Gridless Sound Field Decomposition Based on Reciprocity Gap Functional in Spherical Harmonic Domain

Yuhta Takida, Shoichi Koyama, Natsuki Ueno and Hiroshi Saruwatari (The University of Tokyo)

Abstract

Sound field decomposition

- Goal is to interpolate and reconstruct sound field inside region including sources (ill-posed problem!)
- Sound field should be decomposed into fundamental solutions of Helmholtz eq., i.e., point sources

Current method: Sparse sound field decomposition[1]

- Sparse decomposition by discretizing possible source region into grid points
- Discrete set of point-source dictionary causes off-grid prblem
- High computational cost and large memory are required for large set of grid points

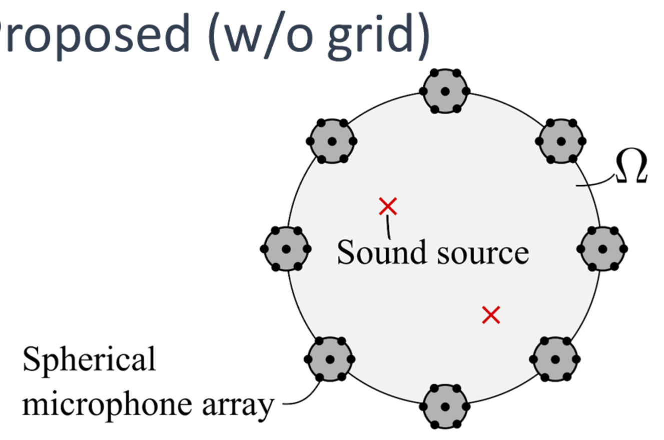
Proposed method

- Reciprocity gap functional (RGF) is applied in spherical harmonic domain
- No more discretization by grids
- Closed-form solution (extremely low computational cost)

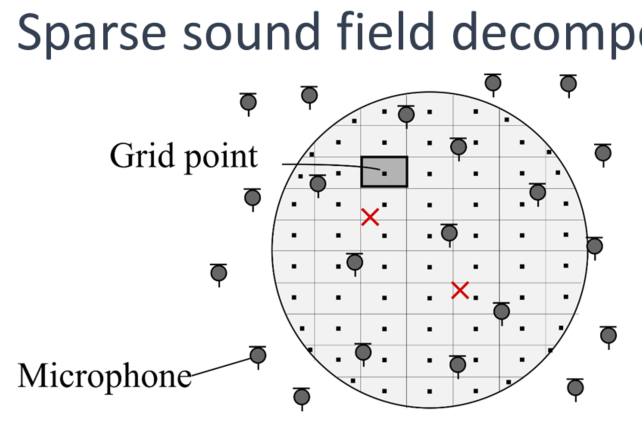
Problem Statement

Sound field decomposition

- Proposed (w/o grid)



- Sparse sound field decomposition



- Source distribution $Q(\mathbf{r},k)$ inside Ω is approximated as a linear combination of J point sources

$$Q(\mathbf{r},k) \approx \sum_{j=1}^{J} c_j \delta(\mathbf{r}-\mathbf{r}_j) \quad \Rightarrow \quad u(\mathbf{r},k) \approx \sum_{j=1}^{J} c_j(k) G(\mathbf{r} \mid \mathbf{r}_j, k)$$

Spatial convolution of source distribution $Q(\cdot)$ with three-dimensional free-field Green's function $G(\cdot)$

Estimating c_j and $\mathbf{r}_j$ makes it possible to reconstruct $u(\cdot)$

Sound Field Decomposition Based on RGF

Concept of RGF

- Test function $w(\cdot)$ and RGF $R(\cdot)$ for $w(\cdot)$ is defined as

$$w(\cdot) \in \mathcal{H}_k := \{w \in \mathcal{H}^1(\Omega) | (\nabla^2 + k^2)w = 0\}$$

$$R(w) := \int_{\Omega} w(\mathbf{r}) Q(\mathbf{r}) d\mathbf{r}$$

- By applying point source assumption and Green's theorem to $R(w)$, the following equation holds

$$\sum_{j=1}^{J} c_j w(\mathbf{r}_j) = \int_{\partial\Omega} \left(u(\mathbf{r}) \frac{\partial w(\mathbf{r})}{\partial \mathbf{n}} - w(\mathbf{r}) \frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} \right) dS \quad \cdots (*)$$

The parameter of sound sources J, $\mathbf{r}_j$ and c_j can be estimated from pressure and velocity values on $\partial\Omega$.

Localization based on RGF

- $w(\cdot)$ is chosen as $w_n(\mathbf{r}) = (x+iy)^n e^{-ikz}$ [2]
- Calculate s_n (Right-hand side of Eq. (*)) from microphone observations
- Compose Hankel matrices

$$\mathbf{H} = \begin{bmatrix} s_1 & s_2 & \cdots & s'_J \\ s_2 & s_3 & \cdots & s_{J'+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{J'} & s_{J'+1} & \cdots & s_{2J'-1} \end{bmatrix}, \quad \mathbf{H}_l = \begin{bmatrix} s_{l+1} & s_{l+2} & \cdots & s_{l+J'} \\ s_{l+2} & s_{l+3} & \cdots & s_{l+J'+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{l+J'} & s_{l+J'+1} & \cdots & s_{l+2J'-1} \end{bmatrix}$$

Determine # of sources J from the eigenvalue of $\mathbf{H}$

Estimate the source locations on x-y plane from the eigenvalue of $\mathbf{H}_1\mathbf{H}_0^{-1}$

A large number of microphone pairs need to be arranged uniformly on $\partial\Omega$ to calculate surface integral.

RGF in Spherical Harmonic Domain

RGF in spherical harmonic domain

- Pressure on $\partial\Omega$ in spherical harmonic domain.

$$u(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \beta_{\nu\mu} h_\nu(kr) Y_{\nu\mu}(\theta,\phi)$$

- Spherical harmonic expansion of $w_n(\cdot)$ is analytically obtained

$$w_{\nu\mu}^{(n)} = \frac{1}{j_\nu(kr)} \int_0^{\pi}\int_0^{2\pi} w_{\nu\mu}^{(n)} Y_{\nu\mu}^*(\theta,\phi) \sin\theta d\theta d\phi$$

$$= \frac{i^{n-\nu}}{k^n} \sqrt{\frac{4\pi(2\nu+1)(\nu+n)!}{(\nu-n)!}} \delta_{\mu,n}$$

Surface integral in spherical harmonic domain

$$s_n = \int_{\partial\Omega} \left(u(\mathbf{r}) \frac{\partial w(\mathbf{r})}{\partial \mathbf{n}} - w(\mathbf{r}) \frac{\partial u(\mathbf{r})}{\partial \mathbf{n}} \right) dS$$

$$= \frac{i}{k} \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} (-1)^{\mu+1} w_{\nu-u}^{(n)} \beta_{\nu\mu}$$

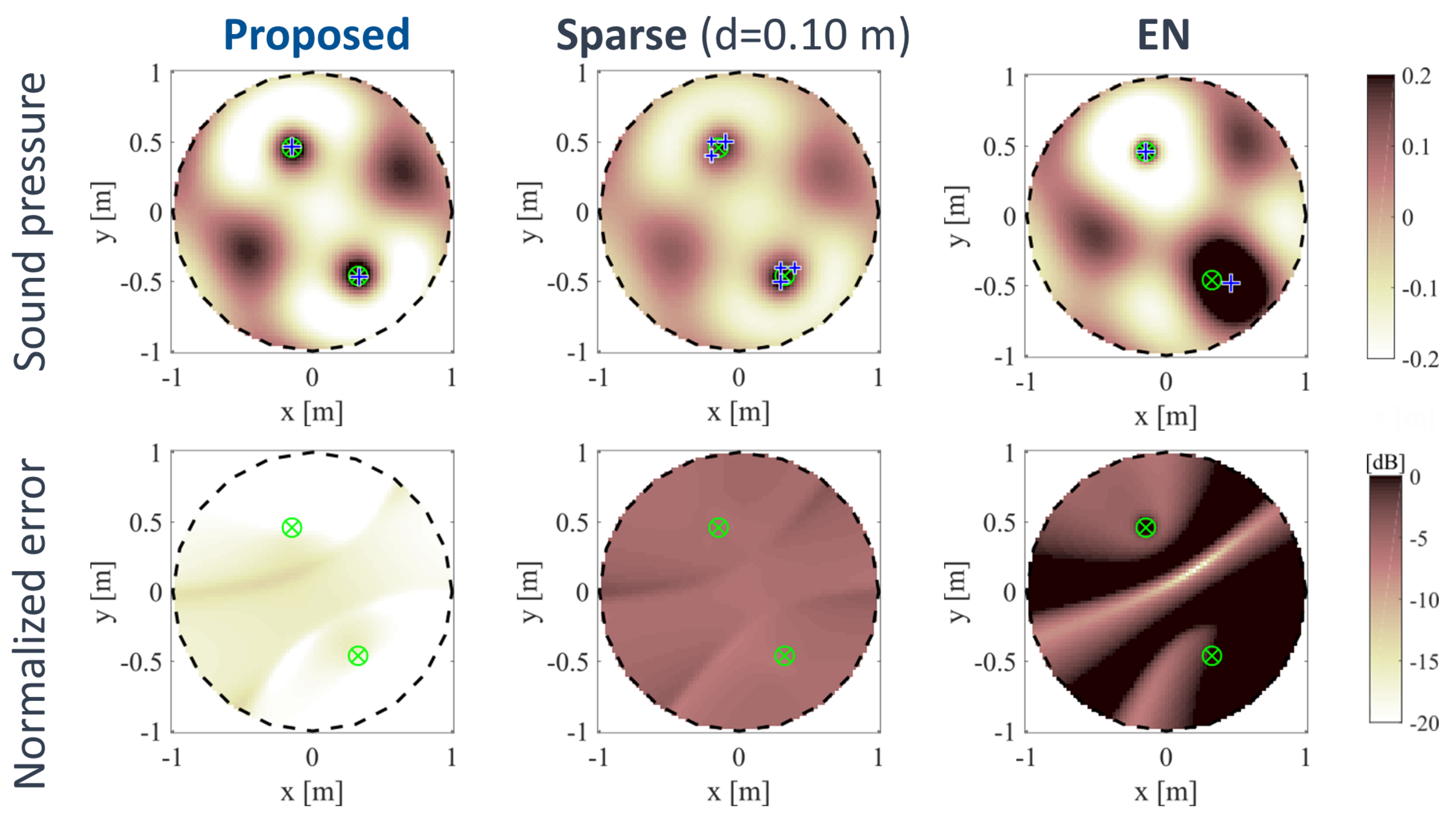
- Two test functions are also used to determine the source location in a 3D space: $w'_n(\mathbf{r}) := w_n(y,z,x)$, $w''_n(\mathbf{r}) := w_n(z,x,y)$
- Analytical spherical harmonic expansion of function $w'_n(\mathbf{r})$, $w''_n(\mathbf{r})$ are obtained from coefficients $w_{\nu\mu}^{(n)}$ and Wigner-D matrix

Experiments

Simulation conditions

- Comparing proposed method (Proposed) with
  - Sparse: sparse sound field decomposition[1] (interval of grid points d of 0.10, 0.15, and 0.20 m)
  - EN: the direct application of RGF[2]
- Microphone array (SNR: 40dB)
  - 24 second-order spherical microphone arrays consisting of 9 microphones for Proposed and Sparse
  - 108 microphone pairs for EN

Sound field reconstruction: 400 Hz



$\mathrm{RMSE_{SL}}$ and SDR are plotted against frequency

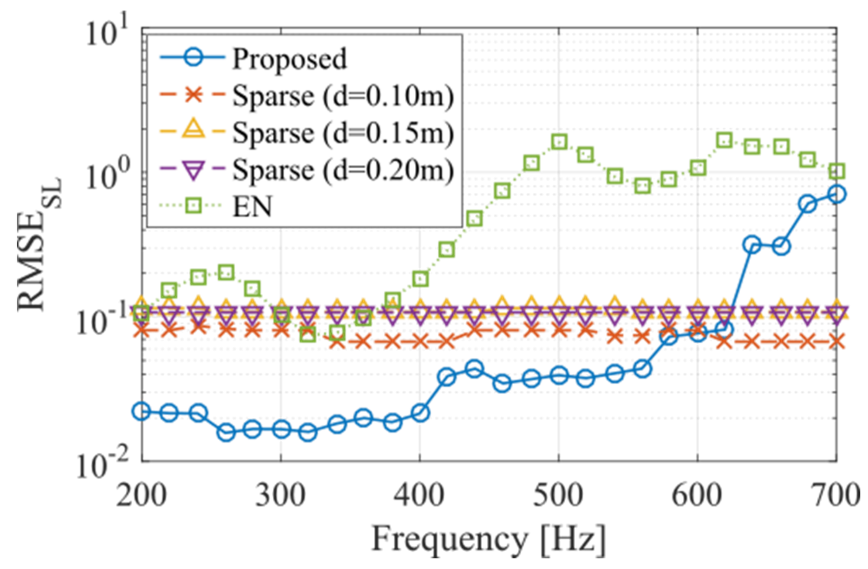
- Evaluation criteria:
  - Root-mean-square error of source location ($\mathrm{RMSE_{SL}}$)
  - Signal-to-distortion ratio (SDR)

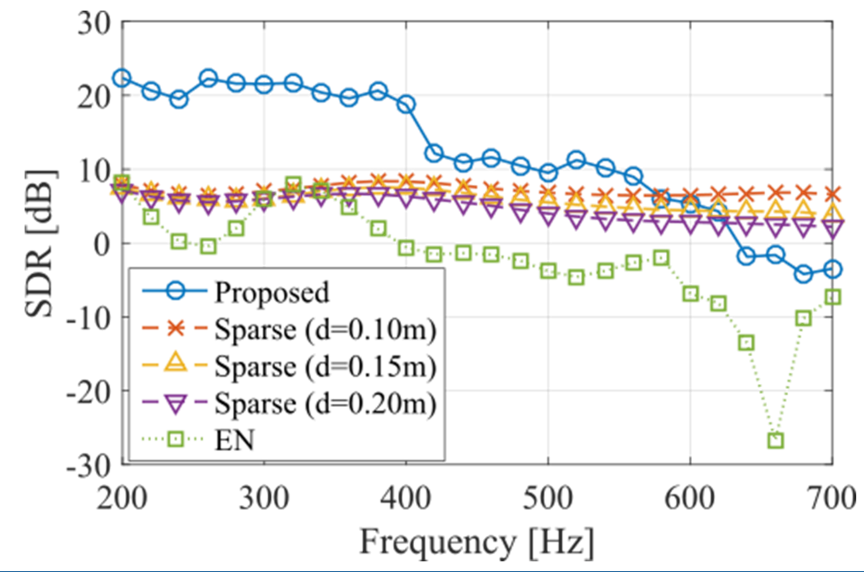
$$\mathrm{RMSE_{SL}} = \sqrt{\frac{1}{J}\sum_{j=1}^{J} \|\mathbf{r}_{j,\mathrm{true}} - \mathbf{r}_j\|_2^2}$$

$$\mathrm{SDR} = 10\log_{10} \frac{\int_\Omega |u_{\mathrm{P,true}}(\mathbf{r})|^2 d\mathbf{r}}{\int_\Omega |u_{\mathrm{P,true}}(\mathbf{r}) - \hat{u}_{\mathrm{P}}(\mathbf{r})|^2 d\mathbf{r}}$$

Computational time (avg.)

- Proposed: 1.97×10^{-4} s
- Sparse
  - d=0.10 m: 1.64×10^{2} s
  - d=0.15 m: 1.88×10^{1} s
  - d=0.20 m: 5.80 s
- EN: 1.77×10^{-4} s





High source localization and sound field reconstruction accuracies are achieved at low frequencies by using proposed method

[1] S. Koyama, S. Shimauchi, and H. Ohmuro, "Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts", in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Florence, May 2014, pp. 4443—4447.

[2] A. El Badia and T. Nara, "An inverse source problem for Helmholtz's equation from the Cauchy data with a single wave number," *Inverse Prob.*, vol. 27, no. 105001, 2011.