3rd International Conference on 3D Systems and Applications
Seoul, Korea, 20~22 June 2011.

Experimental evaluation of hybrid 3D sound field reproduction integrating multi-point sound field control and wave field synthesis

Noriyoshi Kamado, Hiroshi Saruwatari, Kiyohiro Shikano
Nara Institute of Science and Technology
{noriyoshi-k, sawatari, shikano}@is.naist.jp

Abstract

For a reproduced sound field, the competing goals between the listening area and reproduction accuracy in an actual environment is one of the most important problems in sound field reproduction using loudspeakers. In previous work, we proposed the method of balancing these goals with absolute accuracy using an inverse filter of the room acoustics: the null space of a generalized inverse matrix given by a compensation filter of the wave field outside the control points. In this paper, we evaluate the proposed method through the wavefront measurement. To develop an expression for the compensation filter, we use the loudspeaker driving function of wave field synthesis (WFS) to overcomes the compensation limitation of auditory distance and azimuth perception outside the control points. The results of objective evaluations revealed that the proposed method balances the competing goals and has wide applicability in a spatial domain with high accuracy of reproduction under the actual environment.

Keywords— Wave field synthesis, MINT, Null space embedding, Singular value decomposition

1. Introduction

The ultimate objective of 3D sound field reproduction using loudspeakers is to perfectly reproduce the characteristics of natural hearing over the entire spatial and frequency domains. Reproduction methods are classified into two groups: reproduction of multi-point pressure and that of wavefronts. Many of the systems in the first group are based on inverse filtering, and those in the second group are based on wavefront synthesis.

The systems based on the multiple input/output inverse theorem (MINT) [1] is one of the typical systems in the first group. However, outside the control points (*sweet spot*), the inverse filter does not compensate satisfactorily, and it is known that a method based on the inverse filtering of multi-point controlled reproduction is sensitive to user movements.

On another front, in recent years, many of the systems in the second group, based on wavefront synthesis, have been extensively investigated. *Wave field synthesis* (WFS) [2] is premised on an anechoic reproduction environment and offers a large listening area with a high perceptual reproduction quality for multiple listeners. However, the reproduction accuracy decreases in practice owing to the inherent disadvantages of physical inaccuracies found in the Kirchhoff-Helmholtz integral and room reverberation. Therefore, WFS cannot accurately reproduce the sound field at the sweet spot in real environments.

To mitigate the above-mentioned trade-off problem between the accuracy and reproducible region of the reproduced sound field in both of conventional methods, we have proposed an improved inverse filter design which enables high-accuracy reproduction with a wide listening area by integrating our previously proposed multi-point sound field reproduction method and WFS [3]. In the proposed method, the wavefront for the desired spatial cue outside the sweet spot is derived from the approximation of the WFS-synthesized wavefront, and is embeded into the subspace of the inverse filter, preserving the sweet-spot sound perfectly.

In previous work, we revealed that the proposed method can balances the competing goals and has wide applicability in a spatial domain with high accuracy of reproduction through the numerical simulations. In this paper, we evaluate the efficiencies of our proposed method in actual environment through the wavefront measurements.

The rest of this paper is organized as follows. In Section 2, the principle formulations of WFS, MCWS and our proposed method are explained. In Section 3, wavefront measurement experiments in actual environment are described. Following a discussion on the results of the experiments, we present our conclusions in Section 4.

2. Theory

2.1 WFS

In this section, WFS, MCWS and our proposed method fare described theoretically and the equations used for sound field reproduction are derived in detail. The geometric configuration and parameters in WFS are depicted in Fig. 1, where $S_P(\omega)$ and $S_{Sn}(\omega)$ denote the spectra of the primary and nth secondary sources, respectively, on the x-y horizontal plane.

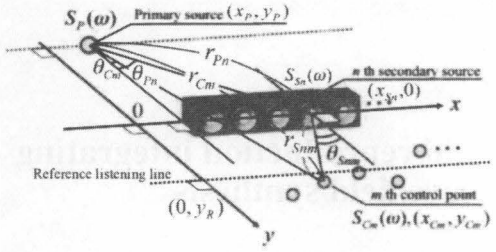

Figure 1 Configuration of WFS and MCWS.

The spectrum of the nth secondary source, which synthesizes the primary spherical wavefront, $S_{Sn}(\omega)$ is expressed as [4]

$$S_{Sn}(\omega) = Q_n^{(\mathrm{WFS})}(\omega) S_P(\omega) = S_P(\omega)\sqrt{\frac{jk}{2\pi}} C(y_R, y_P) \frac{\exp(-jkr_{Pn})}{\sqrt{r_{Pn}}} \frac{\cos\theta_{Pn}}{G(\theta_{Pn},\omega)} \Delta x, \quad (1)$$

where $Q_n^{(\mathrm{WFS})}(\omega)$ is the wavefront synthesis filter, j is the imaginary unit, k is the wavenumber (ω/c), c is the sound velocity, ω is the angular frequency, Δx is the interelement interval among the secondary sources, r_{Pn} is the distance between the primary source and the nth secondary source, and θ_{Pn} is the angle between the y-axis and the line connecting the nth secondary and primary sources. $G(\theta_{Pn},\omega)$ is a distance-independent directivity function defined only under far-field conditions. $C(y_R, y_P)$ is a function that compensates for the level of mismatch due to the stationary phase approximation along the x-direction [5], which is a function of only the reference listening distance y_R and is given as

$$C(y_R, y_P) = \sqrt{\frac{|y_R|}{|y_R - y_P|}}. \quad (2)$$

2.2 MCWS

The geometric parameters of MCWS are shown in Fig. 1. MCWS controls the spatial spectra at the control points, which are located on the x-y horizontal plane in front of the secondary sources, and generates the desired wavefront. Here, $S_{Cm}(\omega)$ denotes the secondary wavefront spectrum at the mth control-point position. Also, θ_{Cm} is the angle between the y-axis and the line connecting the mth control point and the primary source, θ_{Snm} is the angle between the y-axis and the line connecting the mth control point and the nth secondary source, r_{Cm} is the spatial distance between the mth control point and the primary source, r_{Snm} is the spatial distance between the mth control point and the nth secondary source, N is the number of secondary sources, and M is the number of control points.

Here, we derive the spectrum of the secondary source $S_{Sn}(\omega)$, which synthesizes the primary spherical wavefront. The transfer function between the nth secondary monopole source and the mth control point, $H_{nm}(\omega)$, is written as

$$Z_{nm}(\omega) = \frac{\exp(-jkr_{Snm})}{r_{Snm}}, \quad (3)$$

From Eq. (3), we define the transfer function matrix

$$\boldsymbol{Z}(\omega) = \begin{bmatrix} Z_{1,1}(\omega) & Z_{2,1}(\omega) & \cdots & Z_{N,1}(\omega) \\ Z_{1,2}(\omega) & Z_{2,2}(\omega) & \cdots & Z_{N,2}(\omega) \\ & \vdots & & \\ Z_{1,M}(\omega) & Z_{2,M}(\omega) & \cdots & Z_{N,M}(\omega) \end{bmatrix}. \quad (4)$$

We write the secondary wavefront spectrum vector at the mth control-point position as

$$\boldsymbol{S_C}(\omega) = \boldsymbol{Z}(\omega)\boldsymbol{S_S}(\omega), \quad (5)$$

where

$$\boldsymbol{S_C}(\omega) = [S_{C1}(\omega), S_{C2}(\omega), \cdots, S_{CM}]^{\mathrm{T}}, \quad (6)$$

$$\boldsymbol{S_S}(\omega) = [S_{S1}(\omega), S_{S2}(\omega), \cdots, S_{SN}]^{\mathrm{T}}, \quad (7)$$

and the superscript T denotes the transpose of the vector/matrix. If the primary wavefront spectrum is equal to the secondary wavefront spectrum at the control-point position, Eq. (5) can be transformed into

$$\boldsymbol{S_C}(\omega) = \boldsymbol{W}(\omega) S_P(\omega), \quad (8)$$

where

$$\boldsymbol{W}(\omega) = \left[\frac{e^{-jkr_{C1}}}{r_{C1}}, \frac{e^{-jkr_{C2}}}{r_{C2}}, \cdots, \frac{e^{-jkr_{CM}}}{r_{CM}} \right]^{\mathrm{T}}. \quad (9)$$

From Eqs. (5) and (8) and the Moore-Penrose (MP) generalized inverse matrix of $\boldsymbol{Z}(\omega)$, $\boldsymbol{Z}^{+}(\omega)$, we obtain the secondary source spectrum vector $\boldsymbol{S_S}(\omega)$ with a wavefront synthesis filter of MCWS $\boldsymbol{Q}^{(\mathrm{MC})}(\omega)$ in the form,

$$\boldsymbol{S_S}(\omega) = \boldsymbol{Q}^{(\mathrm{MC})}(\omega) S_P(\omega) = \boldsymbol{Z}^{+}(\omega)\boldsymbol{W}(\omega) S_P(\omega). \quad (10)$$

2.3 Proposed method

In the proposed method, the wavefront outside the sweet spot is derived from an approximation of the WFS-synthesized wavefront, and we insert it in the subspace in the inverse filter matrix. As the result, we can simultaneously achieve the perfect sound pressures in the sweet spot (control points) that are not disturbed by the WFS wavefront, and perceive the approximated wavefront reproduced by WFS outside the sweet spot. The detailed algorithm is described below.

Utilizing singular value decomposition, the generalized inverse matrix $\boldsymbol{Z}^{-}(\omega)$ of the transfer impedance matrix $\boldsymbol{Z}(\omega)$ can be denoted as

$$\boldsymbol{Z}^{-}(\omega) = \underbrace{\boldsymbol{V}(\omega)}_{(N\times N)} \underbrace{\begin{bmatrix} \boldsymbol{\Lambda}(\omega) \\ \boldsymbol{S}(\omega) \end{bmatrix}}_{(N\times M)} \underbrace{\boldsymbol{U}^{\mathrm{H}}(\omega)}_{(M\times M)}, \quad (11)$$

where the superscript H denotes the complex conjugate transposition of a matrix, $\boldsymbol{V}(\omega)$ and $\boldsymbol{U}(\omega)$ are the unitary matrices whose columns are the right and left singular vectors of $\boldsymbol{Z}(\omega)$, respectively, and $\boldsymbol{\Lambda}(\omega)$ is

$$\boldsymbol{\Lambda}(\omega) = \mathrm{diag}[\lambda_1(\omega), \ldots, \lambda_M(\omega)], \quad (12)$$

where λ_m is expressed with the singular values σ_m of $\boldsymbol{Z}(\omega)$ as

$$\lambda_m(\omega) = \begin{cases} \frac{1}{\sigma_m(\omega)} & (\text{if } \sigma_m(\omega) \neq 0), \\ 0 & (\text{otherwise}). \end{cases} \quad (13)$$

The MP generalized inverse matrix $\boldsymbol{Z}^{+}(\omega)$ can be obtained by setting $\boldsymbol{S}(\omega)$ to be a zero matrix. However, the MP-type inverse filter is specific to the reproduction at the control points and the reproduction cannot be guaranteed outside the control points. Thus, the sound localization degrades considerably when the user moves from a controlled area.

Next, to approximate $\boldsymbol{T}(\omega)$, which is the wavefront control filter outside the control points in the subspace (or nullspace) of $\boldsymbol{Z}^{-}(\omega)$ with arbitrary components $\boldsymbol{S}(\omega)$ in Eq. (11), obtain the generalized inverse matrix $\boldsymbol{Z}^{-}(\omega)$ closest to $\boldsymbol{T}(\omega)$. We utilize the Frobenius norm as the distance measure and we obtain $\boldsymbol{Z}^{-}(\omega)$ to minimize $F(\omega) = \|\boldsymbol{Z}^{-}(\omega) - \boldsymbol{T}(\omega)\|_{\mathrm{F}}$. Since the Frobenius norm is not changed by the multiplication of unitary matrices, $F(\omega)$ can be rewritten as

$$\begin{aligned} F(\omega) &= \|\boldsymbol{V}^{\mathrm{H}}(\omega)(\boldsymbol{Z}^{-}(\omega) - \boldsymbol{T}(\omega))\boldsymbol{U}(\omega)\|_{\mathrm{F}} \\ &= \left\| \begin{bmatrix} \boldsymbol{\Lambda}(\omega) - \boldsymbol{V}^{\mathrm{H}}_{\mathrm{span}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) \\ \boldsymbol{S}(\omega) - \boldsymbol{V}^{\mathrm{H}}_{\mathrm{null}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) \end{bmatrix} \right\|_{\mathrm{F}}, \end{aligned} \quad (14)$$

where $\boldsymbol{V}_{\mathrm{span}}(\omega)$ is a matrix composed of the first M columns of $\boldsymbol{V}(\omega)$. Since $\boldsymbol{\Lambda}(\omega)$ is a constant matrix, $F(\omega)$ can be minimized if and only if $\boldsymbol{S}(\omega) - \boldsymbol{V}^{\mathrm{H}}_{\mathrm{null}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) = 0$; thus, the optimal inverse filter is obtained as follows by setting $\boldsymbol{S}(\omega) = \boldsymbol{V}^{\mathrm{H}}_{\mathrm{null}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega)$ in Eq. (11):

$$\begin{aligned} \boldsymbol{Z}^{-}_{\mathrm{opt}}(\omega) &= \underset{\boldsymbol{Z}^{-}(\omega)}{\mathrm{argmin}} F(\omega) \\ &= \boldsymbol{V}(\omega) \begin{bmatrix} \boldsymbol{\Lambda}(\omega) \\ \boldsymbol{V}^{\mathrm{H}}_{\mathrm{null}}(\omega)\boldsymbol{T}(\omega)\boldsymbol{U}(\omega) \end{bmatrix} \boldsymbol{U}^{\mathrm{H}}(\omega). \end{aligned} \quad (15)$$

Next, we design the filter to guarantee the sound field accuracy outside the control points. As a method of generating the desired wavefront, WFS was introduced in this study. From Eq. (1), the spectrum at the control points can be written in terms of the impedance matrix $\boldsymbol{Z}(\omega)$ as

$$\boldsymbol{S}_C(\omega) = \boldsymbol{Z}(\omega)\boldsymbol{Q}^{(\mathrm{WFS})}(\omega)\boldsymbol{S}_P(\omega), \quad (16)$$

$$\boldsymbol{Q}^{(\mathrm{WFS})}(\omega) = \left[Q_1^{(\mathrm{WFS})}(\omega), \ldots, Q_N^{(\mathrm{WFS})}(\omega)\right]^{\mathrm{T}}. \quad (17)$$

Equally, the spectrum of MCWS at the same control points can be written as

$$\boldsymbol{S}_C(\omega) = \boldsymbol{Z}(\omega)\boldsymbol{Q}^{(\mathrm{MC})}(\omega)\boldsymbol{S}_P(\omega) = \boldsymbol{Z}(\omega)\boldsymbol{Z}^{+}(\omega)\boldsymbol{W}(\omega)\boldsymbol{S}_P(\omega). \quad (18)$$

Equation (17) is equivalent to Eq. (18) because WFS and MCWS synthesize identical primary wavefronts, and from the equivalence of these equations, the filter $\boldsymbol{T}(\omega)$ has to satisfy the condition $\boldsymbol{Q}^{(\mathrm{WFS})}(\omega) = \boldsymbol{T}(\omega)\boldsymbol{W}(\omega)$. Therefore, the filter $\boldsymbol{T}(\omega)$ is obtained as

$$\boldsymbol{T}(\omega) = \boldsymbol{Q}^{(\mathrm{WFS})}(\omega)\boldsymbol{W}_{\mathrm{ortho}}(\omega), \quad (19)$$

where $\boldsymbol{W}_{\mathrm{ortho}}(\omega)$ is an orthonormal vector of $\boldsymbol{W}(\omega)$ can be written as

$$\boldsymbol{W}_{\mathrm{ortho}}(\omega) = \frac{1}{M}\left[\frac{1}{W_n(\omega)}\right]^{\mathrm{T}}_{\boldsymbol{W}(\omega)}, \quad (20)$$

where $[\cdot]_{\boldsymbol{W}}$ denotes the operator to apply the operations in square brackets for each element of the vector $\boldsymbol{W}$ and n denotes the nth element of the vector operand.

3. Evaluation in actual environment

3.1 Experimental conditions and evaluation criteria

To illustrate the properties of the proposed method in actual environment, the frequency domain and spatial domain descriptions of the synthesized wavefront are used for experiments. The configuration of the acoustic system and the measurement system are shown in Fig. 2. The secondary sources are conducted via 25ch linear-array loudspeakers ($N = 25$), and the loudspeakers are live-sound SM-1A04S loudspeakers. To observe and to visualize the synthesized wavefront, we use the wavefront measurement system [8]. The measurement system uses 10 microphones for the measuring, and the microphones are audio-technica ATM14a omnidirectional microphones. The width and height of the obsevation area are 3.0 m on each axis. To verify the wide applicability of the proposed method in the spatial domain, we calculate the direct wavefront property of the synthesized secondary wavefront

$$\mathrm{SWF}(x_O, y_O, \omega) = \hat{\boldsymbol{Z}}(x_O, y_O, \omega)\boldsymbol{Z}^{-}_{\mathrm{opt}}(\omega)\boldsymbol{W}(\omega) * H(\omega) \quad (21)$$

where x_O and y_O denote the coordinates of the observation point, including the locations of the control points and monitoring sensors, $H(\omega)$ denotes the window function for removal of the wavefront reflected from the room wall surface, $*$ is the convolution operator, $\hat{\boldsymbol{Z}}(x_O, y_O, \omega)$ denotes the spatial impedance vector which can be written as

$$\hat{\boldsymbol{Z}}(x, y, \omega) = \left[\hat{Z}_1(x, y, \omega), \cdots, \hat{Z}_N(x, y, \omega)\right], \quad (22)$$

where $\hat{Z}_n(x, y, \omega)$ denotes the spatial impedance between the nth secondary source and the point geometry (x, y) in measurement environment. In this paper, we use frequency property of Hanning window as $H(\omega)$. To evaluate the accuracy of the reproduction at the control points, we compared the normalized quadratic reproduction error of the proposed method with that of the MCWS at the control points using [6, 7]

$$E_{LS}(\omega) = \sum_{x,y} \frac{|\mathrm{SWF}(x, y, \omega) - \mathrm{PWF}(x, y, \omega)|^2}{|\mathrm{SWF}(x, y, \omega)|^2}, \quad (23)$$

where $\mathrm{PWF}(x, y, \omega)$ denotes the calculated primary point source wavefront at the evaluation point (x, y).

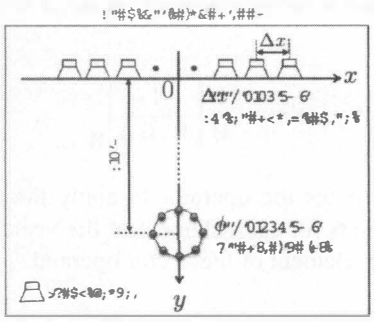

Figure 2 Configurations of loudspeaker array, control points.

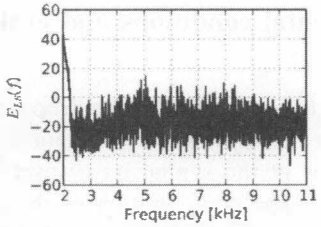

Figure 3 Normalized quadratic reproduction error $E_{LS}(\omega)$ of proposed method with reference wavefront of monopole source at the control points.

3.2 Experimental results

Figure 3 shows the normalized quadratic reproduction error $E_{LS}(\omega)$ of the proposed method with the reference wavefront of monopole source at position $(x_P, y_P, z_P) = (0.0, -1.0, 1.2)$ [m] at the control points. Figures 4(a) and 4(b) respectively show the wave fields synthesized by multi-point controlled wavefront synthesis (MCWS) [8] and the proposed method in the measurement room for a monopole source, where the radiated signal frequency is 1600 Hz. The evaluated wavefront frequency of 1600 Hz is the upper limit of major cues for sound source localization. As can be seen in Fig. 4(a), the reproduction error of MCWS is large because the commonly used inverse filter cannot guarantee the correct wavefront outside the control points. In contrast, in Figs. 2 and 4(b), the reproduction error of the proposed method is smaller than that of MCWS, and is generally smallest in the vicinity of the control points. By using WFS, the proposed method overcomes the compensation limitation of auditory distance and azimuth perception outside the control points.

The results of measurement revealed that the proposed method balances the above goals and has wide applicability in a spatial domain with high accuracy of reproduction in actual environment.

4. Concolusion

In this paper, we evaluated our proposed method through the wavefront measurement. First, we described the theory of our proposed method. To develop an expression for the compensation filter, we use the loudspeaker driving function of wave field synthesis (WFS) to overcomes the compensation limitation of auditory distance and azimuth perception outside the control points. Next, we

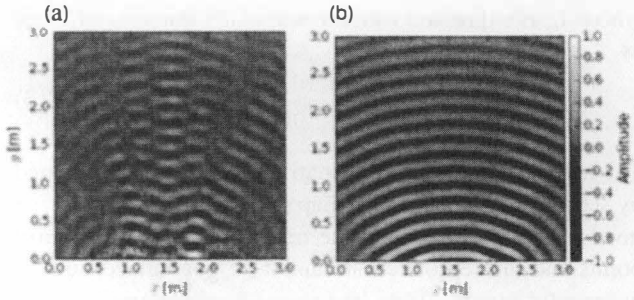

Figure 4 Measured direct wavefront synthesized at 1600 Hz in acoustic isolation room by (a) MCWS and (b) proposed method.

measured the wavefront generated by our proposed method to evaluate the efficiencies. From the measurement results showed that our proposed method can synthesize the primary wavefront with the accuracy of the reproduction at the control-points. Finally, these results revealed that the proposed method balances the competing goals and has wide applicability in a spatial domain with high accuracy of reproduction under the actual environment then numerical calculations.

References

[1] M. Miyoshi, Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. on ASSP*, vol. 36, no. 2, pp. 683–705, 1988.

[2] A. J. Berkhout, D. de Vries, P. Vogel, "Acoustic control by wave field synthesis," *J. Acoust. Soc. Am.*, vol. 93, no. 5, pp. 2764–2778, 1993.

[3] N. Kamado, H. Saruwatari, K. Shikano, "Robust sound field reproduction integrating multi-point sound field control and wave field synthesis," *Proc. ICASSP*, 2011 (to appear).

[4] D. de Vries, "Sound reinforcement by wavefield synthesis: Adaptation of the synthesis operator to the loudspeaker directivity characteristics," *J. Audio Eng. Soc.*, vol. 44, no. 12, pp. 1120–1131, 1996.

[5] E. W. Start, *Direct sound enhancement by wave field synthesis*, Ph.D. thesis, Delft University of Technology, 1997.

[6] O. Kirkeby, P. A. Nelson, "Reproduction of plane wave sound fields," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 2992–3000, 1993.

[7] P. A. Gauthier, A. Berry, W. Woszczyk, "Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain," *J. Acoust. Soc. Am.*, vol. 117, no. 2, pp. 662–678, 2006.

[8] N. Kamado, H. Hokari, S.Shoji, H. Saruwatari, K. Shikano, "Sound field reproduction by wavefront synthesis using directly aligned multi point control," *IEICE Trans. Fundamentals.*, vol. E94-A, no. 3, 2011.