Paper #8883

SPARSITY-BASED SOUND FIELD SEPARATION IN THE SPHERICAL HARMONICS DOMAIN

POLITECNICO MILANO 1863 — VNIVERSITAT DE VALÈNCIA

M. Pezzoli, M. Cobos, F. Antonacci, A. Sarti*

Politecnico di Milano, Italy - Dipartimento di Elettronica, Informazione e Bioingegneria

**Universitat de València, Spain – Departamento de Informática*

spat — ISPL Image and Sound Processing Lab

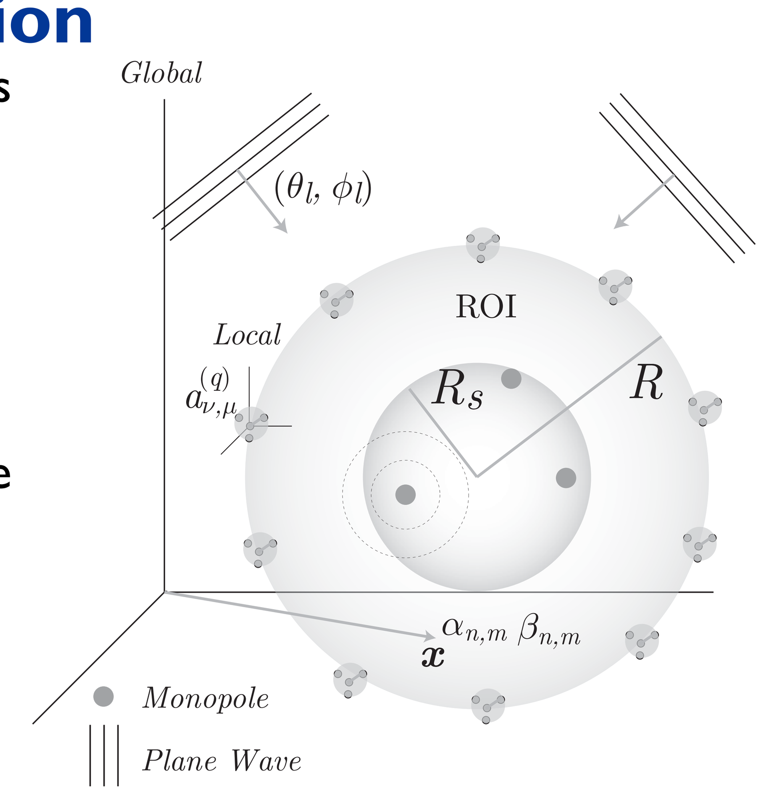
Abstract

Sound field analysis and reconstruction has been a topic of intense research in the last decades for its multiple applications in spatial audio processing tasks. In this context, the identification of the direct and reverberant sound field components is a problem of great interest, where several solutions exploiting spherical harmonics representations have already been proposed. However, the available techniques demand a large number of high-order microphones (HOMs) and high computational power in order to fulfill the necessary spatial sampling requirements, which can only be reduced by prior information obtained through acoustic measurements. Inspired by compressed sensing approaches, this paper proposes an alternative sparse formulation for estimating the exterior and interior sound field components in the spherical harmonics domain that allows to reduce hardware requirements without the need for additional acoustic measurements. The results show that a considerable reduction in the number of HOMs can be achieved while improving the estimation of the sound field components.

1. Data Model and Problem Formulation

The sound field inside a region of interest ROI of radius R is expressed in the spherical harmonics as the sum of the exterior and interior components

$$P_E(\boldsymbol{x},k) = \sum_{n=0}^{N_E}\sum_{m=-n}^{n}\beta_{nm}(k)h_n(kr)Y_{nm}(\theta,\phi),$$
$$P_I(\boldsymbol{x},k) = \sum_{n=0}^{N_I}\sum_{m=-n}^{n}\alpha_{nm}(k)j_n(kr)Y_{nm}(\theta,\phi),$$



- β and α are the "global" exterior and interior coefficients. Their values completely define the sound field inside the ROI.

Given Q Vth order HOMs their spherical harmonics signal is then given in matrix form as

$$\boldsymbol{a} = \boldsymbol{E}\boldsymbol{P}_E\boldsymbol{\beta} + \boldsymbol{E}\boldsymbol{P}_I\boldsymbol{\alpha}$$

$\boldsymbol{P}_E \in \mathbb{C}^{Q(V+1)^2\times(N_E+1)^2}$, $\boldsymbol{P}_I \in \mathbb{C}^{Q(V+1)^2\times(N_I+1)^2}$ model the propagation of the exterior $\boldsymbol{\beta} \in \mathbb{C}^{(N_E+1)^2\times 1}$ and interior $\boldsymbol{\alpha} \in \mathbb{C}^{(N_I+1)^2\times 1}$ coefficients, while $\boldsymbol{E} \in \mathbb{C}^{Q(V+1)^2\times Q(V+1)^2}$ encodes the HOMs signals in the "local" spherical harmonics $\boldsymbol{a}$

GOAL:

Estimate the global coefficients $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ from the HOMs to perform sound field separation

2. Proposed Sparsity-based Spherical Harmonics model (S-SH)

Express exterior and interior coefficients as a sparse set of monopoles and plane waves

- Exterior field: grid of G omnidirectional sources (translated monopoles):
$$\beta_{nm}(k) \approx \sum_{g=1}^{G} c_{0,0}(k)\sqrt{4\pi}\, j_n(kr_g')Y_{mn}^*(\theta_g',\phi_g')$$
- Interior field: distribution of L plane waves :
$$\alpha_{nm}(k) \approx \sum_{l=1}^{L}\sqrt{4\pi}\,(-i)^n Y_{mn}^*(\theta_l,\phi_l)$$

Signal model:

$$\boldsymbol{a} = \boldsymbol{E}\boldsymbol{P}_E\boldsymbol{B}\boldsymbol{y} + \boldsymbol{E}\boldsymbol{P}_I\boldsymbol{W}\boldsymbol{u},$$

$\boldsymbol{B} \in \mathbb{C}^{(N_E+1)^2\times G}$, $\boldsymbol{W} \in \mathbb{C}^{(N_I+1)^2\times L}$ are the dictionaries of monopoles and plane waves, respectively.

- The weights vectors $\mathbf{y} \in \mathbb{C}^{G\times 1}$ and $\mathbf{u} \in \mathbb{C}^{L\times 1}$ are found solving the sparse optimization

$$\arg\min_{y,u} \|\boldsymbol{y}\|_1 + \gamma\|\boldsymbol{u}\|_1, \qquad \text{s.t. } \boldsymbol{a} = \boldsymbol{E}\boldsymbol{P}_E\boldsymbol{B}\boldsymbol{y} + \boldsymbol{E}\boldsymbol{P}_I\boldsymbol{W}\boldsymbol{u},$$

- Similarly to [1], optimal solution $\boldsymbol{y}^*, \boldsymbol{u}^*$ is found using ADMM

Estimate of global coefficients:

$$\hat{\boldsymbol{\beta}} = \boldsymbol{B}\boldsymbol{y}^*, \qquad \hat{\boldsymbol{\alpha}} = \boldsymbol{W}\boldsymbol{u}^*$$

3. Simulation Results

Setup and metrics

- $5\text{m}\times 8\text{m}\times 3\text{m}$ simulated room with 1st order HOM (V=1)
- ROI of radius $R = 1\text{m}$, $G = 100$ grid points and $L = 360$ plane wave directions
- Results compared with:
  - (SS): Sparsity-based sound field reconstruction technique in [1]
  - (VL): Spherical-harmonics-based sound field separation method in [2]
- **Normalized mean square error** evaluated on time-domain signal p at test points $\boldsymbol{x}_t$ inside the ROI:

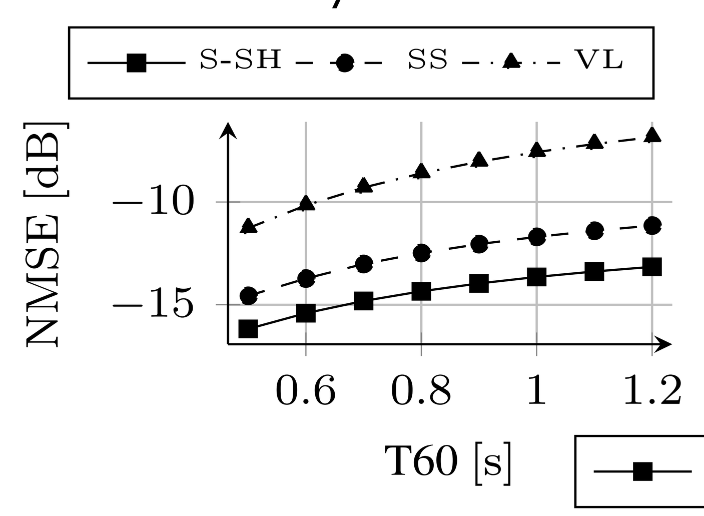
$$\text{NMSE}(\boldsymbol{x}_t) = 10\log_{10}\left(\frac{1}{\text{T}}\frac{\sum_{\tau=1}^{\text{T}}\left(\hat{p}(\boldsymbol{x}_t,\tau) - p(\boldsymbol{x}_t,\tau)\right)^2}{\sum_{\tau=1}^{\text{T}} p(\boldsymbol{x}_t,\tau)^2}\right);$$

- **Direct-to-reverberant ratio** of the estimated direct sound field $\hat{p}_E$

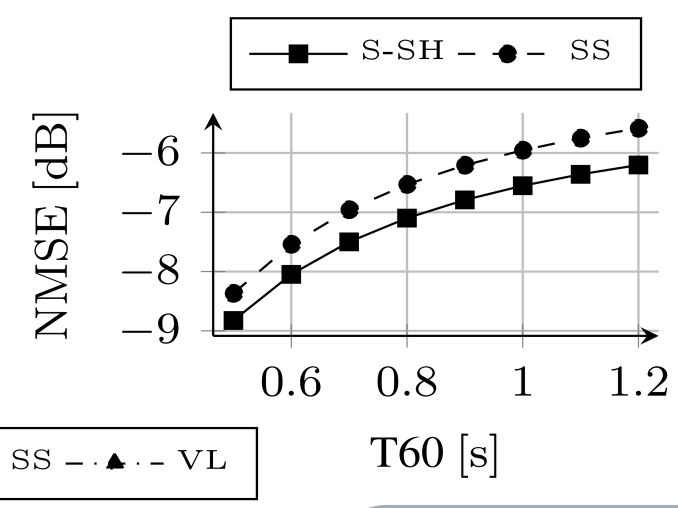
$$\text{DRR}(\boldsymbol{x}_t) = 10\log_{10}\left(\frac{1}{\text{T}}\frac{\sum_{\tau=\tau_0}^{\tau_0+C}\hat{p}_E(\boldsymbol{x}_t,\tau)^2}{\sum_{\tau=\tau_0+C}^{\text{T}}\hat{p}_E(\boldsymbol{x}_t,\tau)^2}\right).$$

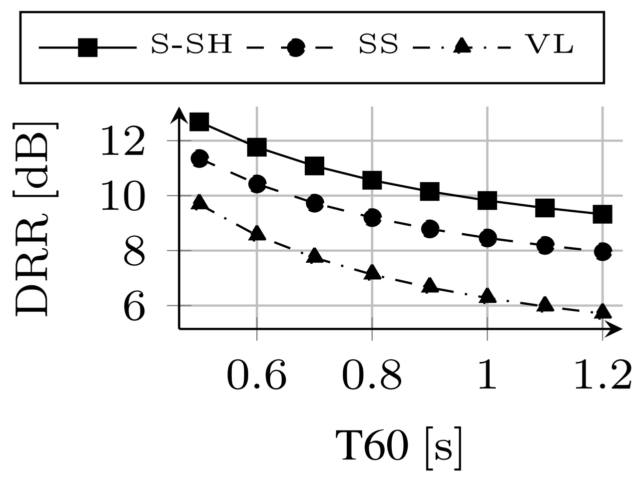
Results varying room T60

Exterior Field only



Full sound field (exterior + interior)

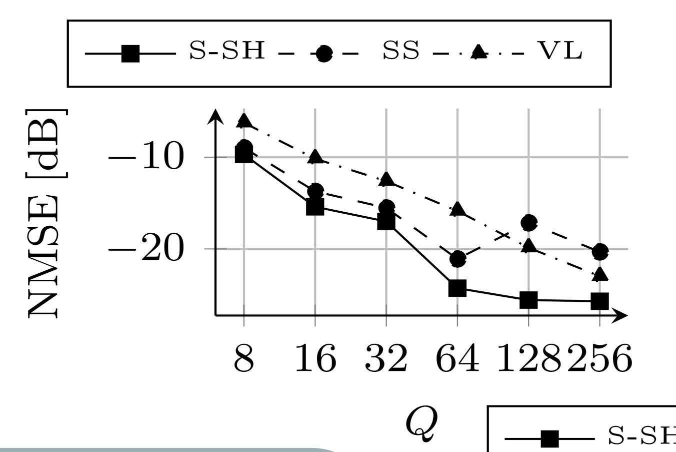




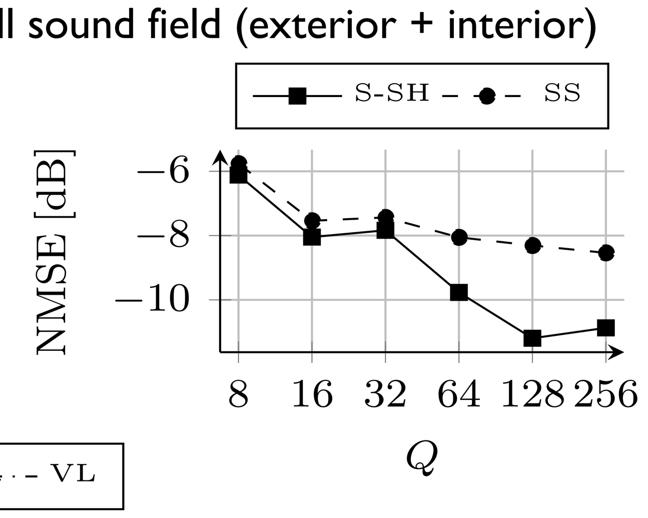
Improved estimation of sound field components & sound field separation

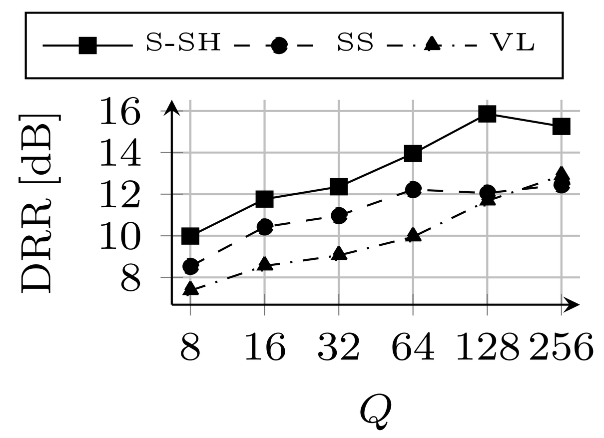
Results varying number of HOMs

Exterior Field only



Full sound field (exterior + interior)





References

[1] S. Koyama and L. Daudet, "Sparse representation of a spatial sound field in a reverberant environment," IEEE J. Sel. Top. Signal Process., vol. 13, no. 1, pp. 172–184, 2019.

[2] F. Borra, S. Krenn, I. D. Gebru, and D. Markovic´, "1st-order microphone array system for large area sound field recording and reconstruction: Discussion and preliminary results," in Workshop Appl. Signal Process. Audio Acoust. IEEE, 2019, pp. 378–382

2022 IEEE International Conference on Acoustics Speech and Signal Processing, Singapore, Singapore, 22-27 May 2022

icassp 2022 Singapore