Sound Field Analysis and Synthesis: Theoretical Advances and Applications to Spatial Audio Reproduction

Shoichi Koyama

Lecturer, The University of Tokyo

Sound field analysis/synthesis and its applications

Room acoustic analysis

VR/AR audio

Signal enhancement

Basic Technologies of Sound Field Analysis and Synthesis

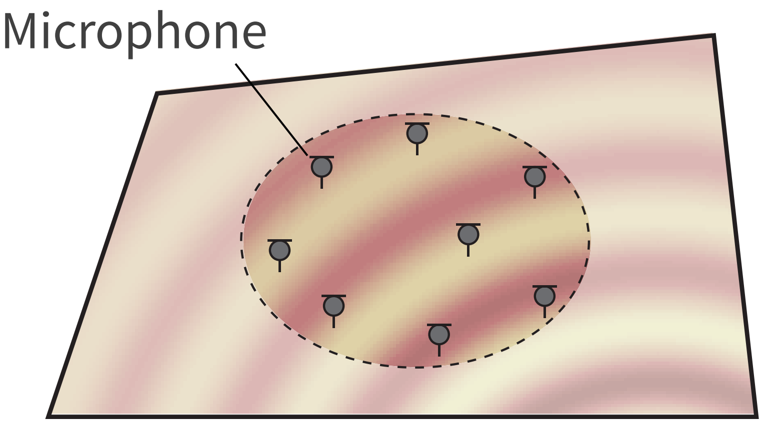
Local-field recording and reproduction

Active noise control

Visualization/auralization

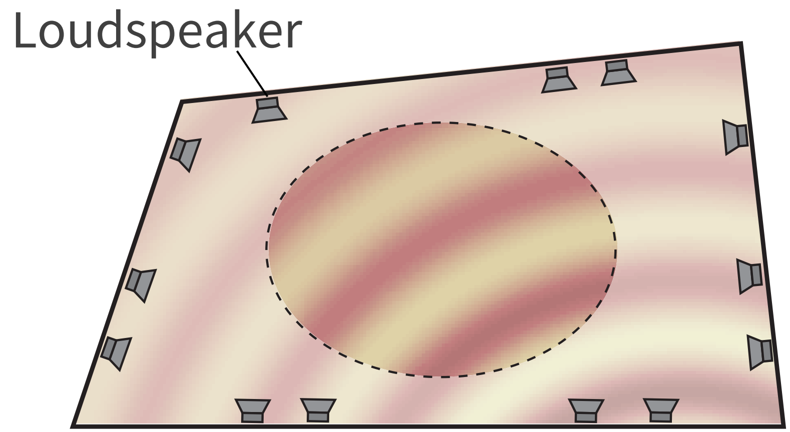
What is sound field analysis/synthesis?

Analysis



Estimating sound field inside target region using multiple mics

Synthesis



Synthesizing desired sound field inside target region using multiple loudspeakers

Sound field analysis techniques incorporating properties of acoustic field into machine learning techniques

Governing equations in acoustic field

Sound propagation is governed by wave equation in time domian and Helmholtz equation in frequency domain

- Sound pressure u at position $\boldsymbol{r} \in \mathbb{R}^3$
  - Wave equation for time t

$$\nabla^2 u(\boldsymbol{r}, t) - \frac{1}{c^2} \frac{\partial^2 u(\boldsymbol{r}, t)}{\partial t^2} = 0$$

⬇ Fourier transform w.r.t. time

  - Helmholtz equation for wave number $k = \omega / c$

$$(\nabla^2 + k^2) u(\boldsymbol{r}, k) = 0$$

➡ Hereafter, all the formulations are in frequency domain

Expansion representations of sound field

- Sound field inside source-free interior regin $u(\boldsymbol{r})$ $(\boldsymbol{r} \in \Omega)$ can be represented by superposition of wave functions
  - Plane wave expansion (a.k.a. Herglotz wave function):

$$u(\boldsymbol{r}) = \int_{\mathbb{S}^2} \tilde{u}(\boldsymbol{\eta}) \mathrm{e}^{\mathrm{j}k\boldsymbol{\eta}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{\eta}$$

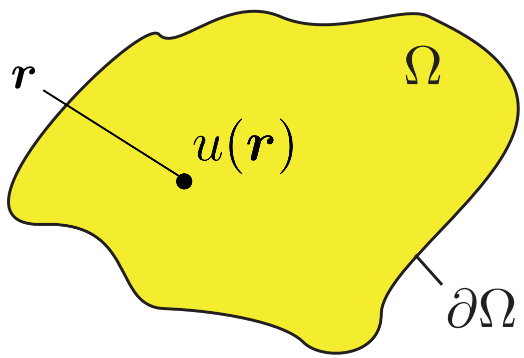
Plane wave function

Unit vector of propagation direction

  - Spherical wave function expansion:

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} j_\nu(kr) Y_\nu^\mu(\theta, \phi)$$

Spherical wave function

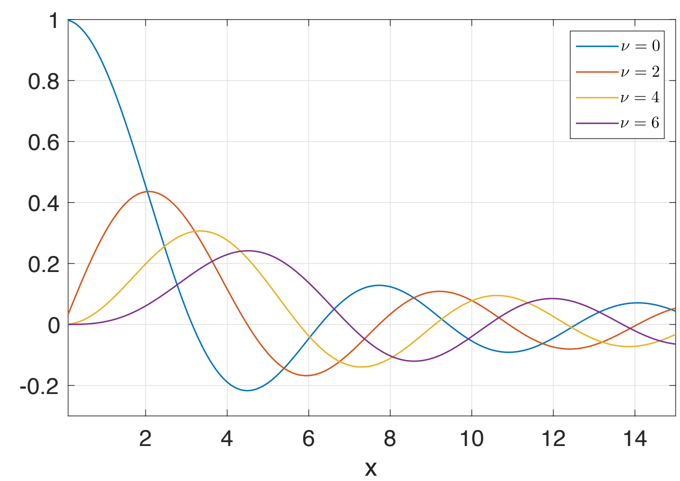


Expansion representations of sound field

- Spherical Bessel function

$$j_\nu(z) = \sqrt{\frac{\pi}{2z}} J_{\nu+1/2}(z)$$

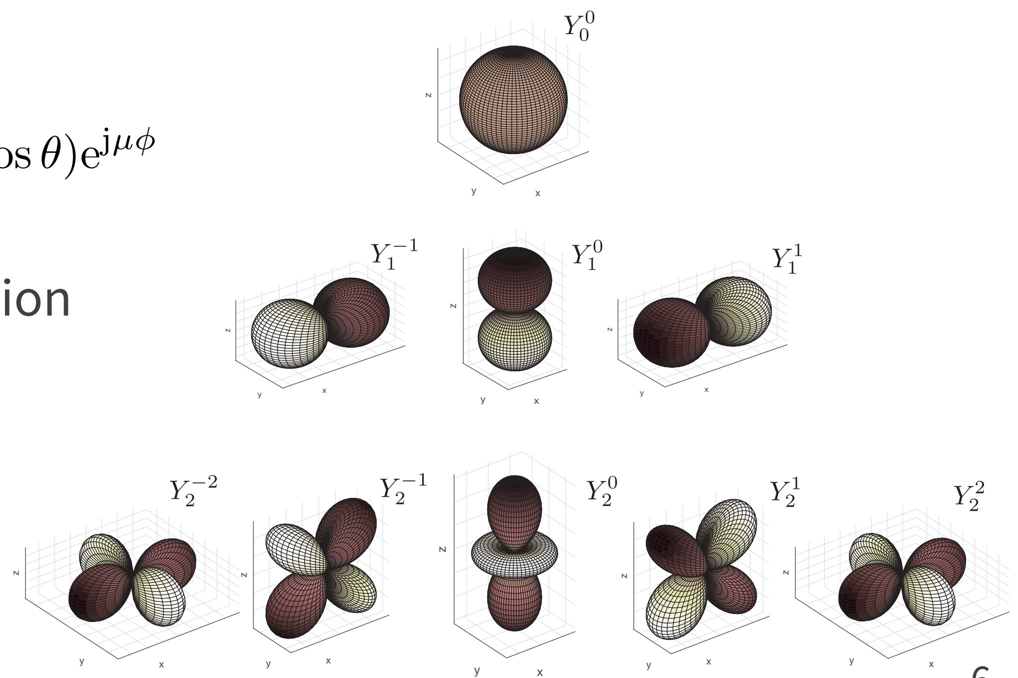
Bessel function



- Spherical harmonic function

$$Y_\nu^\mu(\theta, \phi) = \sqrt{\frac{(2\nu+1)}{4\pi}\frac{(\nu-\mu)!}{(\nu+\mu)!}} P_\nu^\mu(\cos\theta) \mathrm{e}^{\mathrm{j}\mu\phi}$$

Associated Legendre function



Translation in spherical wave function expansion

- Transforming expansion coefficients around $\boldsymbol{r}_0$ into those around $\boldsymbol{r}_0'$
  - Vector of expansion coefficients $\mathring{u}_{\nu,\mu}$ around $\boldsymbol{r}_0$ is defined as $\mathring{\boldsymbol{u}}(\boldsymbol{r}_0) \in \mathbb{C}^\infty$
  - $\mathring{\boldsymbol{u}}(\boldsymbol{r}_0')$ and $\mathring{\boldsymbol{u}}(\boldsymbol{r}_0)$ are related as

$$\mathring{\boldsymbol{u}}(\boldsymbol{r}_0') = \boldsymbol{T}(\boldsymbol{r}_0' - \boldsymbol{r}_0)\mathring{\boldsymbol{u}}(\boldsymbol{r}_0)$$

where $\boldsymbol{T}(\boldsymbol{r}_0' - \boldsymbol{r}_0) \in \mathbb{C}^{\infty\times\infty}$ is translation operator defined by

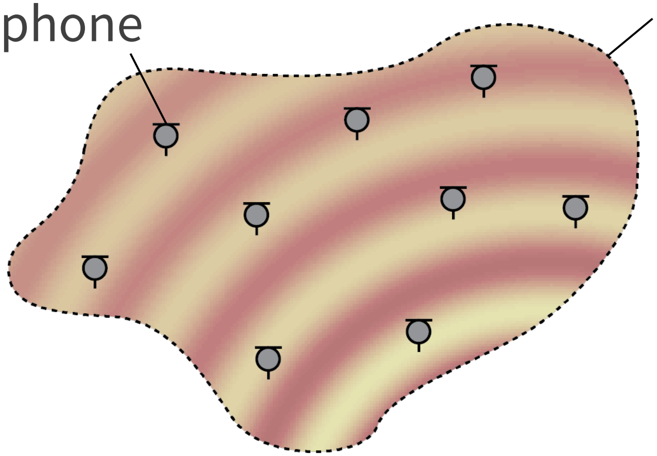
Element of order ν and degree μ

$$[\boldsymbol{T}(\boldsymbol{r})\mathring{\boldsymbol{u}}]_{\nu,\mu} = \sum_{\nu'=0}^{\infty} \sum_{\mu'=-\nu'}^{\nu'} \Bigg[4\pi(-1)^{\mu'} \mathrm{j}^{\nu-\nu'} \times \sum_{l=0}^{\nu+\nu'} \mathrm{j}^l j_n(kr) Y_l^{\mu-\mu'}(\theta,\phi)^* \mathcal{G}(\nu',\mu';\nu,-\mu,l)\Bigg] \mathring{u}_{\nu',\mu'}$$

Gaunt coefficient

Sound field analysis

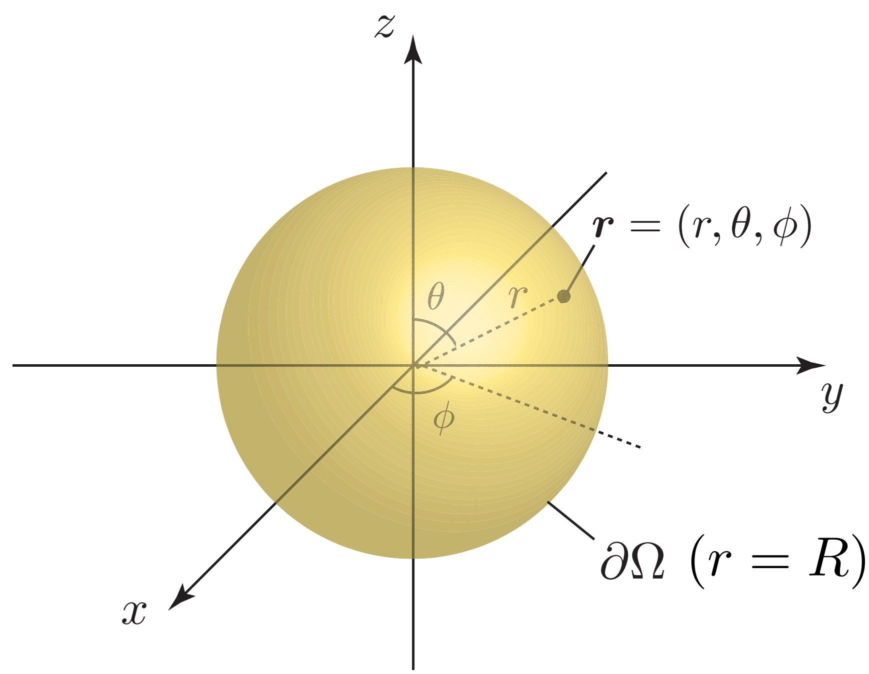
Estimate $u(\boldsymbol{r})$ in target region Ω with observations $\{s_m\}_{m=1}^M$ at discrete set of M mics $\{\boldsymbol{r}_m\}_{m=1}^M$



- Integral-equation-based method
  - Based on discretization of boundary integral equation
- Numerical-optimization-based method
  - Estimation by minimizing some cost function

Sound field analysis for spherical array

- Spherical mic array is typically used in sound field analysis for spatial audio recording
  - Goal is to estimate expansion coefficients of spherical wave functions around array center $\{\mathring{u}_{\nu,\mu}\}$ from observations $\{s_m\}_{m=1}^{M}$ on $\partial\Omega$



Sound field analysis for spherical array

- Integral-equation-based method [Poletti 2005]
  - Spherical harmonic coefficients on $\partial\Omega$ is obtained by

$$U_{\nu,\mu}(R) = \int_0^{2\pi}\int_0^{\pi} u(R,\theta,\phi) Y_\nu^\mu(\theta,\phi)^* \sin\theta \mathrm{d}\theta \mathrm{d}\phi$$

Discretization by M microphone positions on $\partial\Omega$

$$U_{\nu,\mu}(R) = \sum_m \gamma_m u(R,\theta_m,\phi_m) Y_\nu^\mu(\theta_m,\phi_m)^*$$

Weight (γ_m) Observation s_m ($u(R,\theta_m,\phi_m)$)

  - Expansion coefficients $\{\mathring{u}_{\nu,\mu}\}$ are estimated by

$$\hat{\mathring{u}}_{\nu,\mu} = \frac{1}{\sqrt{4\pi} j_\nu(kR)} U_{\nu,\mu}(R)$$

➡ Incomputable when $j_\nu(kR) = 0$! (forbidden frequency problem)

Sound field analysis for spherical array

- Several techniques for avoiding forbidden frequency problem
  1. Mics mounted on rigid spherical baffle
  2. Array of directional mics (e.g., unidirectional mics)
  3. Two (or more) layers of spherical mic array

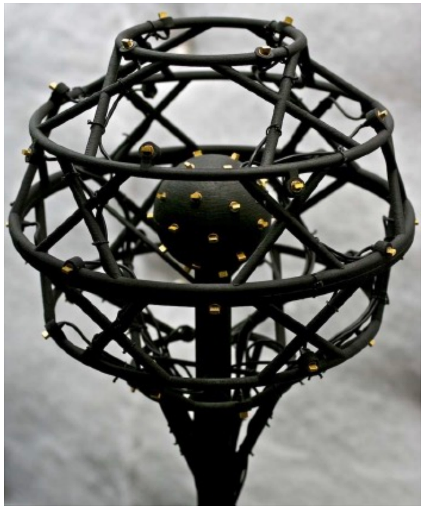
1.

mh acoustics
em32 Eigenmike®

2.

Core Sound
OctoMic™

3.



[Jin+ IEEE/ACM TASLP 2014]

Sound field analysis for spherical array

- Sound field scattered by rigid spherical baffle of radius R

$$u(\boldsymbol{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \mathring{u}_{\nu,\mu} \sqrt{4\pi} \left[j_\nu(kr) - \frac{j'_\nu(kR)}{h'_\nu(kR)} h_\nu(kr) \right] Y_\nu^\mu(\theta, \phi)$$

- Expansion coefficients are estimated by

$$\hat{\mathring{u}}_{\nu,\mu} = \frac{1}{\sqrt{4\pi} \left[j_\nu(kR) - \frac{j'_\nu(kR)}{h'_\nu(kR)} h_\nu(kR) \right]} U_{\nu,\mu}(R)$$

$$= -\frac{\mathrm{j} k^2 R^2}{\sqrt{4\pi}} h'_\nu(kR) U_{\nu,\mu}(R)$$

Does not include spherical Bessel function in denominator

Wronskian

$$j_\nu(z) h'_\nu(z) - j'_\nu(z) h_\nu(z) = \frac{\mathrm{j}}{z^2}$$

➡ Much more robust than open spherical mic array

Sound field analysis for spherical array

- **Numerical-optimization-based method** [Wang+ 1998, Laborie+ 2003]
  - Constructing linear equation of measurements and expansion coefficients by truncating expansion order up to N

$$
\underbrace{\begin{bmatrix} s_1 \\ \vdots \\ s_M \end{bmatrix}}_{\boldsymbol{s} \in \mathbb{C}^M} \approx \underbrace{\begin{bmatrix} \varphi_0^0(\boldsymbol{r}_0) & \cdots & \varphi_N^N(\boldsymbol{r}_0) \\ \vdots & \ddots & \vdots \\ \varphi_0^0(\boldsymbol{r}_M) & \cdots & \varphi_N^N(\boldsymbol{r}_M) \end{bmatrix}}_{\boldsymbol{\Phi} \in \mathbb{C}^{M \times (N+1)^2}} \underbrace{\begin{bmatrix} \mathring{u}_{0,0} \\ \vdots \\ \mathring{u}_{N,N} \end{bmatrix}}_{\mathring{\boldsymbol{u}} \in \mathbb{C}^{(N+1)^2}}
$$

where $\varphi_\nu^\mu(\boldsymbol{r}) = \sqrt{4\pi} j_\nu(kr) Y_\nu^\mu(\theta, \phi)$

  - Expansion coefficients are estimated as least-squares solution

$$
\hat{\mathring{\boldsymbol{u}}} = \left(\boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{\Phi}\right)^{-1} \boldsymbol{\Phi}^{\mathsf{H}} \boldsymbol{s}
$$

➡ Forbidden frequency problem must be avoided in similar way to integral-equation-based method

Summary of conventional analysis methods

- Integral-equation-based method
  - Stable computation and useful for analyzing properties
  - Several well-established methods for avoiding forbidden frequency problem
  - Mics must be on spherical surface, and expansion coefficients at array center are obtained

- Numerical-optimization-based method
  - Arbitrary array geometries are applicable, but expansion center must be set
  - Forbidden frequency problem must be avoided when pressure mics are only on boundary surface
  - Truncation of expansion order is necessary

➡ More accurate and flexible method for sound field analysis?

Infinite-dimensional harmonic analysis

[Ueno+ IEEE SPL 2018]

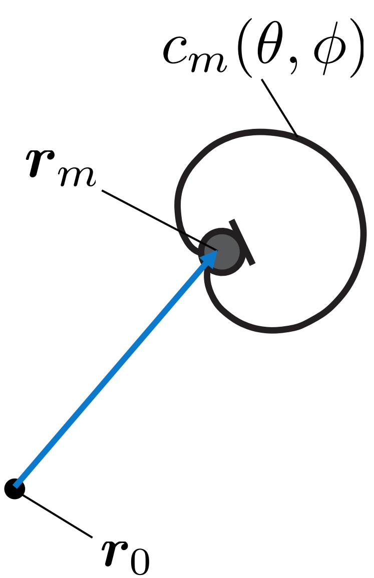
- Consider to use arbitrarily placed directional mics
  - Expansion coefficients of directivity pattern of m th mic $c_m(\theta, \phi)$ is denoted as $\{c_{m,\nu,\mu}\}$

$$c_{m,\nu,\mu} = \frac{(-\mathrm{j})^\nu}{\sqrt{4\pi}} \int_0^\pi \int_0^{2\pi} c_m(\theta,\phi)^* Y_\nu^\mu(\theta,\phi)^* \sin\theta \mathrm{d}\theta \mathrm{d}\phi$$

  - Vectors of $\{\mathring{u}_{\nu,\mu}(\boldsymbol{r}_m)\}$ and $\{c_{m,\nu,\mu}\}$ are denoted as $\mathring{\boldsymbol{u}}(\boldsymbol{r}_m) \in \mathbb{C}^\infty$ and $\boldsymbol{c}_m \in \mathbb{C}^\infty$, respectively
  - Observation of m th mic at $\boldsymbol{r}_m$ is represented by

$$\begin{aligned} s_m &= \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} c_{m,\nu,\mu}^* \mathring{u}_{\nu,\mu}(\boldsymbol{r}_m) \\ &= \boldsymbol{c}_m^{\mathsf{H}} \mathring{\boldsymbol{u}}(\boldsymbol{r}_m) \\ &= \boldsymbol{c}_m^{\mathsf{H}} \boldsymbol{T}(\boldsymbol{r}_m - \boldsymbol{r}_0) \mathring{\boldsymbol{u}}(\boldsymbol{r}_0) \end{aligned}$$

Translation operator



Infinite-dimensional harmonic analysis

[Ueno+ IEEE SPL 2018]

- Modeling of mic observations
  - Observation vector $\boldsymbol{s} \in \mathbb{C}^M$ is written as

$$\boldsymbol{s} = \boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}} \mathring{\boldsymbol{u}}(\boldsymbol{r}_0)$$

where

$$\begin{aligned}\boldsymbol{\Xi}(\boldsymbol{r}_0) &:= \left[(\boldsymbol{c}_1^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_1 - \boldsymbol{r}_0))^{\mathsf{H}} \quad \cdots \quad (\boldsymbol{c}_M^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_M - \boldsymbol{r}_0))^{\mathsf{H}}\right] \\ &= \left[\boldsymbol{T}(\boldsymbol{r}_0 - \boldsymbol{r}_1)\boldsymbol{c}_1 \quad \cdots \quad \boldsymbol{T}(\boldsymbol{r}_0 - \boldsymbol{r}_M)\boldsymbol{c}_M\right]\end{aligned}$$

$$\boldsymbol{T}(\boldsymbol{r})^{\mathsf{H}} = \boldsymbol{T}(-\boldsymbol{r})$$

  - Sensor noise is assumed to be complex Gaussian $\boldsymbol{\epsilon} \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_M)$

$$\boldsymbol{s} = \boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}} \mathring{\boldsymbol{u}}(\boldsymbol{r}_0) + \boldsymbol{\epsilon}$$

➡ Directional mic observations and expansion coefficients around $\boldsymbol{r}_0$ are related

Infinite-dimensional harmonic analysis

[Ueno+ IEEE SPL 2018]

- Expansion coefficients around any position $\boldsymbol{r}$ can be estimated as

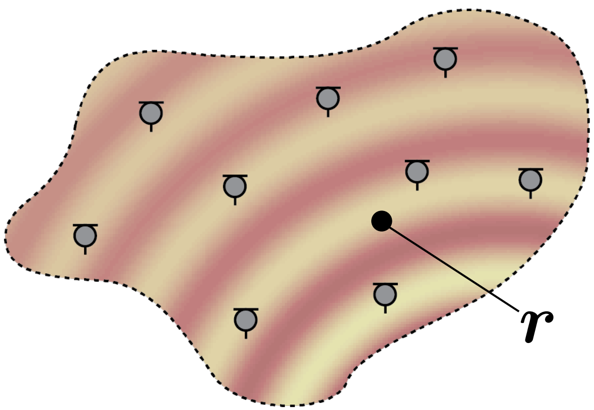
$$\hat{\mathring{\boldsymbol{u}}}(\boldsymbol{r}) = \boldsymbol{T}(\boldsymbol{r}-\boldsymbol{r}_0)\boldsymbol{\Xi}(\boldsymbol{r}_0)\left(\boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}\boldsymbol{\Xi}(\boldsymbol{r}_0)+\sigma^{-2}\boldsymbol{I}_M\right)^{-1}\boldsymbol{s}$$
$$= \boldsymbol{\Xi}(\boldsymbol{r})\left(\boldsymbol{\Psi}+\sigma^{-2}\boldsymbol{I}_M\right)^{-1}\boldsymbol{s}$$

where (m,m') th element of $\boldsymbol{\Psi}=\boldsymbol{\Xi}(\boldsymbol{r}_0)^{\mathsf{H}}\boldsymbol{\Xi}(\boldsymbol{r}_0)\in\mathbb{C}^{M\times M}$ is obtained by

$$(\boldsymbol{\Psi})_{m,m'} = \boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_m-\boldsymbol{r}_0)\boldsymbol{T}(\boldsymbol{r}_0-\boldsymbol{r}_{m'})\boldsymbol{c}_{m'}$$
$$= \boldsymbol{c}_m^{\mathsf{H}}\boldsymbol{T}(\boldsymbol{r}_m-\boldsymbol{r}_{m'})\boldsymbol{c}_{m'}$$

$\boldsymbol{T}(\boldsymbol{r}+\boldsymbol{r}')=\boldsymbol{T}(\boldsymbol{r})\boldsymbol{T}(\boldsymbol{r}')$

Note that $\{c_{m,\nu,\mu}\}$ is generally modelled by low-order coefficients

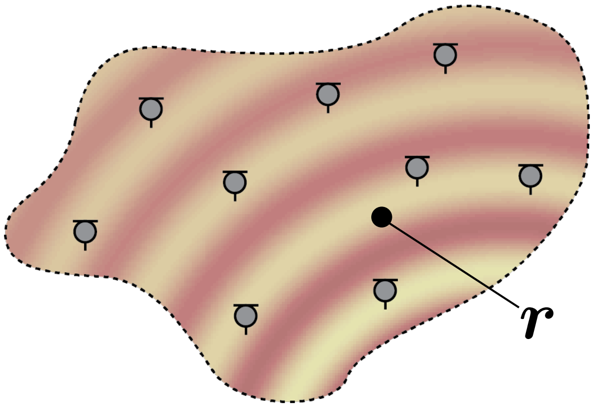


[Ueno+ IEEE SPL 2018]

- Benefits of infinite-dimensional harmonic analysis
  - Arbitrarily placed mics can be used
  - Truncation of expansion order is unnecessary
  - Independent of expansion center

$$\hat{\mathring{\boldsymbol{u}}}(\boldsymbol{r}) = \boldsymbol{\Xi}(\boldsymbol{r}) \left(\boldsymbol{\Psi} + \sigma^{-2} \boldsymbol{I}_M\right)^{-1} \boldsymbol{s}$$

$$(\boldsymbol{\Psi})_{m,m'} = \boldsymbol{c}_m^{\mathsf{H}} \boldsymbol{T}(\boldsymbol{r}_m - \boldsymbol{r}_{m'}) \boldsymbol{c}_{m'}$$



Kernel interpolation of sound field

- Estimation of pressure distribution $u(\boldsymbol{r})$ with pressure (or omnidirectional) mics, i.e., $c_{m,\nu,\mu} = \delta_{\nu,0}\delta_{\mu,0}$

$$
\begin{aligned}
\hat{u}(\boldsymbol{r}) &= \hat{\mathring{u}}_{0,0}(\boldsymbol{r}) \\
&= \begin{bmatrix} j_0(k\|\boldsymbol{r}-\boldsymbol{r}_1\|) & \cdots & j_0(k\|\boldsymbol{r}-\boldsymbol{r}_M\|) \end{bmatrix} \left(\boldsymbol{\Psi} + \sigma^{-2}\boldsymbol{I}_M\right)^{-1} \boldsymbol{s} \\
&= \sum_{m=1}^{M} \left[\left(\boldsymbol{\Psi} + \sigma^{-2}\boldsymbol{I}_M\right)^{-1} \boldsymbol{s}\right]_m j_0(k\|\boldsymbol{r}-\boldsymbol{r}_m\|)
\end{aligned}
$$

$$
\boldsymbol{\Psi} = \begin{bmatrix} j_0(k\|\boldsymbol{r}_1-\boldsymbol{r}_1\|) & \cdots & j_0(k\|\boldsymbol{r}_1-\boldsymbol{r}_M\|) \\ \vdots & \ddots & \vdots \\ j_0(k\|\boldsymbol{r}_M-\boldsymbol{r}_1\|) & \cdots & j_0(k\|\boldsymbol{r}_M-\boldsymbol{r}_M\|) \end{bmatrix}
$$

➡ Corresponding to kernel ridge regression with kernel function of 0th-order spherical Bessel function

Kernel interpolation of sound field

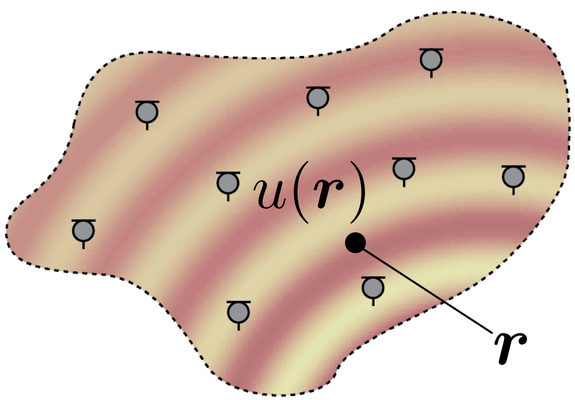
- Consider the following interpolation problem:

$$\underset{u\in\mathcal{H}}{\text{minimize}}\, J := \sum_{m=1}^{M} |u(\boldsymbol{r}_m) - s_m|^2 + \lambda \|u\|_{\mathcal{H}}^2$$

Regularization parameter

  - When solution space $\mathcal{H}$ is reproducing kernel Hilbert space (RKHS), this problem corresponds to kernel ridge regression
  - RKHS is Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle_{\mathcal{H}})$ that has reproducing kernel function $\kappa : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$

$$u(\boldsymbol{r}) = \langle \kappa(\cdot, \boldsymbol{r}), u \rangle_{\mathcal{H}}$$



Kernel interpolation of sound field

- Based on representer theorem, the problem has closed-form solution
  - The solution is represented with $\alpha_m \in \mathbb{C}\ (m = 1, \ldots, M)$ as

$$u(\boldsymbol{r}) = \sum_{m=1}^{M} \alpha_m \kappa(\boldsymbol{r}, \boldsymbol{r}_m)$$

  - Vector of $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_M]^\mathsf{T} \in \mathbb{C}^M$ is obtained by

$$\boldsymbol{\alpha} = (\boldsymbol{K} + \lambda \boldsymbol{I}_M)^{-1} \boldsymbol{s}$$

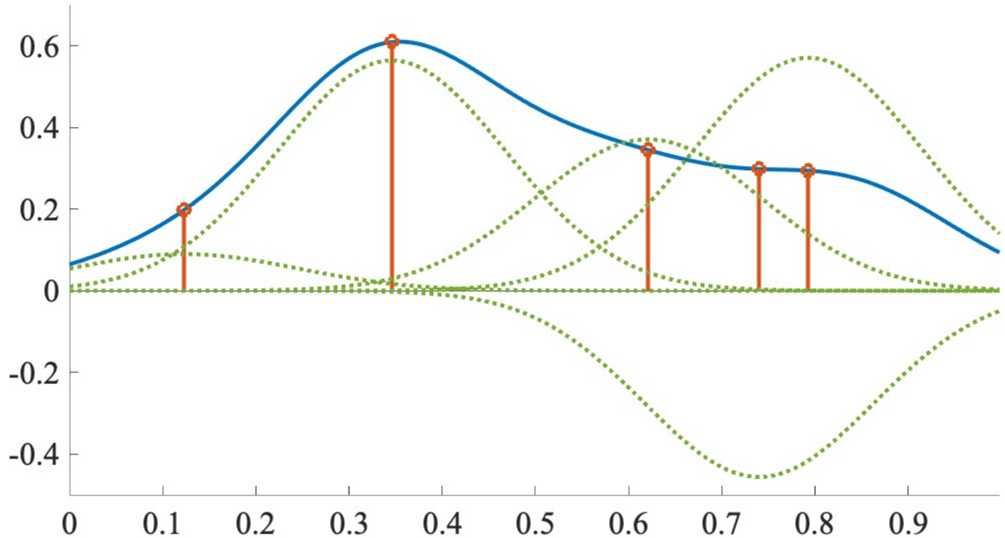
where $\boldsymbol{K} \in \mathbb{C}^{M \times M}$ is Gram matrix defined as

$$\boldsymbol{K} := \begin{bmatrix} \kappa(\boldsymbol{r}_1, \boldsymbol{r}_1) & \cdots & \kappa(\boldsymbol{r}_1, \boldsymbol{r}_M) \\ \vdots & \ddots & \vdots \\ \kappa(\boldsymbol{r}_M, \boldsymbol{r}_1) & \cdots & \kappa(\boldsymbol{r}_M, \boldsymbol{r}_M) \end{bmatrix}$$

Kernel interpolation of sound field

- Interpolation is achieved by linear combination of kernel functions at data points
  - Typically-used kernel function in machine learning is Gaussian kernel

$$\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) = \exp\left(-\frac{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|^2}{\sigma^2}\right)$$



Kernel interpolation of sound field

[Ueno+ IWAENC 2018]

- RKHS for interpolating sound field
  - $(\mathcal{H}, \langle \cdot, \cdot \rangle_{\mathcal{H}})$ is defined as

$$\mathcal{H} := \left\{ \int_{\mathbb{S}^2} \tilde{u}(\boldsymbol{\eta}) \mathrm{e}^{\mathrm{j}k\boldsymbol{\eta}\cdot\boldsymbol{r}} \mathrm{d}\boldsymbol{\eta} \;\middle|\; \tilde{u} \in L_2(\mathbb{S}^2) \right\}$$

$$\langle u_1, u_2 \rangle_{\mathcal{H}} := \int_{\mathbb{S}^2} \tilde{u}(\boldsymbol{\eta})^* \tilde{u}(\boldsymbol{\eta}) \mathrm{d}\boldsymbol{\eta}$$

Any solution of homogeneous Helmholtz equation can be approximated arbitrarily by functions in $\mathcal{H}$

  - Reproducing kernel function for sound field interpolation

$$\kappa(\boldsymbol{r}_1, \boldsymbol{r}_2) = \int_{\mathbb{S}^2} \mathrm{e}^{\mathrm{j}k\boldsymbol{\eta}\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2)} \mathrm{d}\boldsymbol{\eta}$$
$$= j_0(k \|\boldsymbol{r}_1 - \boldsymbol{r}_2\|)$$

Kernel interpolation of sound field

[Ueno+ IWAENC 2018]

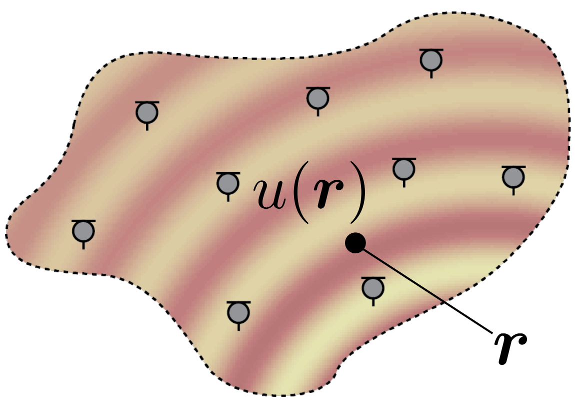
- Pressure distribution is estimated by

$$\hat{u}(\boldsymbol{r}) = \left[(\boldsymbol{K} + \lambda \boldsymbol{I}_M)^{-1} \boldsymbol{s} \right]^{\mathsf{T}} \boldsymbol{\kappa}(\boldsymbol{r})$$

where

$$\boldsymbol{K} := \begin{bmatrix} j_0(k\|\boldsymbol{r}_1 - \boldsymbol{r}_1\|) & \cdots & j_0(k\|\boldsymbol{r}_1 - \boldsymbol{r}_M\|) \\ \vdots & \ddots & \vdots \\ j_0(k\|\boldsymbol{r}_M - \boldsymbol{r}_1\|) & \cdots & j_0(k\|\boldsymbol{r}_M - \boldsymbol{r}_M\|) \end{bmatrix}$$

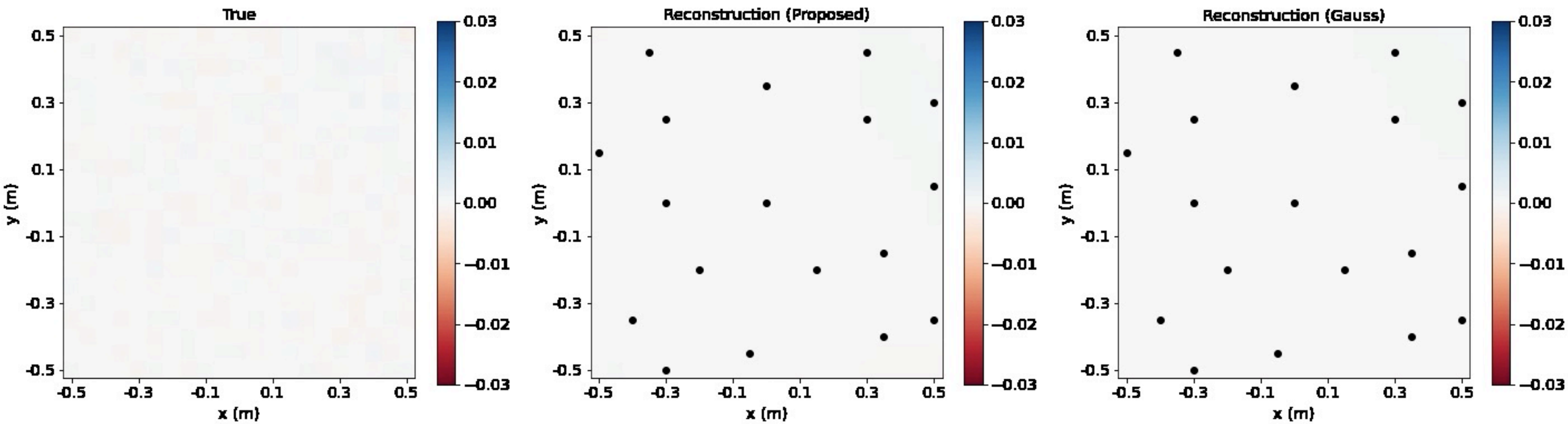
$$\boldsymbol{\kappa}(\boldsymbol{r}) := \begin{bmatrix} j_0(k\|\boldsymbol{r} - \boldsymbol{r}_1\|) & \cdots & j_0(k\|\boldsymbol{r} - \boldsymbol{r}_M\|) \end{bmatrix}^{\mathsf{T}}$$



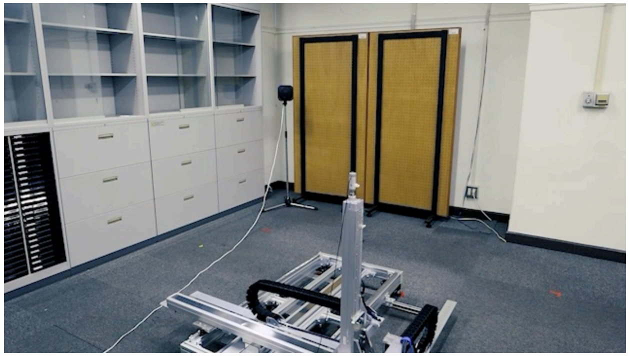
Kernel interpolation of sound field

- Experimental results using real data
  - Reconstructing pulse signal from single loudspeaker w/ 18 mic

True | Proposed | Gaussian kernel



(Black dots indicate mic positions)

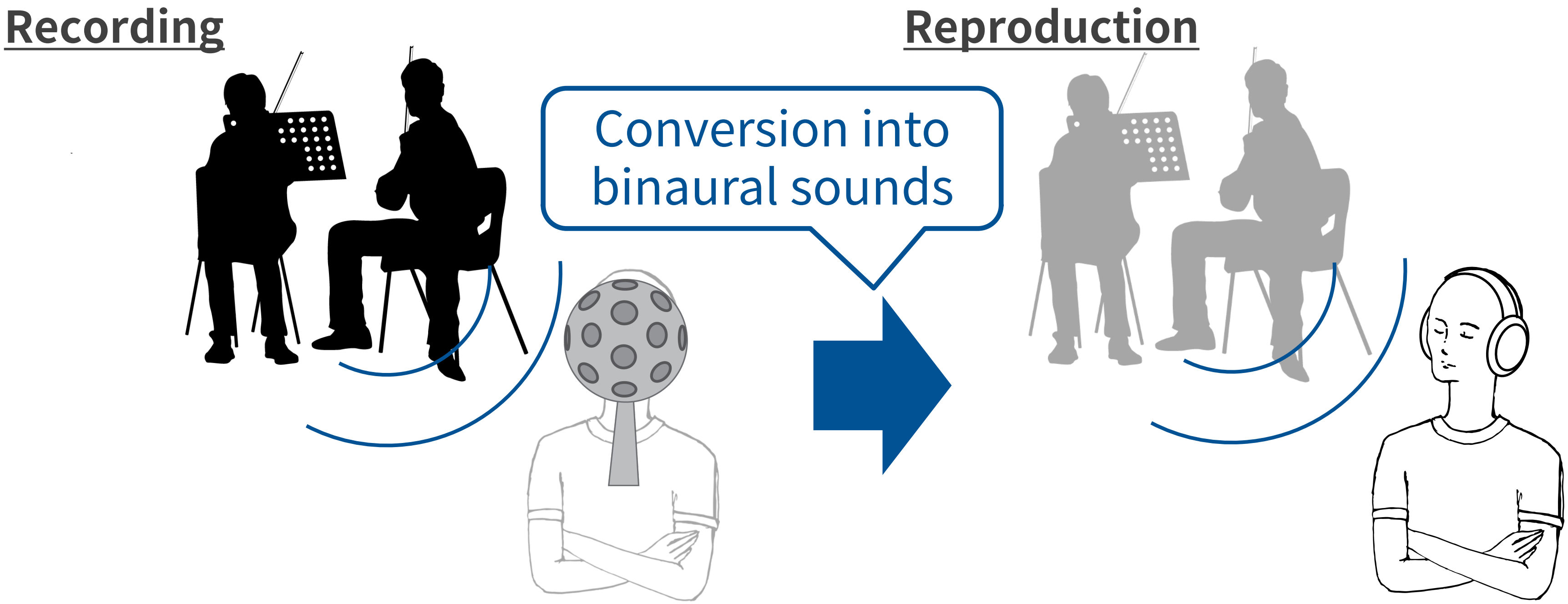


Impulse response measurement system

[Koyama+ IEEE WASPAA 2021]

Application to binaural reproduction

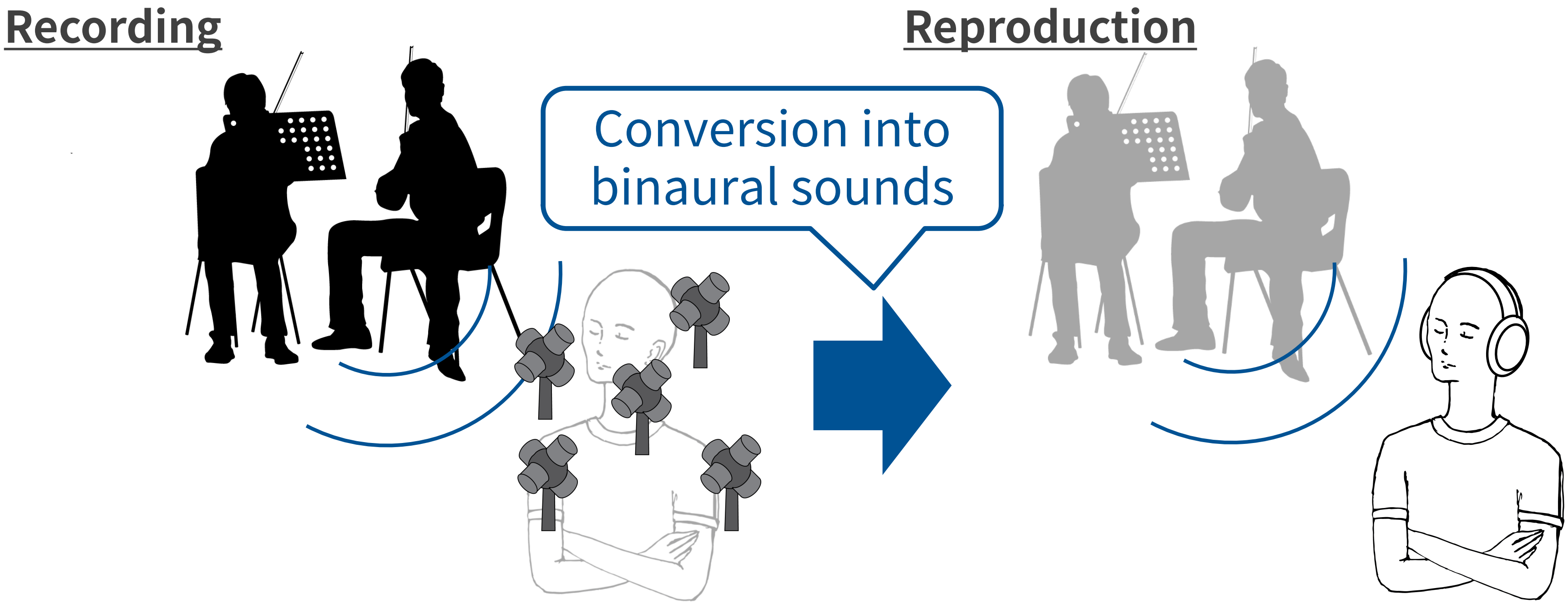
Binaural reproduction from mic array recordings for VR audio



- Binaural reproduction in real world is difficult, comparing with binaural synthesis in VR space
- Conventional spherical array processing requires large-scale system to achieve broad listening area

Application to binaural reproduction

Binaural reproduction from mic array recordings for VR audio

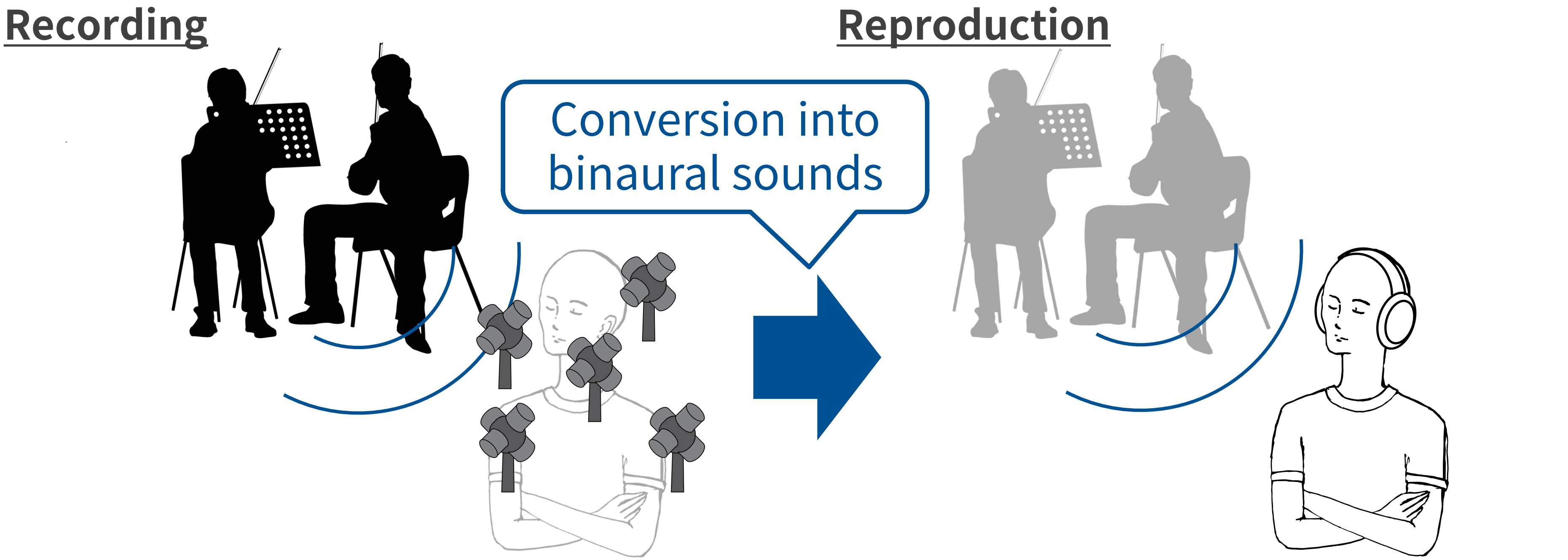


- Binaural reproduction from recordings of multiple small arrays instead of single spherical array
- Broad listening area by using flexible and scalable recording system

[Iijima+ JASA 2021]

Application to binaural reproduction

Binaural reproduction from mic array recordings for VR audio

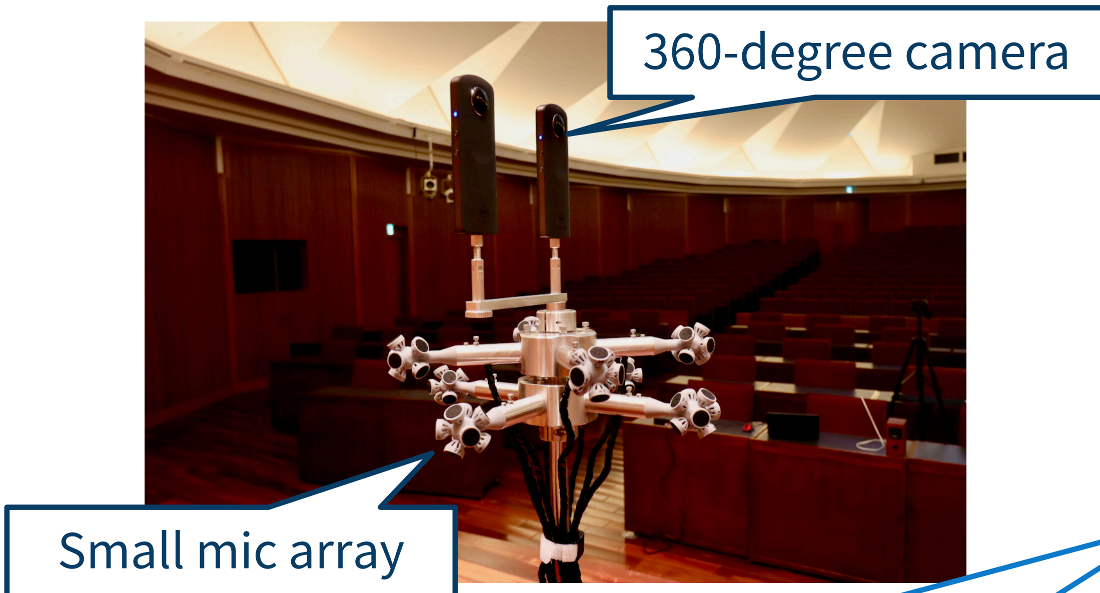


- Recording
  - Sound field estimation by infinite-dimensional harmonic analysis
- Reproduction
  - Binaural rendering from estimated harmonic coefficients with compensation for loudspeaker distance in HRTF measurements

[Iijima+ JASA 2021]

Application to binaural reproduction

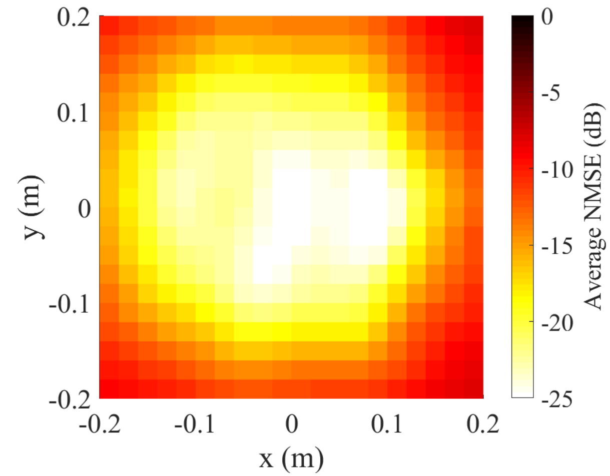
➢ Recording system using multiple small arrays and 360-degree cameras



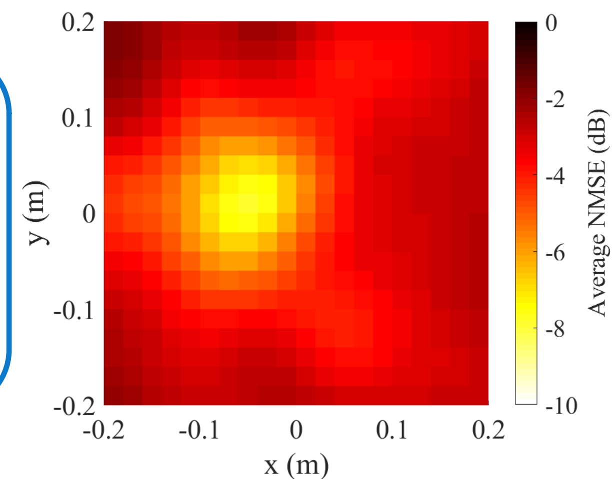
Binaural reproduction error distribution

- High accuracy in broad region
- Rotational and translational head movement is possible (3DoF → 6DoF)

Multiple arrays



Single array



Application to binaural reproduction

- Demo available on Youtube
  - https://youtu.be/tsGIITmQiug
  - Recorded classical music (String Quartet)
  - 360-degree control of viewing and listening directions
  - 2 viewing and listening positions are selectable

Conclusion

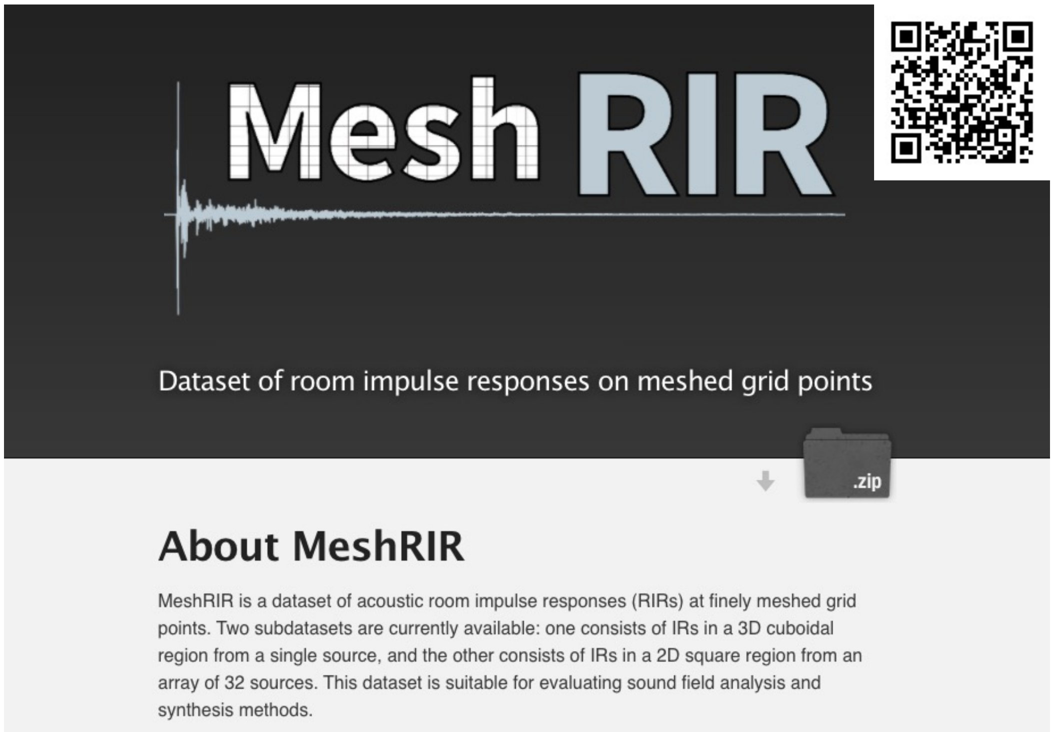
- Recent advances in sound field analysis and synthesis
  - Two categories of sound field analysis methods:
    - Integral-equation-based method
    - Numerical-optimization-based method
  - Introduced infinite-dimensional harmonic analysis
    - Arbitrarily placed mics can be used
    - Truncation of expansion order is unnecessary
    - Independent of expansion center
  - Kernel interpolation of sound field as a special case of infinite-dimensional harmonic analysis
  - Application to binaural reproduction from mic array recordings

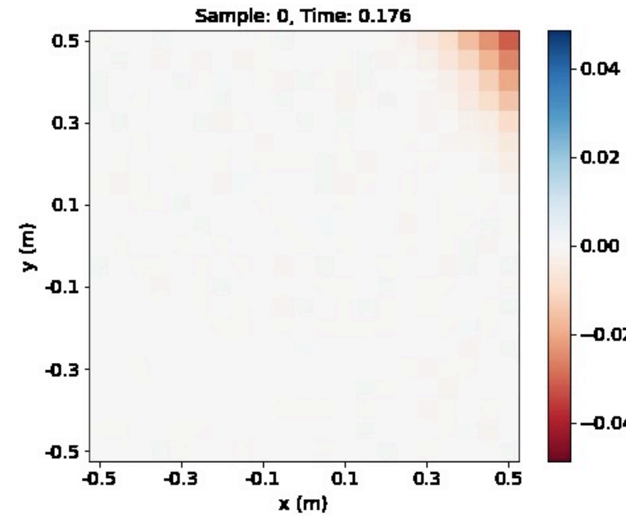
Conclusion

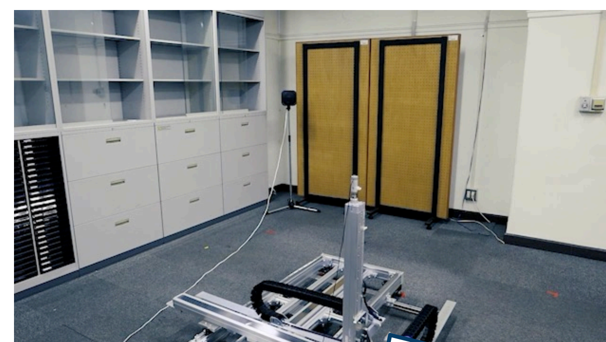
- Related works:
  - Incorporating prior knowledge on approximate source direction by using directional weighting in kernel function to enhance accuracy [Ueno+ IEEE TSP 2021]
  - Adaptation of directional weighting in kernel function from observation data [Horiuchi+ IEEE WASPAA 2021]
  - Application to sound field reproduction by using multiple loudspeakers [Koyama+ I3DA 2021]
  - Application to spatial active noise control [Koyama+ IEEE/ACM TASLP 2021]
  - Example codes of sound field analysis and synthesis techniques are available with room impulse response dataset [Koyama+ IEEE WASPAA 2021]

Dataset of room impulse responses (RIRs)

- Released RIR dataset on meshed grid points with example codes
  - https://sh01k.github.io/MeshRIR/







RIR measurement system

References

- N. Ueno, S. Koyama, and H. Saruwatari, "Sound Field Recording Using Distributed Microphones Based on Harmonic Analysis of Infinite Order," IEEE Signal Process. Lett., 25(1), pp. 135-139, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, "Kernel Ridge Regression With Constraint of Helmholtz Equation for Sound Field Interpolation," in Proc. IWAENC, 2018, pp. 436-440.
- S. Koyama, T. Nishida, K. Kimura, T. Abe, N. Ueno, and J. Brunnström, "MeshRIR: A Dataset of Room Impulse Responses on Meshed Grid Points for Evaluating Sound Field Analysis and Synthesis Methods," in Proc. IEEE WASPAA, 2021.
- N. Ueno, S. Koyama, and H. Saruwatari, "Directionally Weighted Wave Field Estimation Exploiting Prior Information on Source Direction," IEEE Trans. Signal Process., 69, pp. 2383-2395, 2021.
- N. Iijima, S. Koyama, and H. Saruwatari, "Binaural Rendering from Microphone Array Signals of Arbitrary Geometry," J. Acoust. Soc. Am., 150(4), pp. 2479-2491, 2021.
- S. Koyama, K. Kimura, and N. Ueno, "Sound Field Reproduction With Weighted Mode Matching and Infinite-Dimensional Harmonic Analysis: An Experimental Evaluation," in Proc. I3DA, 2021.
- R. Horiuchi, S. Koyama, J. G. C. Ribeiro, N. Ueno, and H. Saruwatari, "Kernel Learning for Sound Field Estimation With L1 and L2 Regularizations," in Proc. IEEE WASPAA, 2021.
- S. Koyama, J. Brunnström, H. Ito, N. Ueno, and H. Saruwatari, "Spatial Active Noise Control Based on Kernel Interpolation of Sound Field," IEEE/ACM Trans. Audio, Speech, Lang. Process., 29, pp. 3052-3063, 2021.