

A multi-zone approach to sound field reproduction based on spherical harmonic analysis

Taku Shimizu*, Jorge Trevino†, Shuichi Sakamoto‡ and Yôiti Suzuki§

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

(Received 12 February 2015, Accepted for publication 20 March 2015)

Keywords: Sound field reproduction, Spatial sound, Multi-zone sound reproduction, Spherical harmonic analysis

PACS number: 43.60.Pt, 43.60.Sx [doi:10.1250/ast.36.441]

1. Introduction

Rapid advances in information and communication technologies have led to increased expectations for modern presentation systems. Ultra-realistic sound reproduction plays a key role towards fulfilling these demands. Several types of reproduction systems have been advanced to present spatial sound. Among them, sound field reproduction methods stand out for achieving high levels of realism by re-creating the relevant physical variables, i.e., the sound pressure field around the listener. Some promising methods in this category are: wave field synthesis (WFS) [1] formulated in terms of the Rayleigh integral, boundary surface control (BoSC) [2] based on the Kirchhoff-Helmholtz integral equation, high-order Ambisonics (HOA) [3] which applies harmonic analysis to encode spatial sound information, and BoSC's binaural implementation, known as acoustic display based on the virtual sphere model (ADVISE) [4]. Existing sound field reproduction technologies can be classified into two groups: those which re-create the sound pressure inside a bounded volume [2,3] and those which focus on an unbounded spatial region [1]. Methods falling in the first group produce a finite listening zone, which may make them unfit for multi-user sound presentation. Systems in the second group can present sound to large audiences, but they also place impractical demands on the reproduction system as they attempt to cover a large space, even regions where no listener is present [1]. The present research letter focuses on the task of presenting spatial sound to multiple users, dispersed over a large region, while making optimal use of the capabilities of the reproduction system. Some research efforts have attempted to deal with the problem of multi-user sound field reproduction. Two existing methods and their weaknesses are reviewed in Sect. 2. A new spatial sound presentation method based on the harmonic analysis of sound fields is introduced in Sect. 3. The new method was designed to avoid the limitations of existing methods. Its performance is evaluated through a numerical simulation in Sect. 4.

2. Existing multi-user reproduction methods

An optimal spatial sound reproduction system for multiple

users should focus on re-creating the sound field around each of the listeners while ignoring the empty space between them. Multi-zone sound field reproduction systems of this kind have been proposed in the past [5,6]. In this section, two of these methods are reviewed.

2.1. Additive approach to multi-zone reproduction

This method, introduced in [5], achieves multi-zone reproduction by superposing the individual fields for each listening zone. This method seeks to re-create the sound field for N listening zones, all of them located inside a loudspeaker array. The first step in this method is to calculate the loudspeaker signals required to produce the target sound field inside one zone (the active zone) while trying to keep the remaining $N - 1$ zones silent (silent zones). This is repeated for each of the listening zones and the results are added to get the final reproduction signals. The loudspeaker signals are calculated by regulating the sound pressure at a large number of control points inside the target zones. A stable implementation requires the number of loudspeakers to match or exceed the number of control points. The target sound for each listening zone is set independently; therefore, different contents can be presented to each user. However, this approach may result in unnatural sound fields that are difficult to reproduce. An example of this occurs when a silent zone is located in the proximity of the active zone. Keeping the sound pressure close to zero in the silent zone imposes a rapid change condition in the resulting sound field. Moreover, this method, as described in [5], can only reproduce sounds in the horizontal plane. There is no fundamental reason for this limitation; however, the number of control points needed to sample the listening zones in 3D space would call for an impractical number of loudspeakers.

2.2. Global field approach to multi-zone reproduction

The second multi-zone sound field reproduction method to be reviewed was proposed in [6]. This method is unique in that it uses a common spatial sound encoding, known as high-order Ambisonics [3], for its inputs. The HOA encoding is based on the spherical harmonic decomposition, which states that the sound pressure field p due to a source at azimuth θ and elevation φ can be expressed as the following infinite series:

$$p(\theta, \varphi, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^n B_n^m(k) Y_n^m(\theta, \varphi), \quad (1)$$

*e-mail: shimizu@ais.riec.tohoku.ac.jp

†e-mail: jorge@ais.riec.tohoku.ac.jp

‡e-mail: saka@ais.riec.tohoku.ac.jp

§e-mail: yoh@riec.tohoku.ac.jp

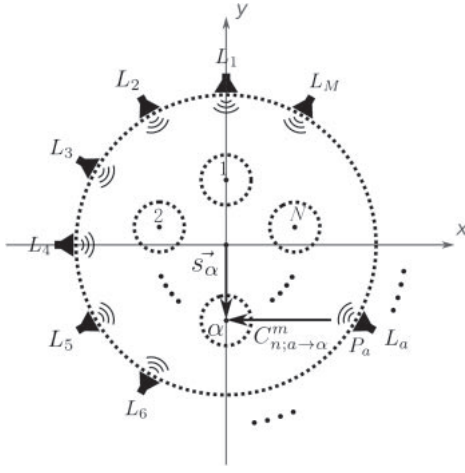


Fig. 1 Geometry of multi-zone reproduction system.

where k denotes the wavenumber and Y_n^m stands for the spherical harmonic function of order n and degree m . The set of coefficients B_n^m fully characterizes the sound pressure field p . The summation in Eq. (1) is typically truncated to a maximum harmonic order N in practical implementations.

The approach introduced in [6] works by constructing a global sound field which can cover all of the listening zones. To achieve this, the harmonic expansion coefficients characterizing the target sound field for each listener are combined. The listening zone positions are taken into account by applying the theorem of spherical harmonic coefficient translation. This theorem, however, can only be applied to the terms where $|m| = n$ in Eq. (1); the rest of the coefficients must be discarded. For this reason, the global field approach is limited to the horizontal plane.

Furthermore, this approach may result in an ill-posed problem when two listening zones are located with a close angular separation, as seen from the center of the reproduction system [6]. Once a global sound field has been derived, this method attempts to reproduce it entirely. This re-creates the target sound fields around each of the listening zones; however, it is not optimal since the empty space between the zones is also covered by the global field. Reproducing an extended field like this may call for an impractical number of loudspeakers, particularly if the distance between listening zones is large.

3. A new approach to multi-zone sound field reproduction

We now propose a new algorithm to reproduce sound fields in multiple listening zones based on spherical harmonic analysis methods. The system geometry assumed in the formulation of our proposal is shown in Fig. 1. The methods of Sect. 2 define separate target sound fields for each listening region. In this research, however, we focus on the problem of presenting a common sound field to all users. Our proposal attempts to overcome the limitations of existing methods. In particular, our goal is to achieve stable reproduction accuracy regardless of the specific positions of the listening zones. The proposed method allows for three-dimensional sound

field reproduction without unrealistic system requirements. Our proposal starts with the spherical harmonic expansion introduced in Eq. (1). We will consider all of the terms in the expansion up to a maximum order. Our proposal consists of two steps: a phase compensation stage and a multi-zone transfer function inversion stage.

3.1. Phase compensation stage

The spherical harmonic expansion encodes the sound field around a single, privileged position. Multi-zone reproduction, however, seeks to re-create the sound pressure at different locations. The sound field, as observed from each of the listening zones, differs in phase from the encoded target field. To account for this, we propose the following phase compensation stage:

$$B_{n;\alpha}^m(k) = e^{-i\vec{k} \cdot \vec{s}_\alpha} B_n^m(k). \quad (2)$$

Here, $B_{n;\alpha}^m$ denotes the set of spherical harmonic coefficients for listening zone α , located at \vec{s}_α . \vec{k} stands for the wavevector of the encoded sound field. Our proposal uses these phase-corrected expansion coefficients to ensure that the sound field covering all of the listening regions is consistent. This prevents abrupt changes in the sound pressure when multiple zones are close to each other, avoiding the position-dependent accuracy problems reported for existing methods [5,6].

3.2. Multi-zone transfer function

The first step is to calculate the expansion coefficients for each zone from the HOA encoding of the target field. This is done using a phase compensation stage. The proposed method attempts to re-create the sound pressure inside multiple listening zones efficiently. The exact positions of the listening zones and their separations should not impact reproduction accuracy in a significant way. To achieve this, all of the transfer functions between each loudspeaker and each of the listening zones are considered simultaneously. The multi-zone transfer function for loudspeaker a , expressed by its harmonic components $C_{n;a}^m(k)$, is given as follows:

$$C_{n;a}^m(k) \rightarrow \frac{1}{N} \sum_{\alpha=1}^N D_{n;a \rightarrow \alpha}^m(k) \quad (3)$$

Here, coefficients $D_{n;a \rightarrow \alpha}^m(k)$ represent the harmonic components for the transfer function between the a -th loudspeaker and the α -th listening region.

3.3. Multi-zone reproduction algorithm

Conventional, single-zone HOA reproduction works by inverting a linear system derived from the transfer functions relating the loudspeaker input signals P_a and the sound pressure around the listener. The equation for loudspeaker a , harmonic order n and degree m is

$$\sum_{a=1}^M C_{n;a}^m(k) P_a = B_n^m(k). \quad (4)$$

The transfer functions are, thus, characterized by the coefficients $C_{n;a}^m(k)$. The second step introduced in our proposed method consists of extending Eq. (4) to include a set of multi-zone transfer functions relating all loudspeakers to all listening zones.

Our proposal puts together the phase compensation stage of Eq. (2) and the multi-zone transfer functions of Eq. (3).

This leads to the following multi-zone sound field re-encoding equation:

$$\sum_{a=1}^M \left[\frac{1}{N} \sum_{\alpha=1}^N e^{i\vec{k} \cdot \vec{s}_a} D_{n;a \rightarrow \alpha}^m(k) \right] P_a = B_n^m(k) \quad (5)$$

This equation defines a linear system relating the HOA encoding of the target sound field and the loudspeaker signals. By inverting this linear system, our proposed method derives the optimal loudspeaker signals to reproduce the target field at all of the listening zones. The proposal described above attempts to re-create the target sound field only in the neighborhoods of the listening zones, while ignoring it in the empty space between them. This allows for a more optimal use of the available loudspeakers when compared with the global sound field approach [6]. The calculation of the loudspeaker signals simultaneously considers all of the listening zones. Furthermore, the phase compensation stage guarantees a consistent sound field around all listeners. This avoids the need to reproduce unnatural sound fields exhibiting rapid changes in magnitude, as required by other approaches [5]. Finally, all steps in the proposed method are applicable to 3D sound reproduction with no more system demands than conventional HOA reproduction.

4. Evaluation

To evaluate the performance of our proposed method, a series of computer simulations were conducted. Our evaluation considers the case of two-zone reproduction of 1 kHz plane waves. The spherical harmonic expansion was limited to a maximum order of 3. The reproduction system consists of 192 loudspeakers regularly distributed on the surface of a sphere. The error in the reproduced field is measured using the following formula:

$$\text{Error} = 10 \log_{10}(p_r - p_i)^2 \quad [\text{dB}]. \quad (6)$$

Here, p_r and p_i represent the reproduced and ideal sound pressure fields, respectively. The use of absolute errors is justified since our simulations normalize the amplitude of the original sound field to unity. Our evaluation considers a reproduction error lower than -6 dB to be acceptable, and defines the regions where this condition is met as the listening zones.

Our evaluation starts by confirming that the proposed method is capable of reproducing the sound field over multiple listening zones. For this, a plane wave incident from the left of the listener in the horizontal plane is considered. We define two listening zones, positioned 0.5 m to the front and back of the center of the array, as observed by the listener. The radius of the reproduction array was set to 2.5 m. Figure 2 shows our simulation results. Here, the listener's head is oriented towards the positive x -axis. The reproduced sound field is shown in panel (a). Panel (b) shows the magnitude of the reproduction error. In this figure, simulation results show two distinct listening zones, defined by the -6 dB error condition. The zones are centered exactly at the desired positions. Both of them have a radius of approximately 17 cm.

Our second simulation considers the impact of inter-zone separation on the size of the listening regions. The conditions are the same as in the previous simulation; however, the

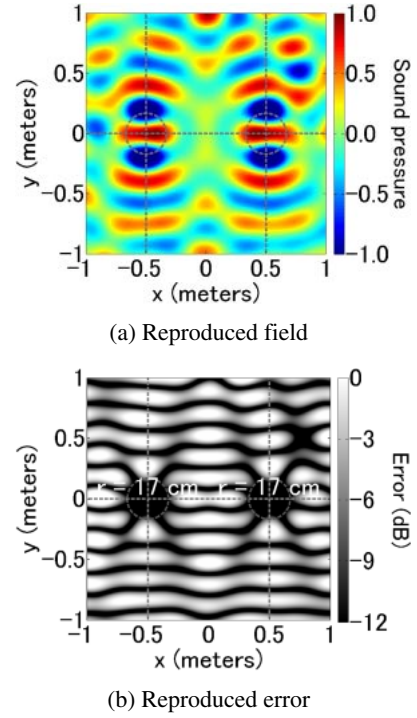


Fig. 2 Results obtained by applying the proposed method to the two-zone reproduction of a 1 kHz plane wave incident from the right.

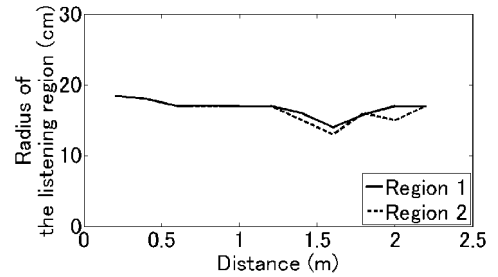


Fig. 3 Size of two listening zones in relation to their separation. The target sound field is a 1 kHz plane wave encoded up to order 3. The radius of reproduction array is 2.5 m.

distance between the listening zones and the center of the reproduction system is changed as a parameter. Results for multiple inter-zone separations are shown in Fig. 3. Our results show that the distance between listening zones has no significant effects on the size of the listening zones. This is an improvement over existing methods where accuracy drops if the listening regions are too close (see Fig. 20 in [5]) to each other. To extend the previous simulation to larger inter-zone distances, the radius of the reproduction array is set to 25 m. Figure 4 shows the size of the two listening zones when they are separated by up to 24 m. Despite the larger separation, the distance between listening zones has no significant effects on the reproduction accuracy. This is an improvement over existing methods where accuracy drops if the listening regions are too far apart (see Fig. 9 in [6]).

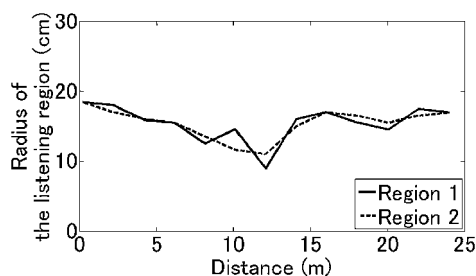


Fig. 4 Size of two listening zones in relation to their separation. The target sound field is a 1 kHz plane wave encoded up to order 3. The radius of reproduction array is 25 m (ten times larger than the one used in Fig. 3).

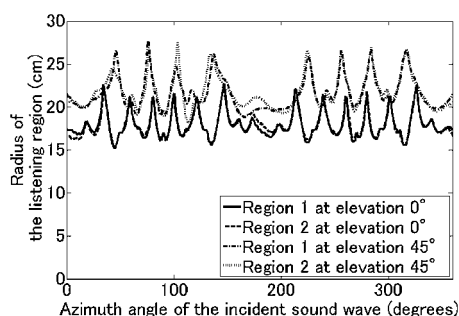


Fig. 5 Radius of two reproduction regions for a 1 kHz plane wave incident from different directions.

Finally, we investigated the effects of sound source position on reproduction accuracy. Figure 5 shows the size of listening zones for plane waves incident from different directions. Our results show that the size of the reproduction zones does not significantly change depending on the direction of the target plane wave. This is an advantage of our proposed method over previous ones (see Figs. 13, 14, and 16 in [5]), in which reproduction accuracy drops if two listening zones lie in the direction of the wavefront.

The proposed method is not limited to the horizontal plane, unlike previous ones [5,6]. This was verified this by simulating the reproduction of a plane wave incident from an elevation of 45°. These results are shown in Fig. 5. This figure

shows the radius of the listening zones measured in the horizontal plane. Elevated sound sources exhibit larger radii since projection onto this plane reduce their apparent wave-vector by a factor of $\cos \varphi$. A detailed analysis of the listening zones' size in 3D space is left for future evaluations.

5. Conclusion

A new method for multi-zone sound field reproduction was introduced and validated through a series of computer simulations. Our proposed method was found to be capable of accurate reproduction inside two separate listening zones using a surrounding loudspeaker array. Furthermore, our method was not found to be affected by the separation between the listening zones or their alignment with the sound sources present.

Acknowledgments

The authors wish to thank Dr. Tomohiko Ise of Alpine Electronics for his assistance and helpful advice, as well as the anonymous reviewer for insightful comments regarding our proposal's evaluation. Parts of this research were supported by a Grant-in-Aid for Scientific Research (A) (No. 24240016) to YS from the Ministry of Education, Culture, Sports, Science and Technology, and the A3 Foresight Program for "Ultra-realistic acoustic interactive communication on next-generation Internet."

References

- [1] A. J. Berkhout, "A holographic approach to acoustic control," *J. Audio Eng. Soc.*, **36**, 977–995 (1988).
- [2] S. Ise, "A principle of sound field control based on the Kirchhoff-Helmholtz integral equation and the theory of inverse systems," *Acustica*, **85**, 78–87 (1999).
- [3] M. A. Poletti, "Three-dimensional surround systems based on spherical harmonics," *J. Audio Eng. Soc.*, **53**, 1004–1025 (2005).
- [4] 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).
- [5] M. Poletti, "An investigation of 2D multizone surround sound systems," *Proc. 125th Audio Eng. Soc. Conv.*, 7551, pp. 1–9 (2008).
- [6] J. W. Yan and D. A. Thushara, "Spatial multizone soundfield reproduction: theory and design," *IEEE Trans. Audio Speech Lang. Process.*, **19**, 1711–1720 (2011).