Robust Gridless Sound Field Decomposition Based on Structured Reciprocity Gap Function 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

Proposed method and its relation to prior works
Gridless sound field decomposition\[1\] Sparse sound field decomposition\[2\]
(w/o grid points) (w/ grid points)

- Based on spherical-harmonic-domain reciprocity gap functional (SHD-RGF)
- Reconstruction accuracy will be strongly affected by noise

Grouping time-frequency bins to improve robustness

- Group-sparse representation\[2\]
- Exploiting group-sparse structure in time-frequency domain
- Off-grid problem is still a major issue

Structured SHD-RGF and Proposed Algorithm

Construction of AF for decomposition
- Construct AF represented by polynomial function
  \( H(q) = \prod_{j=1}^{J} \left( 1 - p_j q^{-1} \right) \sum_{j=1}^{J} h_j q^{-j} \)
  Its roots correspond to source locations \( \{ p_j \}_{j=1}^{J} \), \( p_j = x_j + iy_j \)
- Convolution coefficient sequence \( \{ h_j \}_{j=0}^{\infty} \) and elements \( h_n = \sum_{k=-N}^{N} h_k Y_{\ell}^{m} (\theta, \phi) \)
  \( \ell = 0, \ldots, J \)
- Annihilation holds for multiple time-frequency (T-F) bins when the source locations are assumed to be static for \( T \) time frames.

Proposed AF-based Algorithm for SHD-RGF
- Optimization problem for SHD-RGF using AF
  \( \min_{\mathbf{u}_T} \| \alpha_{\mathbf{f}_T} - \mathbf{T}_{\mathbf{f}_T} \mathbf{u}_T \|_2^2 \) : Minimization of model error
  such that \( h^T s_{\mathbf{f}_T} = 0 \) : Annihilation and regularization

Sound Field Decomposition Based on SHD-RGF

Concept of RGF
- Test function \( w_n(\cdot) \) and RGF \( R(\cdot) \) for \( w_n(\cdot) \) is defined as
  \( w_n(r) := e^{p r} \sin \left( k_z r \right) \)
  \( R(w_n) := \int_{\Omega} w_n(r) R(r) \, dr \)
- By applying point source assumption and Green’s theorem to \( R(w_n) \), the following equation holds:
  \[ \sum_{j=1}^{J} c_j w_n(r_j) = \int_{\Omega} \left( u(r) \frac{\partial u(r)}{\partial n} - w(r) \frac{\partial u(r)}{\partial n} \right) \, dS \quad \cdots (\ast) \]
  The parameter of sound sources \( J, r_j, c_j \) can be estimated from pressure and velocity values on \( \Omega \).

Numerical Simulations

Simulation conditions
- Comparing proposed method (G-RGF) with
  - G-Sparse: group sparse sound field decomposition\[2\]
  - RGF: RGF in spherical harmonic domain\[3\] for single time-frequency bin
  - Microphone array
  - 24-second-order spherical microphone arrays
  - Four-ring geometry (nine arrays on each ring)
  - Evaluation criteria
    - Root-mean-square error of source location

Results

- Highest source localization accuracy is achieved by G-RGF for all group of time-frequency bands
- Off-grid problem is avoided on the basis of RGF
- Robustness against noise is improved by grouping T-F bins