Acoust. Sci. & Tech. **41**, 1 (2020)

ACOUSTICAL LETTER

Spherical harmonic representation of rectangular domain sound fields

Junjie Shi, César D. Salvador, Jorge Treviño*, Shuichi Sakamoto† and Yôiti Suzuki‡

Research Institute of Electrical Communication and Graduate School of Information Sciences, Tohoku University, 2–1–1 Katahira, Aoba-ku, Sendai, 980–8577 Japan

(*Received 30 September 2018, Accepted for publication 9 August 2019*)

Keywords: Spatial audio, Sound field representation
PACS number: 43.60.Gk, 43.60.Tj [doi:10.1250/ast.41.451]

1. Introduction

The feeling of immersion is an essential subject in the virtual reality community, in which audio play a decisive role. However, it is difficult to provide a persuasive spatial audio. ADVISE (auditory display based on the virtual sphere model) [1] aims to present the spatial audio using binaural displays as an interface for listeners. It introduces a virtual sphere of secondary sources to reproduce the generated sound fields around the listener from computational results of the room acoustics. Binaural signals are derived by driving signals for the secondary sources and corresponding HRTFs. As for computation of room acoustics, an efficient and accurate method was proposed [2] based on decomposing a room in to rectangular parts. In this letter, we concentrate on finding the spherical harmonic representation of sound fields [3] in rectangular space with respect to sound wave has a simpler pattern than in an arbitrary space and the convenience of spherical harmonic representation in reproducing the sound field. We derived the representation by combining normal modes in rectangular rooms with rigid boundaries [4] and spherical harmonic expansion of the plane wave [3]. Numerical experiments were conducted to validate the proposed representation. While we focus on the representation of the sound field inside a single rectangular element, the Adaptive Rectangular Decomposition (ARD) can be used with our results to render the whole sound field inside a room [2].

2. Rectangular sound field representation

Consider a 3D rectangular domain with its diagonal extending from the $(0, 0, 0)$ to (l_x, l_y, l_z) and perfectly rigid, reflective walls. The pressure field $p(x, y, z)$ in the rectangular space can be represented as follows [4]

$$p(x, y, z) = \sum_{\eta=(\eta_x,\eta_y,\eta_z)} m_\eta \Phi_\eta(x, y, z). \tag{1}$$

This equation is a triple sum over the modes for each orthogonal axis with mode numbers η_x, η_y and η_z. For brevity, we use the symbol η as a shorthand for these three mode numbers (i.e. to index each of the 3D rectangular domain modes). m_η are, therefore, rectangular mode coefficients and Φ_η are normal modes (the eigenfunctions of the Laplacian) for a rectangular domain. Those modes are given by

$$\Phi_\eta(x, y, z) = \cos\left(\frac{\pi\eta_x}{l_x} x\right) \cos\left(\frac{\pi\eta_y}{l_y} y\right) \cos\left(\frac{\pi\eta_z}{l_z} z\right). \tag{2}$$

In the discrete interpretation, (1) is an inverse Discrete Cosine Transform (*iDCT*) in 3D, with Φ_η being the Cosine basis vectors for the given dimensions [2]. Therefore, we may efficiently transform from pressure values (P) to mode coefficients (M) as

$$P = iDCT(M), \quad M = DCT(P). \tag{3}$$

Thus, for any given sound field, the pressure distributions can be represented by a rectangular mode coefficients matrix.

3. A mapping from rectangular representation to spherical harmonic coefficients

As shown by Eq. (1), the desired field can be represented by rectangular mode coefficients. Each mode is a product of cosines. We apply the Euler formula to express each of the cosines in Eq. (2) as the sum of two complex exponentials. A complex exponential corresponds to a plane wave; therefore, each cosine is the superposition of two plane waves traveling in opposite directions. After carrying out the products and simplifying common terms, Φ_η becomes a summation of eight complex exponential terms or plane waves, written as

$$\Phi_\eta(x, y, z) = \frac{1}{8} \sum_{\ell=1}^{8} e^{i k_{\eta,\ell} \boldsymbol{x}}, \tag{4}$$

where $\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, $k_{\eta,\ell} = \begin{bmatrix} k_{\eta x} & k_{\eta y} & k_{\eta z} \\ k_{\eta x} & k_{\eta y} & -k_{\eta z} \\ k_{\eta x} & -k_{\eta y} & k_{\eta z} \\ \vdots & \vdots & \vdots \\ -k_{\eta x} & -k_{\eta y} & -k_{\eta z} \end{bmatrix}$, and

$k_{\eta x} = \dfrac{\pi\eta_x}{l_x}$, $k_{\eta y} = \dfrac{\pi\eta_y}{l_y}$, $k_{\eta z} = \dfrac{\pi\eta_z}{l_z}$.

In the sound field reproduction system, the listener position should be set at an arbitrary place in the room. To reproduce the sound field around the listener, we place a spherical coordinate taking the listener as the origin, as shown in Fig. 1; we denote the center of the spherical coordinates by O' and denote the vector from O to O' by $\boldsymbol{d}$. Equating (1) and (4), the desired field becomes:

*e-mail: jorge@ais.riec.tohoku.ac.jp
†e-mail: saka@ais.riec.tohoku.ac.jp
‡e-mail: yoh@riec.tohoku.ac.jp

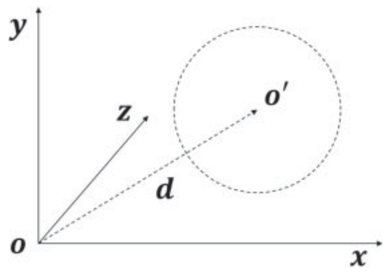

Fig. 1 Coordinate diagram: a spherical coordinate (shown by the dash circle) is located at position O', $\boldsymbol{d}$ denotes the displacement of two coordinates.

$$p(\boldsymbol{r},\omega) = \sum_{\eta=(\eta_x,\eta_y,\eta_z)} m_\eta(\omega) \cdot \frac{1}{8}\sum_{\ell=1}^{8} e^{ik_{\eta,\ell}(\boldsymbol{r}+\boldsymbol{d})}, \tag{5}$$

where $\boldsymbol{r}$ denotes an arbitrary position in the space.

The exponential terms in the above equation can be interpreted as incident plane wave from eight different directions. Note that these directions are taken with respect to O'.

Considering that the spherical harmonic expansion of the plane wave $e^{jk_{\eta,\ell}\boldsymbol{r}}$ was formulated in [3], we write the pressure distribution as follows:

$$\begin{aligned} p(\boldsymbol{r},\omega) = &\sum_{\eta=(\eta_x,\eta_y,\eta_z)} m_\eta(\omega) \cdot \frac{1}{8}\sum_{\ell=1}^{8} e^{ik_{\eta,\ell}\boldsymbol{d}} \\ &\cdot \sum_{n=0}^{\infty}\sum_{m=-n}^{n} 4\pi i^n j_n(kr) Y_n^{m*}(\hat{k}_{\eta,\ell}) Y_n^m(\hat{r}), \end{aligned} \tag{6}$$

where i denotes the imaginary unit, j_n denotes the sperical Bessel function and $k = \frac{\omega}{c}$ denotes the wavenumber.

Because $p(\boldsymbol{r},\omega)$ can also be expressed as $p(\boldsymbol{r},\omega) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} j_n(kr)P_n^m(\omega)Y_n^m(\hat{r})$. Harmonic coefficients P_n^m are finally formulated as:

$$P_n^m(\omega) = 4\pi i^n \sum_{\eta} m_\eta(\omega) \cdot \frac{1}{8}\sum_{\ell=1}^{8} e^{ik_{\eta,\ell}\boldsymbol{d}} Y_n^{m*}(\hat{k}_{\eta,\ell}). \tag{7}$$

This equation maps a rectangular representation m_η to a spherical representation P_n^m, it can be interpreted as the re-encoding of the field. In traditional approaches, P_n^m are obtained by inverse spherical harmonic transformation:

$$P_n^m(\omega) = \frac{1}{j_n(kr)}\int_0^{2\pi}\int_0^{\pi} p(r,\theta,\phi)Y_n^{m*}(\theta,\phi)\sin\theta d\theta d\phi. \tag{8}$$

However, when $j_n(kr) = 0$, P_n^m is undefined if the sound field is observed at a fixed radius. Equation (7) avoids this problem by observing sound field in the whole space.

4. Derivation of the driving signals

Among numerous approaches to calculate driving signals for sound field reproduction systems [5,6], we simply apply mode-matching approach to encoded sound field (refer [5] for further details). In an auditory display based on the virtual sphere model, these driving signals are combined with appropriate HRTFs and summed to produce a binaural signal.

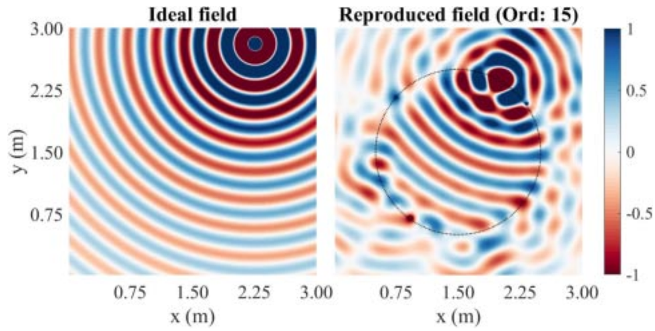

Fig. 2 Ideal field of 1,000 Hz monopole source and reproduced field with 256 radius 1 m of secondary monopole sources.

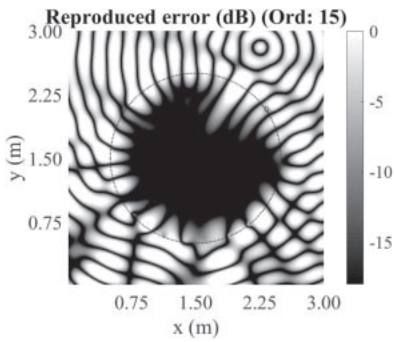

Fig. 3 Reproduced Error: $20\log_{10}(|p_{\mathrm{reproduced}} - p_{\mathrm{ideal}}| \cdot d_{\mathrm{norm}})$.

5. Numerical experiment

A 3D sound field reproduction experiment is conducted to evaluate our proposal. We assume a $3\,\mathrm{m} \times 3\,\mathrm{m} \times 3\,\mathrm{m}$ space in free field in which a 1,000 Hz monopole source located at $(1.5, 0^\circ, 60^\circ)$ (corresponding to $(2.25, 1.5(1+\sqrt{3}), 1.5)$ in the Cartisian Coordinate), the space is discretized and sampled every 0.1 m.

Figure 2 shows the pressure distribution of ideal field and reproduced field on $z = 1.5\,\mathrm{m}$ plane. The expansion order is $n = 15$. Secondary sources (256 in total) are located on a 1 m radius sphere, indicated by the dash circle. The distribution of secondary sources corresponds to a minimum in the energy for a repulsive Coulomb potential [7]. The normalized reproduced error is depicted in Fig. 3, which is calculated by $20\log_{10}(|p_{\mathrm{reproduced}} - p_{\mathrm{ideal}}| \cdot d_{\mathrm{norm}})$, where d_{norm} is the distance between sample points and virtual source. The result shows that the reproduced error is bounded below -36 dB when the distance is less than 10 cm from the expansion center and -25 dB when it is less than 20 cm. These error levels, being comparable to what is observed in state-of-the-art sound field reproduction systems based on spherical harmonic representations, show that the main source of error in the proposal is the use of a finite expansion order. Our results show that reproduction error is as small as -36 dB within a volume comparable to human head size. The proposal may thus be useful as an architecture for future virtual auditory displays.

Acknowledgment

This work was supported by the JSPS Grant-in-Aid for Scientific Research (KAKENHI) No. JP16H01736.

References

[1] S. Takane, Y. Suzuki, T. Miyajima and T. Sone, "A new theory for high definition virtual acoustic display named ADVISE," *Acoust. Sci. & Tech.*, **24**, 276–283 (2003).

[2] N. Raghuvanshi, R. Narain and M. C. Lin, "Efficient and accurate sound propagation using adaptive rectangular decomposition," *IEEE Trans. Vis. Comput. Graph.*, **15**, 789–801 (2009).

[3] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography* (Academic Press, San Diego, 1999).

[4] H. Kuttruff, *Room Acoustics* (CRC Press, Boca Raton, 2016).

[5] M. A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," *J. Audio Eng. Soc.*, **53**, 1004–1025 (2005).

[6] S. Spors, R. Rabenstein and J. Ahrens, "The theory of wave field synthesis revisited," *124th AES Conv.* (2008).

[7] E. B. Saff and A. B. J. Kuijlaars, "Distributing many points on a sphere," *Math. Intell.*, **19**, 5–11 (1997).