社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

Editing distance information in compact microphone array recordings for its binaural rendering

César SALVADOR†, Shuichi SAKAMOTO†, Jorge TREVIÑO†, and Yôiti SUZUKI†

† Research Institute of Electrical Communication and Graduate School of Information Sciences, Tohoku University, Sendai, Japan

E-mail: {salvador,saka,jorge}@ais.riec.tohoku.ac.jp, yoh@riec.tohoku.ac.jp

Abstract Spatial sound information can be recorded using spherical microphone arrays. Realistic binaural renderings of such recordings require the head-related transfer functions (HRTFs), characterized from thousands of surrounding positions to both ears. We propose a 3D audio processing method to make sounds appear closer or farther than their original distance. HRTFs for only some surrounding positions at a fixed distance are required. Binaural signals are synthesized from samplings of the sound pressure field on a rigid spherical surface. The directional structure of both the sound pressure field and the HRTFs are encoded using spherical harmonics. Distance manipulations rely on the use of the spherical Hankel functions.

Key words Sound source distance, sound field recording, binaural reproduction.

1. Introduction

Spherical microphone arrays provide a suitable framework for analyzing the directional structure of sound fields in terms of spherical harmonics functions [1,2]. The microphones are typically mounted on the surface of a rigid spherical baffle. The use of such baffle avoids the dramatic decay of energy for specific harmonic modes at certain frequencies dictated by the array's radius. Compact spherical microphone arrays are therefore useful to obtain stable encodings of sound fields in terms of spherical harmonics.

The distance information of a sound source can also be included in spherical harmonic encodings [3, 4]. It requires the use of spherical Hankel functions, which model the free-field decay of sound intensity along increasing distance. The sound pressure distributions at different distances from an observation point are related by a ratio of spherical Hankel functions [3], which are referred to as acoustic propagators. Spherical harmonic encodings can therefore be enhanced using acoustic propagators so as to make sounds appear closer or farther than their original distance.

On the other hand, the binaural rendering of a sound source aims to re-create the natural listening experience by synthesizing the sound pressure at the listener's ears [5]. These binaural signals contain auditory localization cues to perceive directions and also distances within 1 m [6–10]. They arise from the scattering of sound by the listeners pinna, head and torso. The scattering phenomena are characterized by the so-called Head-Related Transfer Functions (HRTFs). They are typically measured with two microphones placed in the ear canal the listener's ears, and using a surrounding array of loudspeakers to emit sounds in the audible frequency range from 20 Hz to 20 kHz. In conventional binaural systems, an audio signal is filtered with the HRTFs for both ears, and the resulting binaural signals are reproduced over headphones.

Binaural cues can be used to determine directions and distances of nearby lateral sound sources. However, available sets of HRTFs are typically characterized for surrounding sound sources at a fixed distance beyond 1 m [11, 12]. The simplest approximation to synthesize HRTFs for sound sources near the listener's head from available sets of HRTFs uses a head-sized sphere to model distance variations [13]. Better approximations rely on the application of acoustic propagators to the spherical harmonics representation of available HRTFs [14–16].

We introduce in this paper a 3D audio processing method to make sounds appear closer or farther than their original distance during its binaural rendering. Our proposal combines the synthesis of sound fields and HRTFs based on the spherical harmonics decomposition. Section 2 overviews our proposal. Section 3 briefly introduces the propagation of sound. Section 4 describes the edition of distance information on microphone array recordings. Section 5 describes the synthesis of HRTFs at arbitrary positions. Section 6 evaluates our proposal in a practical scenario, where microphones and HRTF's sound sources are placed on spherical samplings. Concluding remarks are finally presented in Section 7.

2. Overview

We propose a 3D audio processing method to edit the distance information of an encoded sound source for its binaural rendering. Our proposal combines the synthesis of sound fields and HRTFs based on the spherical harmonics decomposition, for nearby sound

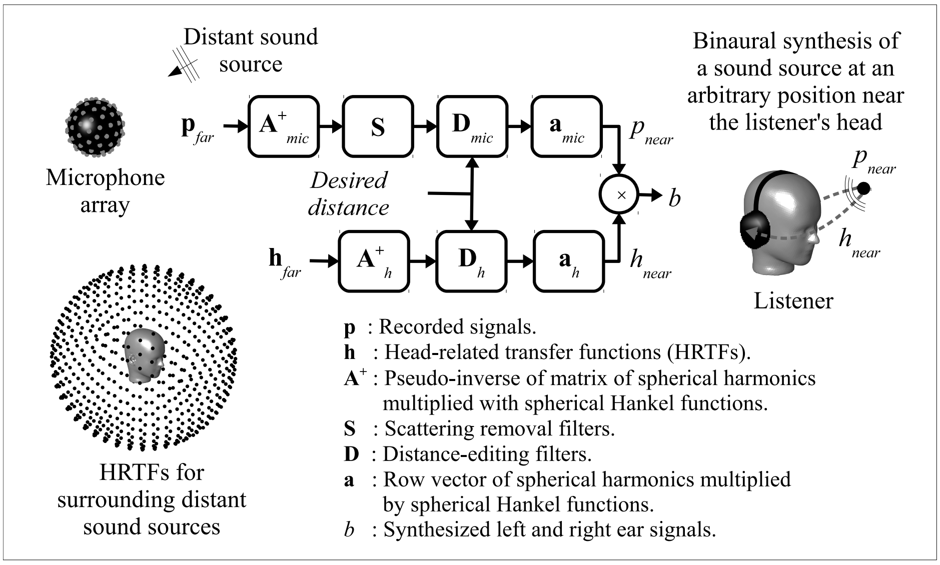

Fig. 1 *Overview of the binaural synthesis method.*

sources at arbitrary positions (see Fig. 1). Our method requires sampling the sound pressure field on a rigid spherical surface, and the HRTFs for only some surrounding positions at a fixed distance. To synthesize the binaural signals at the directions that have not been measured, we rely on the use of spherical harmonic encodings of recorded signals and HRTFs. To edit the distance information on the spherical harmonics representations, we apply acoustic propagators based on spherical Hankel functions.

The spherical coordinate system used in this document to describe our proposal is shown in Fig. 2. A point in space (θ, ϕ, r) is specified by its azimuth $\theta \in [-180°, 180°]$, its elevation $\phi \in [-90°, 90°]$, and its distance r. The listener's ears lie on the horizontal plane at an elevation $\phi = 0°$. The front direction corresponds to azimuth $\theta = 0°$ and elevation $\phi = 0°$. The array of microphones and the surrounding array of sound sources used to characterize the HRTFs are centered at the origin $(0, 0, 0)$.

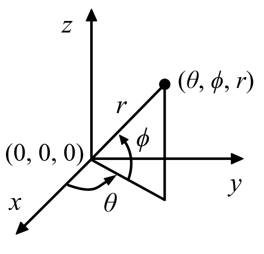

Fig. 2 *Spherical coordinate system used in the formulation of our proposal.*

3. Propagation of sound

The solutions of the acoustic wave equation for sound pressure in spherical coordinates can be expressed as an expansion of the acoustic pressure [3]

$$\psi(\theta, \phi, r, k) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_{nm}(\theta, \phi)\psi_{nm}(r, k), \qquad (1)$$

at position (θ, ϕ, r) and wave number k. Here,

$$Y_{nm}(\theta, \phi) = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_{nm}(\sin\phi)e^{im\theta} \qquad (2)$$

denotes the complex spherical harmonic function of order n and degree m, where P_{nm} are the associated Legendre functions. The coefficients

$$\psi_{nm}(r, k) = c_{nm}(k)h_n(kr), \qquad (3)$$

are referred to as the spherical wave spectrum, where h_n is the spherical Hankel function of order n, and c_{nm} denotes the spherical expansion coefficients.

For small kr, that is, for low frequencies or small distances from the origin, the solutions in Eq. (1) can be approximated with small errors by truncating the sum along n to a finite value N. By using the asymptotic expansion of spherical Hankel functions for large orders, it has been shown in [17] that small and bounded truncation errors are obtained when maximum orders are limited according to

$$N = \lceil \frac{ekr}{2} \rceil, \qquad (4)$$

where e is the base of the natural logarithm. Conversely, given a maximum order N, acoustic pressures can be represented up to a maximum frequency f_{max} given by

$$f_{max} = \frac{N\nu}{e\pi r}, \qquad (5)$$

where ν is the velocity of sound in air.

For a given wave number k, the acoustic pressure on a spherical surface is entirely determined by its spherical expansion coefficients, as can be noted from Eqs. (1) and (3). Such continuous set of pressures form an acoustic field. The spherical wave spectra of the acoustic field measured on two concentric spherical surfaces of different radii r_1 and r_2 are therefore related by [3]

$$\psi_{nm}(r_2,k) = \frac{h_n(kr_2)}{h_n(kr_1)}\psi_{nm}(r_1,k). \tag{6}$$

The ratio of spherical Hankel functions are the acoustic propagators from radius r_1 to radius r_2. The synthesis of the acoustic field at an arbitrary distance is thus performed in two processing steps: (a) the spherical wave spectrum is determined by expanding the acoustic field on a sphere of radius r_1, and (b) the acoustic pressure field at a desired radius r_2 is calculated by acoustic propagation from radius r_1 applying the corresponding acoustic propagators.

We describe the application of this approach to the edition of distance information in compact microphone array recordings in Section 4, and to the edition of distance information in available HRTFs in Section 5. Available sets of HRTFs are typically measured at far distances. Binaural cues are used by listeners to determine the range of sound sources nearby the listener's head. We therefore limit our description in this document to the acoustic propagation from far distances $r_1 = r_{far}$ to nearby distances $r_2 = r_{near}$.

4. Editing distance in sound field recordings

Using a system of equations like Eq. (1), the vector of pressures $\mathbf{p}_{far}$ at positions $(\theta_{mic}, \phi_{mic}, r_{mic})$ of M microphones, due to a sound source at a position $(\theta_{far}, \phi_{far}, r_{far})$ faraway from the microphones, can be written

$$\mathbf{p}_{far} = \mathbf{A}_{mic}\mathbf{c}_{mic}. \tag{7}$$

The column vector $\mathbf{c}_{mic}$ contains the $(N_{mic}+1)^2$ spherical expansion coefficients to be determined. The $M \times (N_{mic}+1)^2$ matrix $\mathbf{A}_{mic}$ contains the spherical harmonics evaluated on $(\theta_{mic}, \phi_{mic})$ multiplied by the corresponding spherical Hankel functions evaluated at r_{mic}, both limited to a maximum order

$$N_{mic} = \min(\lceil \frac{ekr_{mic}}{2} \rceil, \lfloor \sqrt{M} - 1 \rfloor). \tag{8}$$

The vector of coefficients $\mathbf{c}_{mic}$ in Eq. (7) can be obtained by applying a least-squares method with Tikhonov regularization to compute the pseudo-inverse matrix of $\mathbf{A}_{mic}$. Regularization yields approximated solutions that are less sensitive to data perturbations than the ones obtained with the least-squares method. A particular solution is emphasized by choosing properly the regularization matrix $\mathbf{R}$ that minimizes the residual $\|\mathbf{A}_{mic}\mathbf{c}_{mic} - \mathbf{p}_{far}\|^2 + \|\mathbf{R}\mathbf{c}_{mic}\|^2$. The pseudo-inverse matrix of $\mathbf{A}_{mic}$ reads [18]

$$\mathbf{A}^{+}_{mic} = \left(\mathbf{A}^{H}_{mic}\mathbf{W}_{mic}\mathbf{A}_{mic} + \epsilon_{mic}\mathbf{R}_{mic}\right)^{-1} \mathbf{A}^{H}_{mic}\mathbf{W}_{mic}, \tag{9}$$

where $\mathbf{W}_{mic}$ is the $M \times M$ diagonal matrix of weighting coefficients, $\mathbf{R}_{mic}$ is the $(N_{mic}+1)^2 \times (N_{mic}+1)^2$ regularization matrix, and ϵ_{mic} is the regularization parameter.

Weighting coefficients are applied to the recorded signals according to the area covered by the corresponding microphones. They are not part of the regularization itself, but are analogous to the integration quadratures used on a spherical surface. We used weighting coefficients that are proportional to the area of each microphone's neighborhood. The neighborhood of a microphone is defined as all points on the sphere that are closer to it than to other microphones.

The regularization matrix is typically chosen as the identity matrix, giving preference to solutions with smaller norms. We used, though, a regularization matrix that depends on the decomposition order n. It is given by the diagonal matrix [14]

$$\begin{aligned} &\mathbf{R}_{mic} = diag\left(\rho_1,\ \ldots, \rho_{(N_{mic}+1)^2}\right), \\ &\rho_{n^2+n+m+1} = 1 + n(n+1), \end{aligned} \tag{10}$$

where n and m are respectively the order and degree of the spherical harmonics of the columns of matrix $\mathbf{A}_{mic}$. In this way, high-degree harmonics are damped more than low-degree ones, which was seen to reduce truncation errors [14]. In order to get a good compromise between the error and the energy of the source, we set the regularization parameter $\epsilon_{mic} = \|\mathbf{A}_{mic}\| \times 10^{-8}$, with the norm of $\mathbf{A}_{mic}$ equal to the largest singular value [19].

In particular, the solution to Eq. (7) when using a rigid spherical baffle reads [20]

$$\mathbf{c}_{mic} = \mathbf{S}\mathbf{A}^{+}_{mic}\mathbf{p}_{far}, \tag{11}$$

where $\mathbf{S}$ is the $(N_{mic}+1)^2 \times (N_{mic}+1)^2$ diagonal matrix whose entries are filters to compensate for the presence of the baffle, and $\mathbf{A}^{+}_{mic}$ is given by Eq. (9). To remove the scattering effects introduced by the baffle of radius r_{mic}, we used the filters proposed in [1]. These filters were derived assuming that the total radial velocity equals zero at the surface of the solid sphere. The scattering removing matrix is the diagonal matrix

$$\begin{aligned} &\mathbf{S} = diag\left(\sigma_1,\ \ldots, \sigma_{(N_{mic}+1)^2}\right), \\ &\sigma_{n^2+n+m+1} = -kr^2_{mic}\frac{h'_n(kr_{mic})}{h_n(kr_{far})}, \end{aligned} \tag{12}$$

where n and m index to the columns of $\mathbf{A}_{mic}$ like in Eq. (10), and h'_n denotes the derivative of the spherical Hankel function of order n.

The distance information contained in spherical expansion coefficients $\mathbf{c}_{mic}$ can now be edited based on Eq. (6). The sound pressure of a synthesized sound source near the compact microphone array is subsequently expressed in a way similar to Eq. (7). This results in

$$p_{near} = \mathbf{a}_{mic}\mathbf{D}_{mic}\mathbf{c}_{mic}, \tag{13}$$

where $\mathbf{a}_{mic}$ is the row vector of $(N_{mic}+1)^2$ spherical harmonics evaluated on the direction $(\theta_{far}, \phi_{far})$ multiplied with spherical Hankel functions evaluated at r_{near}, and $\mathbf{D}_{mic}$ is the $(N_{mic}+1)^2 \times (N_{mic}+1)^2$ diagonal matrix of distance-editing filters

$$\mathbf{D}_{mic} = diag\left(\delta_1, ..., \delta_{(N_{mic}+1)^2}\right),$$
$$\delta_{n^2+n+m+1} = \frac{h_n(kr_{near})}{h_n(kr_{far})}, \tag{14}$$

where n and m index to the columns of $\mathbf{A}_{mic}$ like in Eq. (10).

To summarize this section, for a given wave number k, the sound pressure of a proximal sound source p_{near} can be synthesized from the pressure field due to a distant source $\mathbf{p}_{far}$, as follows:

$$p_{near} = \mathbf{a}_{mic}\mathbf{D}_{mic}\mathbf{S}\mathbf{A}_{mic}^{+}\mathbf{p}_{far}, \tag{15}$$

where $\mathbf{A}_{mic}^{+}$ is computed with Eq. (9), $\mathbf{S}$ with Eq. (12), $\mathbf{D}_{mic}$ with Eq. (14), and $\mathbf{a}_{mic}$ is the row vector used in Eq. (13). The expression in Eq. (15) allows for the edition of distance information in compact microphone array recordings. This distance-editing capability can enhance the recordings by making sounds appear closer than their original distance.

5. Editing distance in HRTFs

A similar analysis to the one described in previous Section can be applied to the measurement of HRTFs. The Helmholtz' principle of reciprocity allows to formulate the measurement of HRTFs as an acoustic propagation problem [21]. Two original sound sources are assumed to be located at the listener's ears. A surrounding array of measurement points at a radius r_{far} is centered on the listener's head. Here, all the sources of scattering from the pinna, head and torso of the listener, together with the original sound source, all of them constitute the source field. When torso is not considered, the head's radius r_{head} is defined as the radius of the smallest sphere containing the head, and hence containing the source field too.

Using Eq. (1), the set of HRTFs $\mathbf{h}_{far}$ characterized for L surrounding measurements at far positions (θ_h, ϕ_h, r_h) can be written

$$\mathbf{h}_{far} = \mathbf{A}_h\mathbf{c}_h, \tag{16}$$

where $\mathbf{c}_h$ is the column vector of $(N_h + 1)^2$ spherical expansion coefficients to be determined, and $\mathbf{A}_h$ is the $L \times (N_h + 1)^2$ matrix of spherical harmonics evaluated on (θ_h, ϕ_h) multiplied with the corresponding spherical Hankel functions evaluated at r_h, both limited to a maximum order

$$N_h = \min(\lceil \frac{ekr_{head}}{2} \rceil, \lfloor \sqrt{L} - 1 \rfloor). \tag{17}$$

The HRTFs h_{near} for a sound source at an arbitrary position nearby the listener's head is synthesized from the set of HRTFs $\mathbf{h}_{far}$ characterized by measurements at far distances, as follows [16]:

$$h_{near} = \mathbf{a}_h\mathbf{D}_h\mathbf{A}_h^{+}\mathbf{h}_{far}. \tag{18}$$

The pseudo-inverse matrix $\mathbf{A}_h^{+}$ and the diagonal matrix of distance-editing filters $\mathbf{D}_h$ are also calculated with Eqs. (9) and (14), respectively, but considering the L surrounding measurements points, and therefore the dimension L instead of M, and the dimension N_h instead of N_{mic}. The desired direction is that of the recorded far source. The row vector $\mathbf{a}_h$ is hence evaluated on $(\theta_{far}, \phi_{far}, r_{near})$.

The left and right ear signals for a sound source nearby the listener's head can therefore be synthesized by filtering the sound pressure in Eq. (15) with the corresponding HRTFs of Eq. (18)

$$b_{near,left} = p_{near} \times h_{near,left},$$
$$b_{near,right} = p_{near} \times h_{near,right}. \tag{19}$$

6. Accuracy evaluation by computer simulations

We have performed an evaluation of the numerical accuracy when the synthesized sound sources lie on the horizontal plane. For this purpose we have used an average model of a human head. The left-right symmetry of the model used allowed us to evaluate the synthesis accuracy by examining the signals of only one ear.

6.1 Conditions for the evaluation

To avoid spatial aliasing, microphones and sound sources to characterize HRTFs should be placed on regular samplings of the sphere. Regular spherical samplings, though, are only possible for the platonic solids. Among existing almost-regular samplings of the sphere, we have chosen the constructions based on subdivisions of the icosahedron's edges. Figure 3 shows an example of an icosahedral grids, where dots indicate the positions of microphones or sound sources used to characterize HRTFs, and the lines enclose their neighborhoods. It can be seen that icosahedral grids provide almost constant weighting coefficients, which are proportional to the area of each neighborhood.

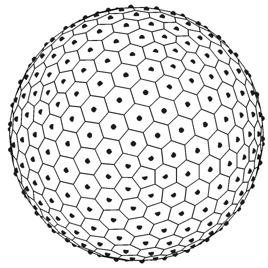

Fig. 3 *An icosahedral grid. This kind of spherical sampling was used to distribute microphones and far sound sources to characterize HRTFs.*

At this stage, we have performed an initial evaluation assuming the same number of microphones and characterizing sound sources: $M = L = 1962$. The available data allowed for spherical harmonics expansions of orders up to $N = 43$. Microphones were assumed to be placed on a baffle of radius $r_{mic} = 8.5$ cm. We consider this is the size of an average human head [22]. The smallest sphere containing the head model has a radius $r_{head} = 14.5$ cm. Under these considerations, sound fields were expected to be accurately represented by spherical harmonics up to 20 kHz, and HRTFs up to 12.4 kHz.

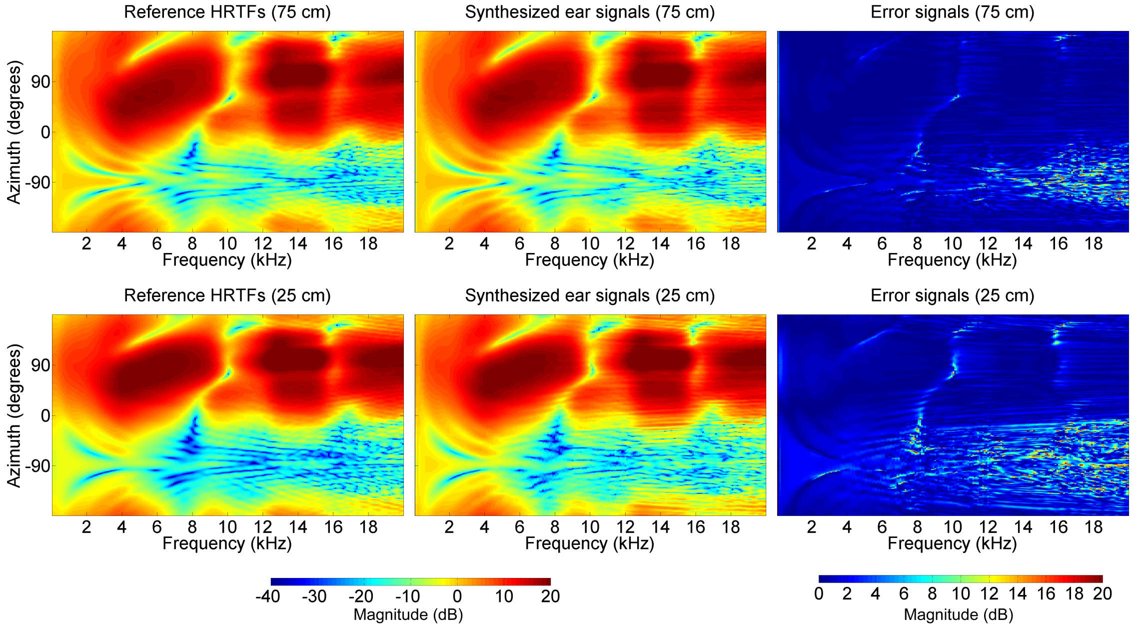

Fig. 4 *Reference HRTFs (left), synthesized signals (middle) and error signals (right) for the left ear and sound sources in the horizontal plane of elevation* $0°$. *Sound sources are at desired distances of 75 cm (top) and 25 cm (bottom). Reference HRTFs were numerically computed for a dummy head. We assumed 1962 microphones and initial HRTFs, and hence spherical harmonics decompositions of orders up to N = 43.*

We calculated the microphone signals $\mathbf{p}_{far}$ produced by a distant sound source using a model of the acoustic scattering from a rigid sphere [23]. We considered sets of microphone signals corresponding to 360 far sound sources equiangularly distributed on the horizontal plane at a radius $r_{far} = 1.5$ m. Initial sets $\mathbf{h}_far$ of HRTFs for sound sources at a distance $r_{far} = 1.5$ m were computed numerically for a dummy head using the Boundary Element Method (BEM) [24].

Transfer functions for the whole binaural synthesis process, denoted by $b(\theta, r_{near}, f)$, were characterized by using Eqs. (15), (18) and (19), for several frequencies and desired positions in the horizontal plane. We refer to this transfer functions as the synthesized ear signals. On the other hand, a reference set of near-field HRTFs, denoted by $h_{ref}(\theta, r_{near}, f)$, was also numerically computed using BEM. The resulting transfer functions for the whole binaural synthesis process were finally compared with the reference near-field HRTFs.

For a desired distance r_{near}, accuracy was calculated using the error signals between the reference HRTFs $h_{ref}(\theta, f)$ and the synthesized binaural signals $b(\theta, f)$. The error signals are defined by [25]:

$$\epsilon(\theta, f) = \left| 20 \log_{10} \left| \frac{h_{ref}(\theta, f)}{b(\theta, f)} \right| \right|. \tag{20}$$

6.2 Simulation results

Fig. 4 shows two sets of reference HRTFs (left), synthesized signals (middle), and the corresponding error signals (right). The sets corresponds to the left ear and sound sources at desired distances of 75 cm (top) and 25 cm (bottom). Sound sources lie in the horizontal plane of elevation $0°$. A visual comparison between the reference HRTFs and the synthesized ear signals shows that the synthesis for sound sources placed on the same side of the ear (azimuth from $0°$ to $180°$) is performed with good accuracy. Nevertheless, clearly decreasing accuracies appear for sound sources placed on the opposite side of the ear (azimuth from $-180°$to $0°$). In this region, the head shadowing produces signals that show rapid variations along frequency and azimuths.

Rapid variations can also be seen in positive azimuths around 10 kHz. In this region, decreasing accuracies also appear specially when the desired distance is very closed to the listener's head, as can be seen in the bottom panels in Fig. 4. However, this can also be related to the limitation of the HRTF expansion up to a maximum frequency of 12.4 kHz. A future evaluation should consider initial sets of HRTFs characterized for at least 5000 surrounding sound sources. Using this spatial resolution, the spectral cues that arise from the scattering of sound by the listener's head would be covered in full audible frequency range up to 20 kHz.

This initial evaluation was based on the pressure of a sound source near the listener's head. It was synthesized from the spherical harmonic encodings of the same sound source at the same direction but placed faraway from the listener's. The spherical harmonics encodings, though, contains the information to synthesize the sound pressure field on any spherical surface centered on the origin of the coordinate system. They can therefore also be decoded for a set of positions surrounding the listener's head by applying matrices with desired directions instead of the row vectors $\mathbf{a}_{mic}$ and $\mathbf{a}_h$ used in Eqs. (15) and (18), respectively. This would lead to a reproduction technique based on an array of surrounding virtual loudspeakers [26], which can be of interest for an evaluation in the future.

7. Conclusions

We proposed a 3D audio processing method to make sounds appear closer or farther during their binaural rendering. Our proposal combines the spherical harmonics representation of both sound field recordings and HRTFs. Our method requires a rigid spherical microphone array and the HRTFs for only some surrounding positions at a fixed distance.

Accuracy was evaluated on the horizontal plane. Accurate synthesis was obtained when sound sources lie on the same side as the ear. We noticed that the spherical harmonics expansion does not yield good approximations in regions where signals show rapid variations as functions of frequency and azimuth. However, further evaluations based on higher spatial resolution data need still to be done.

8. Acknowledgments

This study was supported by Grant-in-Aid of JSPS for Scientific Research (no. 24240016), the Foresight Program for "Ultra-realistic acoustic interactive communication on next-generation Internet", and the Cooperative Research Project Program of RIEC Tohoku University (H24/A14). The authors wish to thank Makoto Otani for his efforts in developing the BEM solver used to generate the reference HRTF data.

References

[1] Jens Meyer and Gary Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, USA, May 2002, vol. II, pp. 1781–1784.

[2] Boaz Rafaely, "Analysis and design of spherical microphone arrays," *IEEE Trans. Speech & Aud. Proc.*, vol. 13, no. 1, pp. 135–143, Jan. 2005.

[3] Earl G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*, Academic Press, London, UK, 1999.

[4] Jerome 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*, Denmark, May 2003.

[5] Henrik Møller, "Fundamentals of binaural technology," *Applied Acoustics*, vol. 36, no. 3–4, pp. 171 – 218, 1992.

[6] M. Morimoto, Y Ando, and Z Maekawa, "On head-related transfer function in distance perception," in *Proceedings of the Congress of the Acoustical Society of Japan*, Japan, 1975, pp. 137–138, (in Japanese).

[7] Douglas S. Brungart and William M. Rabinowitz, "Auditory localization of nearby sources. hear-related transfer functions," *J. Acoust. Soc. Am.*, vol. 106, no. 3, pp. 1465–1479, Sept. 1999.

[8] David R. Moore and Andrew J. King, "Auditory perception: The near and far of sound localization," *Curr. Biol.*, vol. 9, no. 10, pp. R361–R363, May 1999.

[9] Hae-Young Kim, Yôiti Suzuki, Shouichi Takane, and Toshio Sone, "Control of auditory distance perception based on the auditory parallax model," *Appl. Acoust.*, vol. 62, no. 3, pp. 245–270, Mar. 2001.

[10] Pavel Zahorik, Douglas S. Brungart, and Adelbert W. Bronkhorst, "Auditory distance perception in humans: A summary of past and present research," *Acta Acust. United Ac.*, vol. 91, pp. 409–420, 2005.

[11] W. G. Gardner and K. D. Martin, "HRTF measurements of a KEMAR," *J. Acoust. Soc. Am.*, vol. 97, pp. 3907–3908, June 1995.

[12] V.R. Algazi, R.O. Duda, D.M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, 2001, pp. 99–102.

[13] Alan Kan, Craig Jin, and Andre van Schaik, "A psychophysical evaluation of near-field head-related transfer functions synthesized using a distance variation function," *J. Acoust. Soc. Am.*, vol. 125, no. 4, pp. 2233–2242, Apr. 2009.

[14] Ramani Duraiswami, Dmitry N. Zotkin, and Nail A. Gumerov, "Interpolation and range extrapolation of HRTFs," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2004, vol. 4, pp. 45–48.

[15] Wen Zhang, Thushara Abhayapala, and Rodney A. Kennedy, "Insights into head-related transfer function: spatial dimensionality and continuous representation," *J. Acoust. Soc. Am.*, vol. 127, no. 4, pp. 2347–2357, Apr. 2010.

[16] Martin Pollow, Khoa-Van Nguyen, Olivier Warusfel, Thibaut Carpentier, Markus Müller-Trapet, Michael Vorländer, and Markus Noisternig, "Calculation of head-related transfer functions for arbitrary field points using spherical harmonics," *Acta Acust. United Ac.*, vol. 98, no. 1, pp. 72–82, Jan. 2012.

[17] T.D. Abhayapala, T.S. Pollock, and R.A. Kennedy, "Characterization of 3D spatial wireless channels," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, Oct. 2003, vol. 1, pp. 123–127 Vol.1.

[18] A. Neumaier, "Solving ill-conditioned and singular linear systems: A tutorial on regularization," *SIAM Review*, vol. 40, no. 3, pp. 636–666, 1998.

[19] Bård Støfringsdal and Peter Svensson, "Conversion of discretely sampled sound field data to auralization formats," *J. Aud. Eng. Soc.*, vol. 54, no. 5, pp. 380–400, May 2006.

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

[21] Dmitry N. Zotkin, Ramani Duraiswami, Elena Grassi, and Nail A. Gumerov, "Fast head-related transfer function measurement via reciprocity," *J. Acoust. Soc. Am.*, vol. 120, no. 4, pp. 2202–2215, Oct. 2006.

[22] Shuichi Sakamoto, Satoshi Hongo, R. Kadoi, and Yôiti Suzuki, "SENZI and ASURA: new high-precision sound-space sensing systems based on symmetrically arranged numerous microphones," in *Proceedings of the Second International Symposium on Universal Communication*, 2008, pp. 429–434.

[23] J. J. Bowman, T. B. A. Senior, and P.L.E. Uslenghi, *Electromagnetic and acoustic scattering by simple shapes*, Hemisphere, New York, NY, USA, 1987.

[24] Makoto Otani and Shiro Ise, "Fast calculation system specialized for head-related transfer function based on boundary element method," *J. Acoust. Soc. Am.*, vol. 119, no. 5, pp. 2589–2598, May 2006.

[25] Takanori Nishino, Naoya Inoue, Kazuya Takeda, and Fumitada Itakura, "Estimation of HRTFs on the horizontal plane using physical features," *Applied Acoustics*, vol. 68, pp. 897–908, Feb. 2007.

[26] Markus Noisternig, Markus Sontacchi, Alois Musil, and Robert Holdrich, "A 3D ambisonic based binaural sound reproduction system," in *Audio Engineering Society 24th International Conference: Multichannel Audio, The New Reality*, Graz, Austria, June 2003, Audio Engineering Society.