Sparsity Based Framework for Spatial Sound Reproduction in Spherical Harmonic Domain

Gyanajyoti Routray and Rajesh M Hegde
Indian Institute of Technology, Kanpur
Email: {groutray, rhegde} @iitk.ac.in

***Abstract*—In this paper, a novel sparsity based framework is proposed for accurate spatial sound field reproduction in spherical harmonic domain. The proposed framework can effectively reduce the number of loudspeakers required to reproduce the desired sound field using higher order ambisonics (HOA) over a fixed listening area. Although HOA provides accurate reproduction of spatial sound, it has a disadvantage in terms of the restriction on the area of sound reproduction. This area can be increased with the increase in the number of loudspeakers during reproduction. In order to limit the use of a large number of loudspeakers the sparse nature of the weight vector in the HOA signal model is utilized in this work. The problem of obtaining the weight vector is first formulated as a constrained optimization problem which is difficult to solve due to orthogonality property of the spherical harmonic matrix. This problem is therefore reformulated to exploit the sparse nature of the weight vector. The solution is then obtained by using the Bregman iteration method. Experiments on sound field reproduction in free space using the proposed sparsity based method are conducted using loudspeaker arrays. Performance improvements are noted when compared to least squares and compressed sensing methods in terms of sound field reproduction accuracy, subjective, and objective evaluations.**

I. Introduction

Reproduction of spatial sound field, with high accuracy within a desired area is a fundamental problem in spatial audio processing. Plane wave decomposition using spherical harmonics provides a basis for rendering in 3D space. Some of the popular reproductions techniques based on loudspeaker arrays are higher-order Ambisonics (HOA) [1], wave field synthesis (WFS) [2], and model approaches such as vector base amplitude panning (VBAP) [3]. These techniques spatially encode audio signals, and decode them to generate feeds to loudspeaker arrays or headphones. Decoding of HOA signal can be done by matching the spherical harmonics modes produced by the loudspeakers with the modes of ambisonic sound field decomposition [4]–[7]. However in such decoding schemes there is an undesired variation in the loudness levels, which can be overcome by preserving the decoded energy [8], maintaining constant angular spread across source direction [9]. These rendering techniques suffer from smaller error free reproduction area (sweet spot) [6], [10]. With higher frequency this sweet spot reduces to smaller than the radius of human head. However with higher order ambisonics, spatial resolution can be improved [8], [11], [12]. Another approach is to sample the spherical harmonics at the loudspeaker positions, and obtain a suitable panning function [13]–[16]. Such methods need a dense sampling for accurate spatial sound field reproduction. Such dense sampling requirements intuitively leads to an investigation into sparsity based methods for solving this problem in an efficient manner.

In this paper, a sparsity based framework is proposed to compute loudspeaker weights and subsequently the speaker feeds of a loudspeaker array for rendering spatial sound fields. This frame work is based on the fact that loudspeakers act as point sources in near field [10]. The theoretical contribution of this work lies in the development of a constrained optimization framework to formulate and compute the sparse weight vector corresponding to the loudspeaker feeds. Additionally this method can be used in the accurate reproduction of sound fields at any listening position in free space using a reduced number of loudspeakers. Performance comparisons using reduced number of loudspeakers, with results obtained by direct matrix inversion (HOA Reconstruction), least square (LS), and compressed sensing(CS) [5], [7] methods are indeed motivating.

The rest of paper is organized as follows, Section 2 introduces the system model for sound field reproduction in free space and describes the proposed method for finding optimum weights of loudspeaker array. In Section 3 performance of proposed method is evaluated and compared with existing methods. Section 4 concludes the paper.

II. System Model and Problem Formulation

In this section the HOA based system model for sound field reproduction is first developed. The weight vector in the system model is assumed to be sparse and a constrained optimization problem is formulated to compute it. Bregman iteration method is then used solve this problem and obtain the weights for loudspeaker feeds. An algorithm for sound field rendering using the proposed method is also discussed in this context.

A. System Model for Sound Field Reproduction using Loudspeaker Array

Consider a normalized plane wave having frequency f in the direction $\Theta_s \triangleq (\theta_s, \phi_s)$. Where θ and ϕ describes the elevation from top and azimuth respectively. The incident sound field

This work was funded by the SERB - DST under project no. SERB/EE/2017242

ISBN 978-90-827970-1-5 © EURASIP 2018

at a point $\mathbf{r} = (r, \Theta)$ due to a plane wave in the direction Θ_s expressed as [10]

$$p(k;\mathbf{r}) = e^{-ik\mathbf{r}} = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} 4\pi i^n j_n(kr)[Y_n^m(\Theta_s)]^* Y_n^m(\Theta) \tag{1}$$

where $\mathbf{Y}_n^m(\Theta_s) = [Y_0^0(\Theta_s), Y_1^{-1}(\Theta_s), \ldots, Y_N^m(\Theta_s)]^T$, is the spherical harmonic coefficients in which $n = 1, \ldots, N$ and $m = -n, \ldots, n$ be the order and mode respectively. The spherical harmonics are defined as

$$Y_n^m(\Theta_q) = \frac{1}{2}\sqrt{\frac{(2n+1)(n-|m|)!}{\pi(n+|m|)!}} P_n^{|m|}(\cos\theta_q) e^{im\phi_q} \tag{2}$$

where $P_n^m(\cdot)$ is the associated Legendre function, $j_n(\cdot)$ is the nth order spherical Bessel function of first kind, and related to ordinary Bessel function as $j_n(x) = \sqrt{\frac{\pi}{2x}} J_{n+1/2}(x)$. Reproduction of a sound field with an array of loudspeakers arranged on sphere having radius r_p can be given as [10]

$$T(k;\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} X_n(kr) \sum_{l=1}^{L} a(k;\Theta_l)[Y_n^m(\Theta_l)]^* Y_n^m(\Theta) \tag{3}$$

where $X_n(kr)$ is defined as

$$X_n(kr) = \begin{cases} 4\pi(-i)kh_n(kr_p)j_n(kr) & \text{for } r < r_p \\ 4\pi(-i)kh_n(kr)j_n(kr_p) & \text{for } r > r_p \end{cases} \tag{4}$$

and $h_n(x) = \sqrt{\frac{\pi}{2x}}[J_{n+1/2}(x) - i\mathcal{J}_{n+1/2}(x)]$ is the nth order spherical Hankel function of second kind, $\mathcal{J}_n(\cdot)$ is the nth order Bessel function of second kind. $\|\cdot\|$ represents the Euclidean norm. Equating (1) and (3) and writing in matrix form

$$\mathbf{Pa} = \mathbf{u} \tag{5}$$

Truncating the order of spherical harmonics to N in [10], the loudspeaker weights can be obtained by solving the linear expression in (5).

B. Problem Formulation

Computation of the loudspeaker weights $a(k;\Theta_l)$ in (5) can be formulated as a constrained optimization problem of the form

$$\min_{\mathbf{a}} \quad \|\mathbf{a}\|_p \qquad \text{s.t.} \quad \mathbf{Pa} = \mathbf{u} \tag{6}$$

where $\|\cdot\|_p$ is the l_p vector norm, defined as $\|\mathbf{a}\|_p = (\sum_i |a_i|^p)^{1/p}$. The solution for the weighting coefficients depends on the choice of norms. Selecting $p = 2$ leads to the conventional least square (LS) minimization problem. The LS solution yields a closed form expression given by

$$\mathbf{a} = \mathbf{P^H}(\mathbf{PP^H} + \lambda\mathbf{I})^{-1}\mathbf{u} \tag{7}$$

where $\mathbf{I}$ is the identity matrix and λ is the regularization parameter. When $\lambda = 0$, the solution is the pseudoinverse of the spherical harmonics matrix, which is also known as HOA based decoding in literature. Although HOA provides accurate reproduction of spatial sound, it has a disadvantage in terms of the restriction on the area of sound reproduction. This area can be increased with the increase in the number of loudspeakers during reproduction. Inorder to limit the use of a large number of loudspeakers the sparse nature of the weight vector in the HOA signal model can be exploited. One of the standard approaches to obtain a sparse solution in this context is compressed sensing (CS). In this approach the l_1 norm is minimized. This problem can be modeled by relaxing the equality constraint using error tolerance $\epsilon \geq 0$ as [17]

$$\min_{\mathbf{a}} \quad \|\mathbf{a}\|_1 \qquad \text{s.t.} \quad \|\mathbf{Pa} - \mathbf{u}\|_2 \leq \epsilon \tag{8}$$

This method gives the optimized result in far field where the sound sources act as plane wave sources. In spatial sound field productions the arrangement of loudspeakers make the speakers act as point source which further decrease the performance of these methods.

C. Development of Sparse Iterative Framework for Decoding

In order to develop the proposed sparsity based framework we reformulate the problem in (8) as a LASSO formulation

$$\min_{\mathbf{a}} \quad \frac{1}{2}\|\mathbf{Pa} - \mathbf{u}\|_2^2 + \lambda\|\mathbf{a}\|_1 \tag{9}$$

With increase in the number of loudspeakers the panning vector $\mathbf{a}$ becomes sparse. On the other hand the matrix $\mathbf{P}$ containing the spherical harmonics also becomes sparse in the spatial domain. Thus it can be expressed as

$$\mathbf{P} = \mathbf{DR} \tag{10}$$

where $\mathbf{D} \in \mathbb{C}^{(N+1)^2 \times \tau}$ is an overcomplete dictionary matrix and $\mathbf{R} \in \mathbb{R}^{\tau \times L}$ is the sparse coefficient matrix. The Spherical harmonic matrix exhibits orthogonality i.e. $\mathbf{PP}^H \approx \mathbf{I}$. A necessary condition is that the matrix $\mathbf{P}$ should orthonormal i.e. $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. This condition is further incorporated in the problem formulation as

$$\begin{aligned} \min_{\mathbf{P},\mathbf{R}} \quad & \tfrac{1}{2}\|\mathbf{u} - \mathbf{Pa}\|_2^2 + \lambda\|\mathbf{R}\|_1 \\ \text{s.t.} \quad & \mathbf{P} = \mathbf{DR} \\ & \mathbf{R}^T\mathbf{R} = \mathbf{I} \\ \mathbf{a} = \quad & (\mathbf{P^H P})^{\mathbf{-1}}\mathbf{P^H u} \end{aligned} \tag{11}$$

The above problem involves orthogonality constraints and the formulation therefore becomes a non convex problem. For solving such problems iterative methods have been proposed in [18]–[20]. The Bregman iteration method is used herein to solve for the weights under question.

1) Bregman Iteration method for Loudspeaker Array Gain Calculation: The above problem is solved using Bregman splitting iteration formulation in line with our previous work in [19]. The problem in (11) is first rewritten as

$$\begin{aligned} 1.\quad & (\mathbf{P}^n, \mathbf{R}^n) = \min_{\mathbf{P},\mathbf{R}} \frac{1}{2}\|\mathbf{u} - \mathbf{Pa}\|_2^2 + \lambda\|\mathbf{R}\|_1 \\ & \qquad + \frac{\alpha}{2}\|\mathbf{R} - \mathbf{\Psi}^{n-1} + \mathbf{B}^{n-1}\|_F^2 \\ & \text{s.t.} \quad \mathbf{P} = \mathbf{DR} \\ 2.\quad & \mathbf{\Psi}^n = \min_{\Psi} \frac{\alpha}{2}\|\mathbf{\Psi} - (\mathbf{R}^n + \mathbf{B}^{n-1})\|_F^2, \\ & \text{s.t.} \quad \mathbf{\Psi^T \Psi} = \mathbf{I} \\ 3.\quad & \mathbf{B}^n = \mathbf{B}^{n-1} + \mathbf{R}^n - \mathbf{\Psi}^n \\ 4.\quad & \mathbf{a} = (\mathbf{P^H P})^{\mathbf{-1}}\mathbf{P^H u} \end{aligned} \tag{12}$$

In this formulation the subproblem in (12) is an unconstrained convex problem, and $\mathbf{\Psi}$ has a closed from solution given by $\mathbf{\Psi}_{opt} = \mathbf{UVT}^{-1/2}\mathbf{V}^T$,where $\mathbf{U} = \mathbf{R}^n + \mathbf{B}^{n-1}$, and $\mathbf{T}$ is a diagonal matrix satisfying SVD factorization $\mathbf{U}^T\mathbf{U} = \mathbf{VTV}^T$. The spherical harmonic function $Y_0^0(\Theta)$ has a constant value irrespective of the location (azimuth and elevation). Hence additional linear constraint can be imposed in the formulation as

$$-\mathbf{1}^T\epsilon \leq (\mathbf{P}_n^T\mathbf{e}_1 - 1^T\mathbf{Y}_0^0) \leq \mathbf{1}^T\epsilon \tag{13}$$

where $\mathbf{e}_1 = [1, \ldots, 0]^T$ is an eigen vector. This constraint provides an initial information about spherical harmonic coefficient matrix $\mathbf{P}$ and helps in obtaining an optimal solution. Applying these constraints a standard optimization problem can be formulated as

$$\begin{aligned} \min_{\mathbf{a},\mathbf{R}} \quad & \tfrac{1}{2}\|\mathbf{u} - \mathbf{Pa}\|_2^2 + \lambda\|\mathbf{R}\|_1 + \tfrac{\alpha}{2}\|\mathbf{R} - \mathbf{\Psi}^{n-1} + \mathbf{B}^{n-1}\|_F^2 \\ \text{s.t.} \quad & \mathbf{P} = \mathbf{DR} \\ & \mathbf{R}^T\mathbf{R} = \mathbf{I} \\ & -\mathbf{1}^T\epsilon \leq (\mathbf{P}_n^T\mathbf{e}_1 - 1^T\mathbf{Y}_0^0) \leq \mathbf{1}^T\epsilon \end{aligned} \tag{14}$$

It is important to note that orthonormal constraints also appear in the above formulation making the objective function non-convex. Algorithm1 describes the complete steps involved in obtaining the optimum weights and the corresponding feeds for the array of loudspeakers in a sparse manner.

Algorithm 1 Algorithm for sound field rendering using the proposed sparsity based method

```
Choose the radius of the reproduction area and corresponding spherical harmonic order N
Obtain the observations $\mathbf{u}$
Create an overcomplete dictionary $\mathbf{D}$ with required resolution in $\theta$ and $\phi$
Choose the optimum parameter $\alpha, \lambda$,and $\epsilon$
Choose the number of iterations ittr
Initialize $\mathbf{a^0}$
Initialize matrix $\mathbf{\Psi^0}$ randomly and $\mathbf{B}^0$ to a matrix of zeros
for n := 1 to ittr do
    Solve the minimization problem in (12) to find $\mathbf{R}^n$
    Assign $\mathbf{U} = \mathbf{R}^n + \mathbf{B}^{n-1}$
    Compute the SVD form$\mathbf{U}^T\mathbf{U} = \mathbf{VTV}^T$
    Update $\mathbf{\Psi}^n = \mathbf{UVT}^{-1/2}\mathbf{V}^T$
    Update $\mathbf{B}^n = \mathbf{B}^{n-1} + \mathbf{R}^n - \mathbf{\Psi}^n$
    Update $\mathbf{P}^n = \mathbf{DR}^n$
    Compute loudspeaker array gain $\mathbf{a} = (\mathbf{P^H P})^{-1}\mathbf{P^H u}$
end for
Locate the loudspeaker position arranged on the sphere in accordance with the dictionary matrix $\mathbf{D}$.
Obtain the the speaker feeds from original sound $\mathbf{a}$ for rendering its spatial version.
```

III. Performance Evaluation

The performance of the proposed sparse iterative (SI) method is evaluated using sound field reconstruction analysis, subjective, and objective evaluations. The performance evaluation results are also compared to HOA, LS, and conventional compressive sensing methods.

A. Loudspeaker Array Configuration

For conducting sound field reconstruction experiments, an array setup of 130 loudspeakers is simulated. All these loudspeakers are placed uniformly over a sphere of radius 1m in an icosahedron pattern. Weights for loudspeakers are calculated using three techniques and it is observed that few entries of the weight vectors are zero. It signifies that actual number of speakers is less than the fully populated loudspeaker array. LS method uses only 123 loudspeakers to reproduce the sound field where as compressed sensing and proposed sparse iterative method uses 81 and 76 loudspeakers respectively in the experiments performed herein. The placement of the active speakers are therefore not uniform and are shown in figure 1a, 1b, and 1c respectively. In equation (14), ϵ, λ was fixed to 10^{-2} and 0.02 respectively.

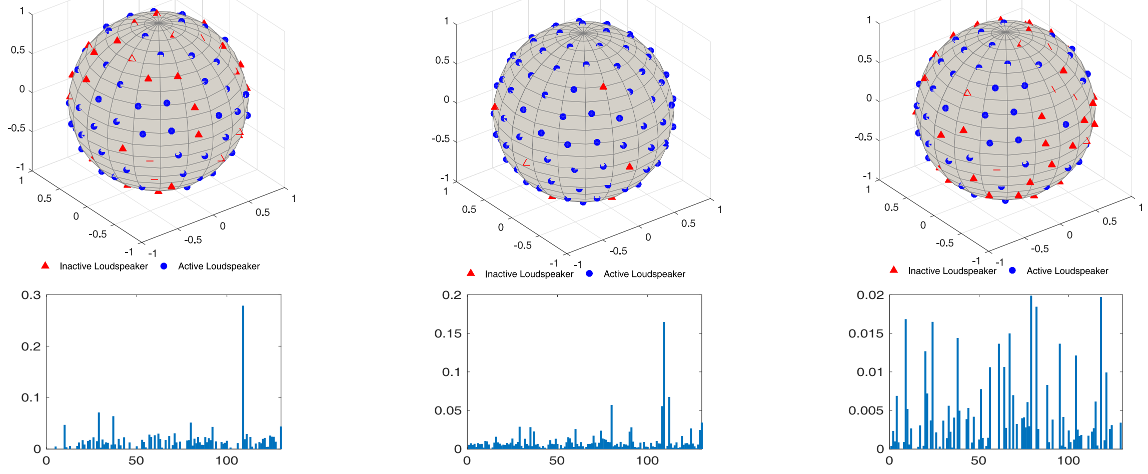
B. Experimental Results

1) Reconstructed Sound Field Analysis: In the experimental simulation, we consider a monochromatic plane wave of frequency 2kHz, incident from $[\theta, \phi] = [45°, 30°]$ with a reproduction sphere of radius $x_0 = 0.2m$. The loudspeaker weights were obtained by equation (7), (8) and (14) for LS, CS and sparse iterative (SI) method respectively. With these weights the sound pressures are obtained by using equation (3) and plotted in Fig.2 in an area $0.64m^2$ centered around the listening position.Figures 2 show density plots both for real and imaginary parts. Acoustic pressure less than -1 are showing black, greater than 1 are showing white, and pressure levels in between are appropriately shaded. In these figures the circle inside the plots represents the desired error free reconstruction area (sweet spot) where we compare the reproduction techniques. The reconstructed sound field by the proposed SI method is more closer to the reference sound in comparison to other methods. The normalization error for each technique is calculated as

$$e(k) = \frac{\int |S(\mathbf{x};k) - T(\mathbf{x};k)|^2 d\hat{\mathbf{x}}}{\int |S(\mathbf{x};k)|^2 d\hat{\mathbf{x}}} \tag{15}$$

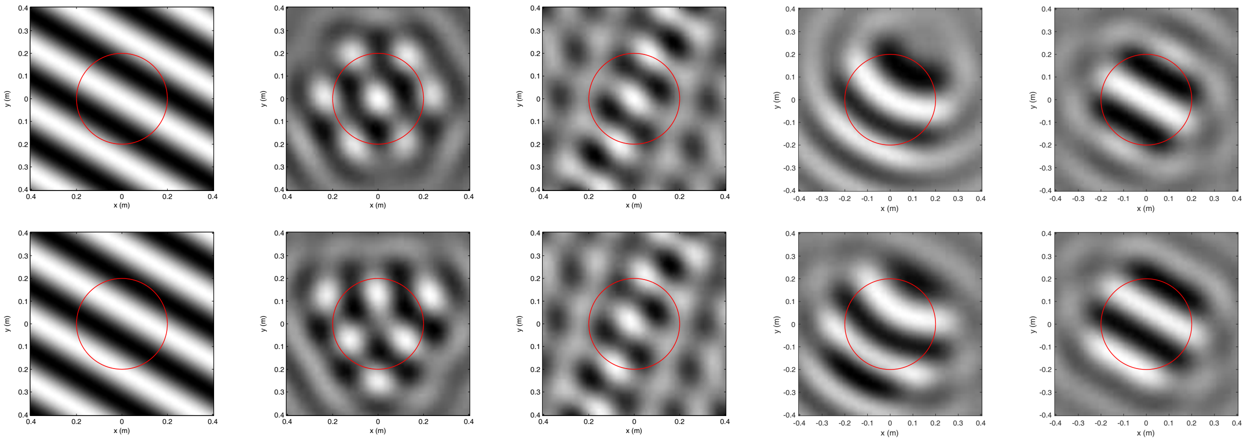
where the integration is taken over the surface of reproduction sphere. These error values are shown in Table-I for different radii of reproduction area around the listening spot for various sound field reconstruction methods.

2) Objective Evaluation: For objective evaluations, the loudspeaker gains are obtained by the different methods such as least square (7), compressed sensing (8), and sparse iterative (14). Perceptual Evaluation of Audio Quality (PEAQ) [21] and Perceptual Similarity Measure (PMS) [22] measures are used to quantify the spatial audio quality. These measures are provided in Table-II. From Table-II it can be concluded the spatial sound reproduced using proposed sparse iterative method is perceptually better than the other methods compared herein. The Average Error Distribution (AED) for the various sound field methods are also computed and plotted in Fig.3.



(a) CS
(b) LS
(c) Sparse Iterative

Fig. 1: The estimated loudspeaker positions illustrated on a sphere (top row) and their corresponding weights (bottom row) for CS, LS, and proposed sparse iterative methods respectively.



(a) Reference
(b) HOA
(c) CS
(d) LS
(e) Sparse Iterative

Fig. 2: Sound pressure plots for various methods (top and bottom row represents the imaginary and real part of reconstructed sound field respectively). The inner circle represents the area of perceptual interest ($x_0 = 0.2m$) for a frequency of 2kHz.

TABLE I: Decoding error obtained for sound field reconstruction by varying radii of spherical region of interest

x_0(in m) →	0.10		0.15		0.20		0.25		0.30	
Methods ↓	Real	Imag	Real	Imag	Real	Imag	Real	Imag	Real	Imag
HOA	0.1061	0.0819	0.2058	0.5036	0.2788	0.5671	0.2964	0.6357	0.3342	0.5783
CS	0.0488	0.6423	0.1271	0.4713	0.1984	0.4851	0.2067	0.5315	0.2490	0.4891
LS	0.0106	0.0129	0.0528	0.0520	0.1675	0.1252	0.2789	0.2186	0.3002	0.2869
SI	9.21E-04	4.71E-05	0.0091	0.0018	0.0645	0.0188	0.1422	0.0716	0.1660	0.1327

The error variance for the proposed sparse iterative method is less in comparison to other methods as shown in Fig.3.

3) Subjective Evaluation: The subjective evaluation is quantified as Mean Opinion Scores (MOS). The reproduced sound was perceived by ten subjects and these subjects were asked to rate the spatial attributes on a scale of 1 to 5, where very annoying was indicated by 1 and imperceptible by 5. The spatial attributes used are listed below [23]

- Naturalness: How true to life the audio listening was.
- Presence: Presence in audio source environment.
- Preference: Degree of pleasantness or harshness.
- Source Envelopment: Sound being all around a person.
- Perception of motion: The precision and correctness with which the trajectory of the source is perceived.

The scores obtained from the different subjects were averaged to obtain MOS and are listed in Table-III. A t-test is also performed to compare the audio attributes of spatial sound reproduced by different methods and shown in Fig.4.

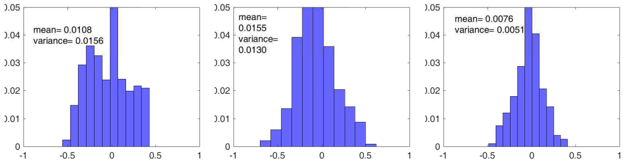


Fig. 3: Average Error Distributions for CS, LS, and SI methods respectively

TABLE II: PEAQ and PSM scores of reconstructed sound sources.

Method	PEAQ	DI	PSM	PSMt
LS	-3.0281	-1.2274	0.9954	0.939
CS	-3.1247	-1.3667	0.9945	0.9192
SI	-2.8460	-0.9946	0.9965	0.953

TABLE III: Subjective evaluation results for various sound field reproduction methods.

Method→	CS		LS		SI		Reference	
Attributes ↓	μ	σ^2	μ	σ^2	μ	σ^2	μ	σ^2
Naturalness	3.2	0.22	2.9	0.66	3.4	0.76	3.4	0.43
Presence	2.5	0.51	3.0	1.01	3.1	0.51	3.1	0.69
Preference	2.0	0.37	2.4	0.44	2.8	0.97	2.6	1.07
Source Envelopment	3.0	0.43	2.8	0.65	3.4	0.56	3.1	0.70
Perception of Motion	2.5	0.61	2.6	1.62	2.9	0.42	2.3	1.28

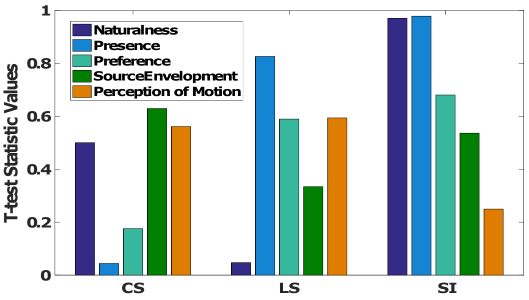


Fig. 4: T-test results for various sound field reproduction methods.

IV. Conclusion

In this work, a constrained optimization problem is formulated to determine the sparse panning vector to reconstruct the sound field accurately in spherical harmonic domain. The results are compared with conventional HOA techniques. Analysis of reconstructed sound field indicates that sparse iterative method performs better in term of spatial resolution and area of reconstruction. Extension of the proposed method to varechoic environment and improving the computational complexity will be investigated in future.

References

[1] Jérôme Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*, Ph.D. thesis, University of Paris VI, France, 2000.

[2] Augustinus J Berkhout, "A holographic approach to acoustic control," *Journal of the audio engineering society*, vol. 36, no. 12, pp. 977–995, 1988.

[3] Ville Pulkki, "Virtual sound source positioning using vector base amplitude panning," *Journal of the audio engineering society*, vol. 45, no. 6, pp. 456–466, 1997.

[4] Mark A. Poletti, "A unified theory of horizontal holographic sound systems," *J. Audio Eng. Soc*, vol. 48, no. 12, pp. 1155–1182, 2000.

[5] Andrew Wabnitz, Nicolas Epain, André van Schaik, and Craig Jin, "Time domain reconstruction of spatial sound fields using compressed sensing," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*. IEEE, 2011, pp. 465–468.

[6] Mark A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc*, vol. 53, no. 11, pp. 1004–1025, 2005.

[7] Efren Fernandez-Grande and Angeliki Xenaki, "Compressive sensing with a spherical microphone array," *The Journal of the Acoustical Society of America*, vol. 139, no. 2, pp. EL45–EL49, 2016.

[8] Franz Zotter, Hannes Pomberger, and Markus Noisternig, "Energy-preserving ambisonic decoding," *Acta Acustica united with Acustica*, vol. 98, no. 1, pp. 37–47, 2012.

[9] N Epain, CT Jin, and F Zotter, "Ambisonic decoding with constant angular spread," *Acta Acustica united with Acustica*, vol. 100, no. 5, pp. 928–936, 2014.

[10] Darren B Ward and Thushara D Abhayapala, "Reproduction of a plane-wave sound field using an array of loudspeakers," *IEEE Transactions on speech and audio processing*, vol. 9, no. 6, pp. 697–707, 2001.

[11] Franz Zotter, Matthias Frank, and Hannes Pomberger, "Comparison of energy-preserving and all-round ambisonic decoders," *Fortschritte der Akustik, AIA-DAGA,(Meran)*, 2013.

[12] Franz Zotter and Matthias Frank, "All-round ambisonic panning and decoding," *Journal of the audio engineering society*, vol. 60, no. 10, pp. 807–820, 2012.

[13] Ole Kirkeby and Philip A Nelson, "Reproduction of plane wave sound fields," *The Journal of the Acoustical Society of America*, vol. 94, no. 5, pp. 2992–3000, 1993.

[14] Y. J. Wu and T. D. Abhayapala, "Theory and design of soundfield reproduction using continuous loudspeaker concept," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 1, pp. 107–116, Jan 2009.

[15] Jérôme Daniel, "Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format," in *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*. Audio Engineering Society, 2003.

[16] Wen Zhang, Parasanga N. Samarasinghe, Hanchi Chen, and Thushara D. Abhayapala, "Surround by sound: A review of spatial audio recording and reproduction," *Applied Sciences*, vol. 7, no. 5, 2017.

[17] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, June 2010.

[18] X. Yong, R. K. Ward, and G. E. Birch, "Generalized morphological component analysis for eeg source separation and artifact removal," in *2009 4th International IEEE/EMBS Conference on Neural Engineering*, April 2009, pp. 343–346.

[19] Sachin N Kalkur, Sandeep Reddy C, and Rajesh M Hegde, "Joint source localization and separation in spherical harmonic domain using a sparsity based method," in *Sixteenth Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2015.

[20] Rongjie Lai and Stanley Osher, "A splitting method for orthogonality constrained problems," *Journal of Scientific Computing*, vol. 58, no. 2, pp. 431–449, 2014.

[21] Thilo Thiede, William C Treurniet, Roland Bitto, Christian Schmidmer, Thomas Sporer, John G Beerends, and Catherine Colomes, "Peaq-the itu standard for objective measurement of perceived audio quality," *Journal of the Audio Engineering Society*, vol. 48, no. 1/2, pp. 3–29, 2000.

[22] Rainer Huber and Birger Kollmeier, "Pemo-qa new method for objective audio quality assessment using a model of auditory perception," *IEEE Transactions on audio, speech, and language processing*, vol. 14, no. 6, pp. 1902–1911, 2006.

[23] C Sandeep Reddy and Rajesh M Hegde, "Horizontal plane hrtf interpolation using linear phase constraint for rendering spatial audio," in *Signal Processing Conference (EUSIPCO), 2016 24th European*. IEEE, 2016, pp. 1668–1672.