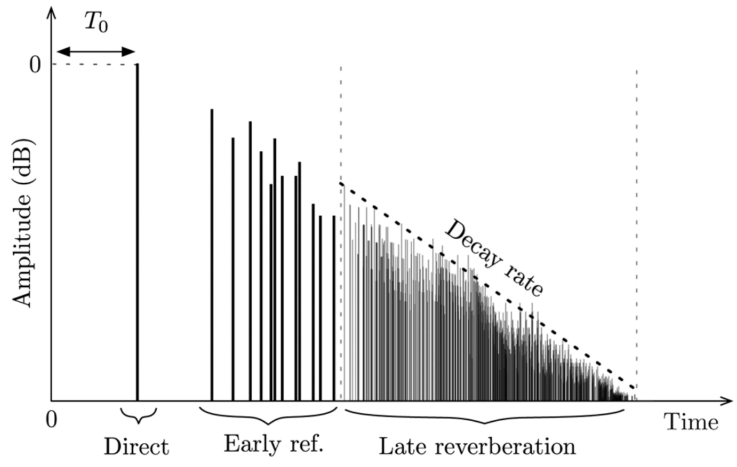


RIR: play REV: play

Rendering of dynamic scenes with geometric models

- Even then, we still need to
  1. recalculate RIR when moving source/observer
  2. run a convolution (computationally expensive for real-time applications)
- If physical accuracy not needed, perceptual methods provide better option

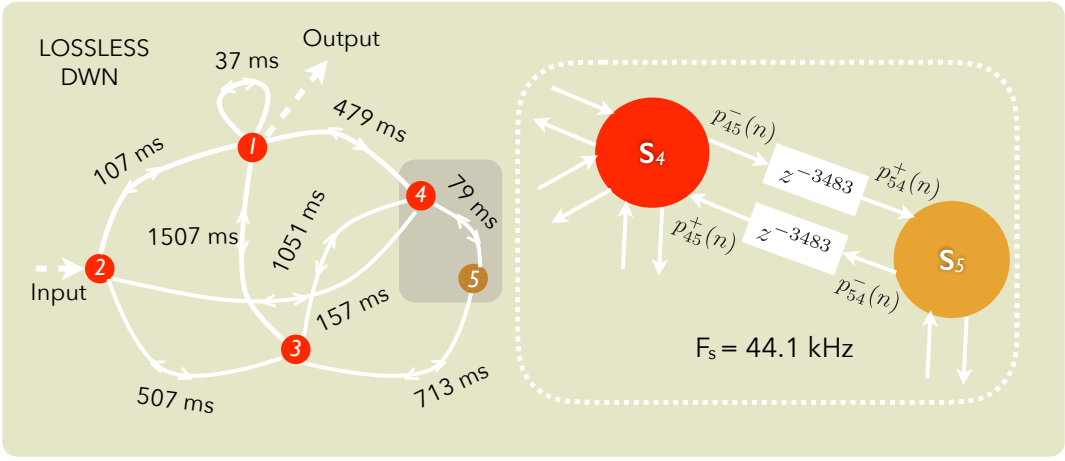
Room acoustics perception



RIR components:

- Direct line-of-sight
- Early reflections: important for perception of size and shape
- Late reverberation: important for envelopment and perception of size; we are not sensitive to exact structure
- Echo density has to be high enough for perceived texture
- Mode density high enough so that it is not metallic

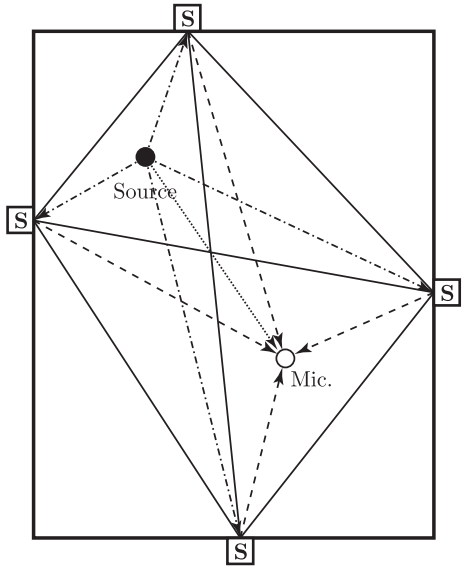
Digital waveguide networks (DWN)



- Network of bi-directional delay lines connected at scattering junctions [Smith, 1985]
- Can be interpreted as network of acoustic tubes
- **Question:** How to set parameters (delay line lengths, network connections, scattering matrix..)?

Scattering delay network (SDN) [De Sena et al., 2015]

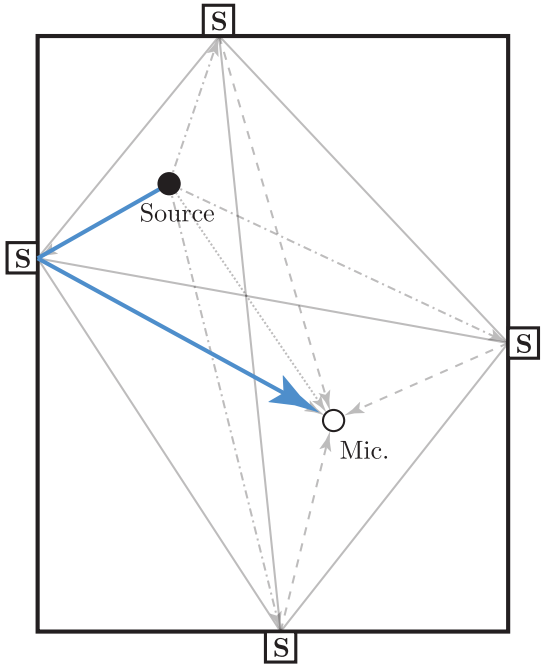
- Design DWN based on characteristics of a physical room



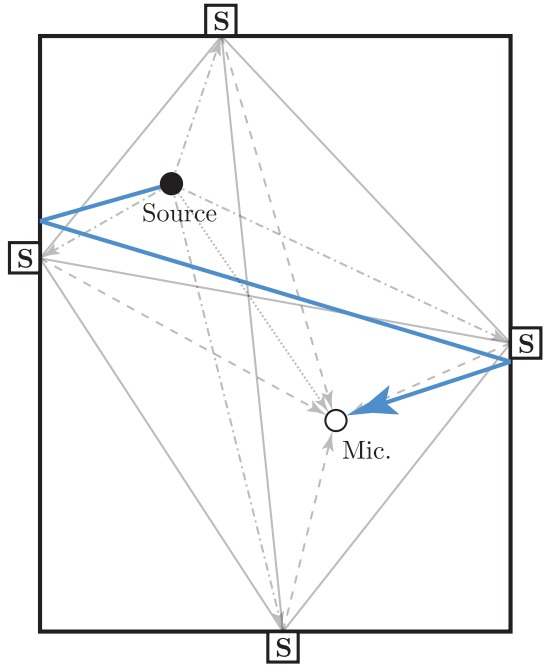
- Position nodes at first-order reflection points
- Fully connected DWN network
- Mono-directional lines for source-junction and junction-mic

SDN: approximation of geometric acoustics

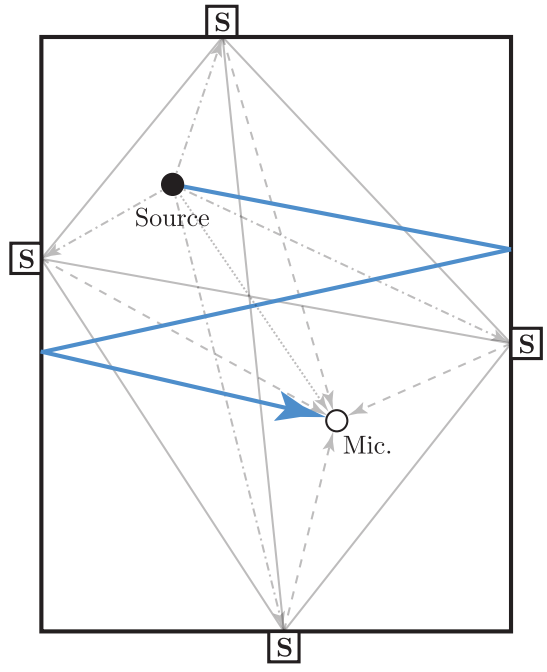
- Correct rendering of LOS and first-order reflections in time, amplitude and direction
- Approximation of second and higher-order reflections, less important perceptually



I-order reflection



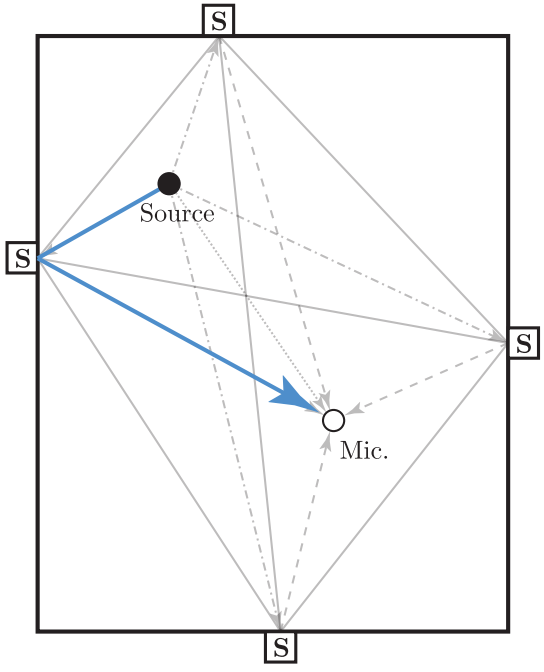
II-order reflection



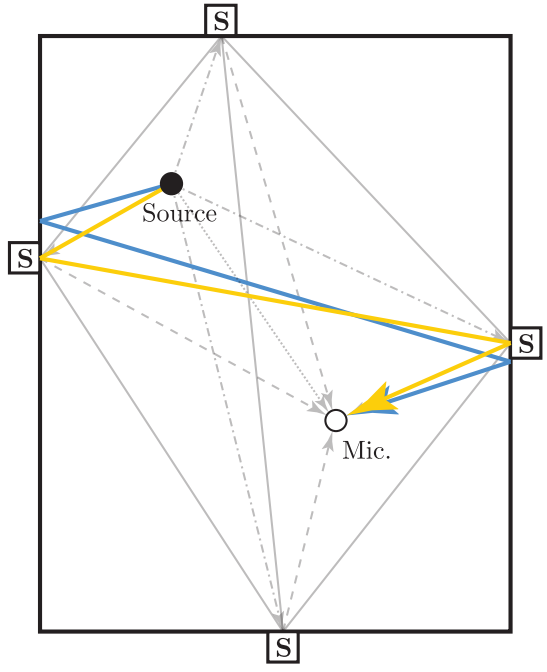
Another II-order reflection

SDN: approximation of geometric acoustics

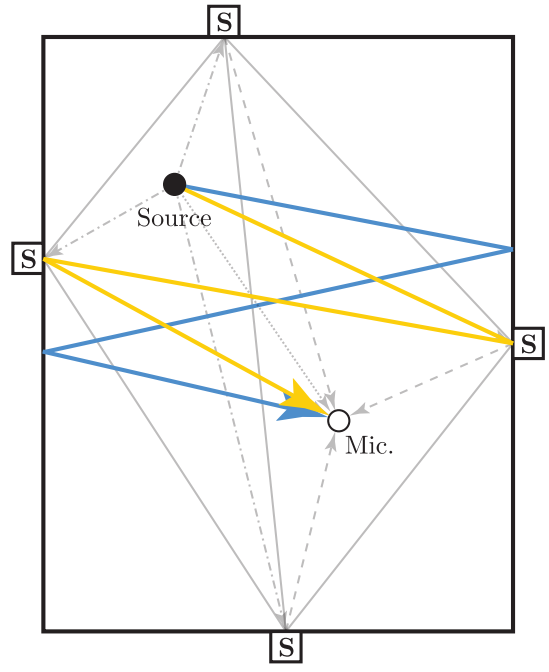
- Correct rendering of LOS and first-order reflections in time, amplitude and direction
- Approximation of second and higher-order reflections, less important perceptually



I-order reflection



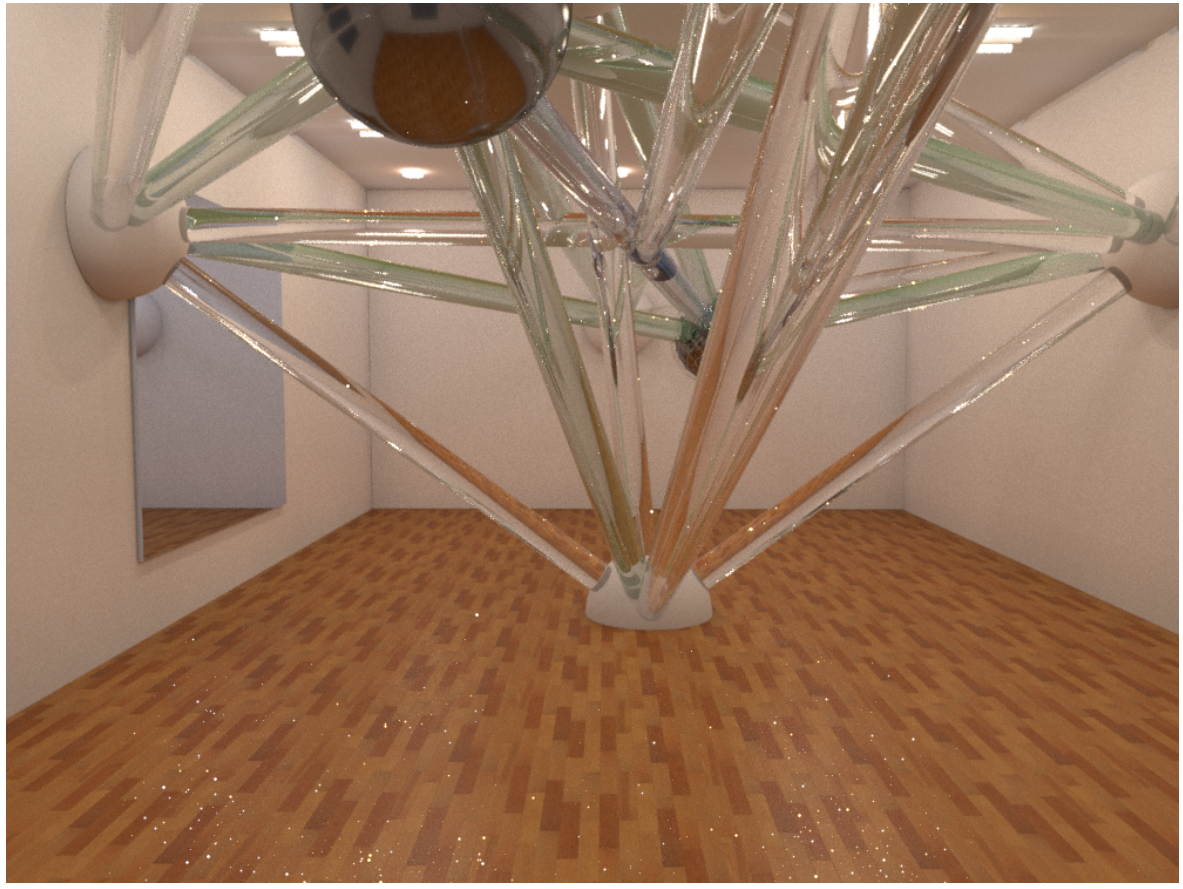
II-order reflection



Another II-order reflection

SDN: alternative interpretation

- Can also be interpreted as model of network of acoustic tubes

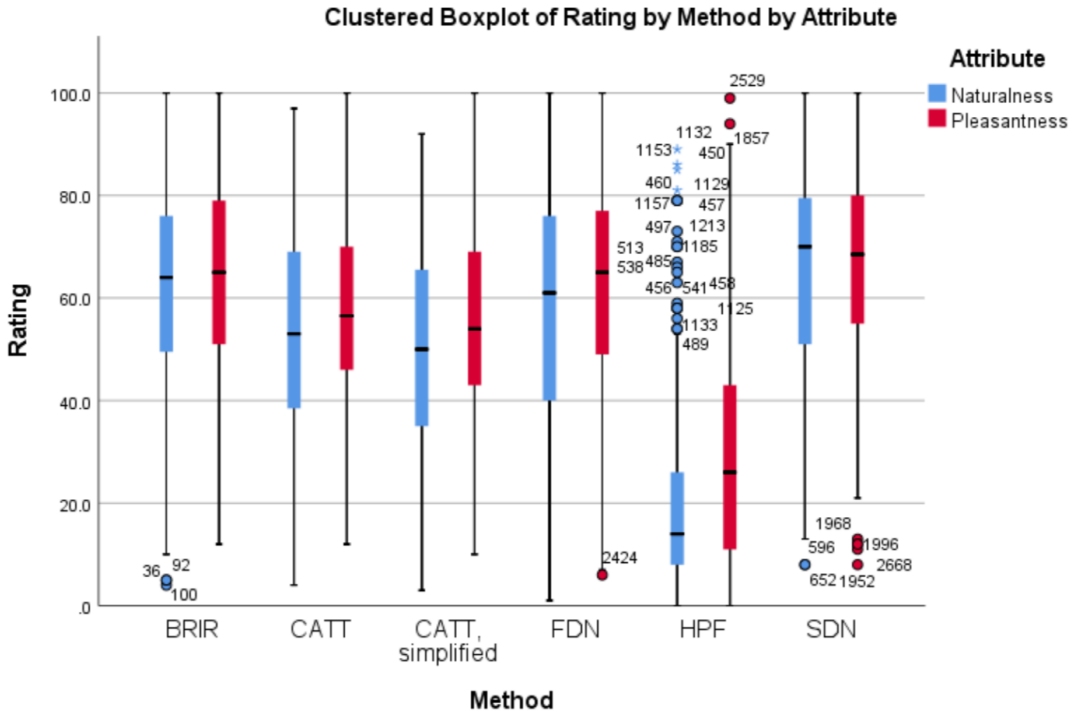


Advantages

- Also, not shown here:
  - similar frequency-dependent RT60 to full-scale models
  - similar echo density to full-scale models
  - sufficient modal density
  - axial resonant modes of room well approximated
- Orders of magnitude faster than FFT convolution (alone!)
- All parameters of model derived from physical properties

Perceptual evaluation [Djordjevic, 2019]

- Headphone-based (binaural) comparison (28 subjects)
- Higher pleasantness ($p < 0.001$) and naturalness ($p < 0.001$) than comparable delay-network based method

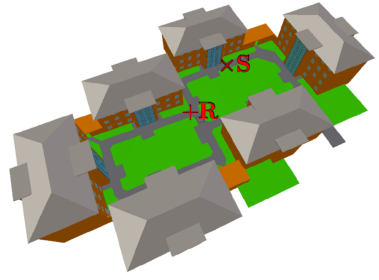


Comparison of SDN-IM-FDTD

See https://youtu.be/1hdhhrM4juQ

Recent advancements in SDN

- Stevens et al. (2017):
  - Extension to exact second-order reflections
  - Implementation of direction-dependent scattering (e.g. modelling of trees)
  - Modelling of outdoor scenes (sky absorbing nodes)
- Schlecht and Habets (2017):
  - Showed scattering matrix is "unilossless"



- SCReAM (SCalable Room Acoustics Modelling): £407k EPSRC project, hiring 3-year postdoc (deadline 16 Nov)

Outline

Contributions

Recording and reproduction

- Consistent ICLD/ICTD pairs $\Rightarrow$ sources easier to localise
- Small ICTDs $\Rightarrow$ larger sweet spot by delaying occurrence of inconsistent ICLD/ICTD pairs
- Beamforming used to bypass need for explicit DOA estimation

Room Acoustics Simulation

- When using image method, beware of sweeping echoes
- SDN is simple/fast network of delay lines that renders accurately what is more important perceptually
- Orders of magnitude faster than convolution (alone!)

- That's all folks! :) Questions?

Further Reading

Spatial Sound Overview

H. Hacihabiboglu, E. De Sena, Z. Cvetkovic, J. Johnston, J. O. Smith III, "Perceptual Spatial Audio Recording, Simulation, and Rendering: An overview of spatial-audio techniques based on psychoacoustics," IEEE Signal Processing Magazine, 34(3), 36-54, 2017. F. Rumsey, *Spatial audio,* Focal Press, 2012.

S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, F. Zotter "Spatial sound with loudspeakers and its perception: A review of the current state," *Proc. IEEE*, **101**(9):1920–1938, 2013.

Physical Sound-Field Reconstruction

Mark A Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.,* **53**(1):1004–1025, 2005.

J. Daniel, "Spatial sound encoding including near field effect," *Proc. AES 23rd Int. Conf.,* paper #16, 2003.

T. Betlehem, T. D. Abhayapala, "Theory and design of sound field reproduction in reverberant rooms," *J. Acoust. Soc. Amer.,* **117**(4):2100–2111, 2005.

M. Kolundzija, C. Faller, M. Vetterli, "Reproducing Sound Fields Using MIMO Acoustic Channel Inversion," *J. Audio Eng. Soc.*, **59**(10):727–734, 2011.

A. J. Berkhout, D. de Vries, P. Vogel, "Acoustic Control by Wave Field Synthesis," *J. Acoust. Soc. Amer.,* **93**(5):2764–2778, 1993.

E. Hulsebos, D. de Vries, E. Bourdillat, "Improved microphone array configurations for auralization of sound fields by wave-field synthesis," *J. Audio Eng. Soc.* **50**(10):779-790, 2002.

J. Ahrens, S. Spors, "A Modal Analysis of Spatial Discretization of Spherical Loudspeaker Distributions Used for Sound Field Synthesis," *IEEE TrASLP*, **20**(9):2564–2574, 2012.

J. Ahrens, S. Spors, "Sound Field Reproduction Using Planar and Linear Arrays of Loudspeakers," *IEEE TrASLP*, **18**(8):2038–2050, 2010.

Further Reading

Perceptual Sound Field Reconstruction

J. D. Johnston, Y. H. Lam, "Perceptual soundfield reconstruction," *109th AES Conv.,* paper #2399, 2000.

H.-K. Lee, F. Rumsey, "Investigation into the effect of interchannel crosstalk in multichannel microphone technique," *118th AES Conv.*, paper #6405, 2005.

M. Williams, G. Le Du, "Microphone array analysis for multi- channel sound recording," *107th AES Conv.*, paper #4997, 1999.

M. Williams, G. Le Du, "Miultichannel sound recording: Multichannel Microphone Array Design (MMAD)," 2010.

N. V. Franssen, *Stereophony*, Eindhoven, The Netherlands: Philips Research Laboratories, 1964.

L. Simon, R. Mason, F. Rumsey,"Localisation curves for a regularly-spaced octagon loudspeaker array," *127th AES Conv.*, paper #7915, 2010.

S. P. Lipshitz, "Stereo Microphone Techniques... Are the Pursuits Wrong?," *J. Audio Eng. Soc.*, **34**(9):716–744, 1986.

H. Fletcher, *Speech and Hearing in Communication*, New York, NY, USA: van Nostrand, 1953.

Y. Ando, K. Kurihara, "Nonlinear response in evaluating the subjective diffuseness of sound fields," *J. Acoust. Soc. Amer.,* **80**(3)833–836, 1986.

E. De Sena, H. Hacihabiboglu, Z. Cvetkovic, "Analysis and Design of Multichannel Systems for Perceptual Sound Field Reconstruction," *IEEE TrASLP.*, **21**(8):1653–1665, 2013.

E. De Sena, Z. Cvetkovic, "A Computational Model for the Estimation of Localisation Uncertainty," *Proc. ICASSP*, pp. 388–392, 2013.

E. De Sena, Z. Cvetkovic, H. Hacihabiboglu, M. Moonen, and T. van Waterschoot, "Localization Uncertainty in Time-Amplitude Stereophonic Reproduction," *IEEE TrASLP.*, **28**:1000 - 1015, 2020.

Further Reading

E. De Sena, *Analysis, Design and Implementation of Multichannel Audio Systems,* PhD Thesis, King's College London, 2013.

V. Pulkki, "Virtual Source Positioning Using Vector Base Amplitude Panning," *J. Audio Eng. Soc.*, **45**(6):456–466, 1997.

V. Pulkki, "Spatial Sound Reproduction with Directional Audio Coding," *J. Audio Eng. Soc.*, **55**(6):503–516, 2007.

E. De Sena, H. Hacıhabiboğlu, and Z. Cvetković, "On the design and implementation of higher-order differential microphones," IEEE Trans. on Audio, Speech and Language Process., vol. 20, no. 1, pp 162-174, Jan. 2012.

E. De Sena, H. Hacıhabiboğlu, Z. Cvetković, and J. O. Smith III "Efficient Synthesis of Room Acoustics via Scattering Delay Networks," IEEE/ACM Trans. Audio Speech Language Process., vol. 23, no. 9, pp 1478 - 1492, Sept. 2015.

E. De Sena, Niccoló Antonello, Marc Moonen, and Toon van Waterschoot, "On the modeling of rectangular geometries in room acoustic simulations," IEEE/ACM Trans. Audio Speech Language Process., vol. 23, no. 4, Apr. 2015.

Rose, P. Nelson, B. Rafaely, and T. Takeuchi, "Sweet spot size of virtual acoustic imaging systems at asymmetric listener locations," J. Acoust. Soc. Amer., vol. 112, no. 5, pp. 1992- 2002, 2002

F. Rumsey, S. Zielinski, R. Kassier, S. Bech, "On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality", J. Acoust. Soc. Amer., 2005

Stojan Djordjevic, B.Sc. Thesis, University of Surrey, 2019

Ashley Andrew-Jones, B.Sc. Thesis, University of Surrey, 2019

Further Reading

E De Sena, Z Cvetkovic, H Hacihabiboglu, M Moonen, T van Waterschoot, "Localization Uncertainty in Time-Intensity Stereophonic Reproduction", arXiv preprint arXiv:1907.11425 (submitted to IEEE TrASLP), 2019.

E Erdem, E De Sena, H Hacihabiboglu, Z Cvetkovic, "Perceptual Soundfield Reconstruction in Three Dimensions via Sound Field Extrapolation", ICASSP, 2019

F Stevens, D Murphy, L Savioja, V Valimaki "Modeling Sparsely Reflecting Outdoor Acoustic Scenes using the Waveguide Web" *IEEE TrASLP.* 2017.

S Schlecht, E Habets, "On lossless feedback delay networks" *IEEE TrSP*, **65(6)**:1554 - 1564, 2017.

M Williams, "Microphone Array Analysis for Stereo and Multichannel Sound Recording ", Editrice il Rostro, 2004.

E De Sena, H Hacıhabiboğlu, and Z Cvetković, "On the design and implementation of higher-order differential microphones," *IEEE TrASLP*, **20(1)**, 162-174, 2012.

J Meyer, G Elko, "A spherical microphone array for spatial sound recording", *JASA*, **111(5)**, 2346-2346, 2002.

J Smith, "A new approach to digital reverberation using closed waveguide networks ", Proc. 11th Int. Comput. Music Conf., 1985.