

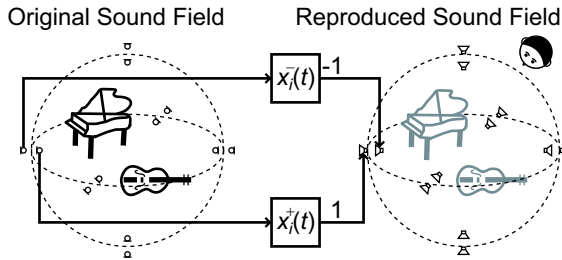
# 波面合成法を用いた近接三次元音場再生の数値解析\*

◎木村敏幸, 山肩洋子, 勝本道哲 (独立行政法人情報通信研究機構)

## 1 はじめに

波面合成法 [1-3] は原音場内に設定した制御領域の境界面上に配置したマイクロホンアレイで音を収録し, 再生音場内に設定した聴取領域の境界面上に配置したスピーカアレイで音を再生することによって, Huygens の原理に基づき制御領域の波面を聴取領域で忠実に合成する三次元音場再生技術である. この技術は聴取者の両耳位置ではなく三次元空間上の領域の音場を制御するので, バイノーラル [4, 5], トランスオーラル [6, 7] とは異なり, 聴取者はヘッドホンのような音響デバイスを装着せずに領域内の好きな場所で音を聞くことができる.

### (a) Dipole Control



### (b) Directional Point Control

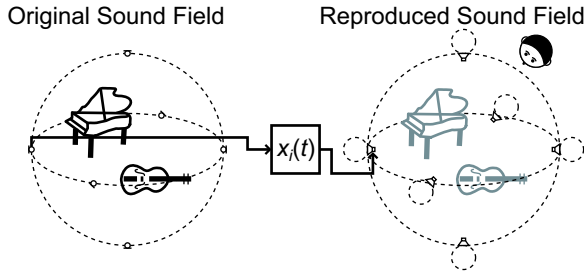


Fig. 1 Near 3D sound field reproduction system using wave field synthesis

我々は先に Fig. 1 に示すような聴取者が音源の周りの近接位置で音を聞くことができる波面合成法を用いた近接三次元音場再生技術 (双極子制御手法, 指向点制御手法) を提案しているが [8], 本報告ではその性能を計算機シミュレーションによって検討する.

## 2 計算機シミュレーション

### 2.1 シミュレーション環境

Fig. 2 に示すように 162 点の制御点  $\mathbf{r}_i$  を半径 0.4 m の球面上に, 162 点の合成点  $\mathbf{r}$  を半径 0.8 m の球面上に均等に配置した.

音源信号  $s(t)$  は振幅  $A$ , 周波数  $f$  の正弦波信号 ( $= A \sin 2\pi f t$ ) とした. 原音場における合成点  $\mathbf{r}$  での音圧  $p_0(\mathbf{r}, f, t)$  は以下のように表わされる.

$$p_0(\mathbf{r}, f, t) = \frac{1}{d_0} s\left(t - \frac{d_0}{c}\right) = \frac{A}{d_0} \sin\left\{2\pi f \left(t - \frac{d_0}{c}\right)\right\} \quad (1)$$

但し,  $d_0 (= |\mathbf{r} - \mathbf{r}_0|)$  は音源から合成点までの距離,  $\mathbf{r}_0$  は音源の位置ベクトル,  $c$  は音速を表す.

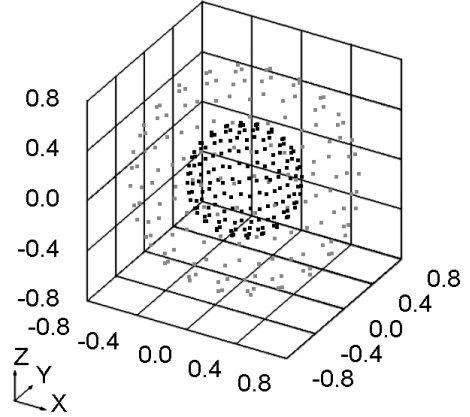


Fig. 2 Position of control points (black) and synthesis points (gray)

双極子制御手法を用いる場合,  $i$  番目のマイクロホン対で収録する信号  $x_i^+(t)$ ,  $x_i^-(t)$  は以下のように表される.

$$x_i^+(t) = \frac{1}{d_i^+} s\left(t - \frac{d_i^+}{c}\right) = \frac{A}{d_i^+} \sin\left\{2\pi f \left(t - \frac{d_i^+}{c}\right)\right\} \quad (2)$$

$$x_i^-(t) = \frac{1}{d_i^-} s\left(t - \frac{d_i^-}{c}\right) = \frac{A}{d_i^-} \sin\left\{2\pi f \left(t - \frac{d_i^-}{c}\right)\right\} \quad (3)$$

但し,  $d_i^+ (= |\mathbf{r}_i^+ - \mathbf{r}_0|)$ ,  $d_i^- (= |\mathbf{r}_i^- - \mathbf{r}_0|)$  は音源から  $i$  番目のマイクロホン対までの距離,  $\mathbf{r}_i^+$ ,  $\mathbf{r}_i^-$  は以下の式で定義される  $i$  番目のマイクロホン対の位置ベクトルを表す.

$$\mathbf{r}_i^+ = \mathbf{r}_i - \frac{\Delta d_i}{2} \mathbf{n}_i, \mathbf{r}_i^- = \mathbf{r}_i + \frac{\Delta d_i}{2} \mathbf{n}_i \quad (4)$$

但し,  $\mathbf{r}_i$  は制御点の位置ベクトル,  $\mathbf{n}_i$  は  $\mathbf{r}_i$  における境界面の外側方向への法線単位ベクトル,  $\Delta d_i (= |\mathbf{r}_i^+ - \mathbf{r}_i^-|)$  はマイクロホン対間の距離である. 従って, 再生音場において双極子制御手法によって合成点  $\mathbf{r}$  で合成される音圧  $p(\mathbf{r}, f, t)$  は  $x_i^+(t)$ ,  $x_i^-(t)$  から以下のように表される.

$$p(\mathbf{r}, f, t) = \sum_{i=1}^M \left\{ \frac{1}{d_i^+} x_i^+\left(t - \frac{d_i^+}{c}\right) - \frac{1}{d_i^-} x_i^-\left(t - \frac{d_i^-}{c}\right) \right\} \\ = \sum_{i=1}^M \left[ \frac{A}{d_i^+ d_i^0} \sin\left\{2\pi f \left(t - \frac{d_i^+ + d_i^0}{c}\right)\right\} - \frac{A}{d_i^- d_i^0} \sin\left\{2\pi f \left(t - \frac{d_i^- + d_i^0}{c}\right)\right\} \right] \quad (5)$$

但し,  $M$  はスピーカ対の総数,  $d_i^+ (= |\mathbf{r} - \mathbf{r}_i^+|)$ ,  $d_i^- (= |\mathbf{r} - \mathbf{r}_i^-|)$  は  $i$  番目のスピーカ対から合成点までの距離を表す.

指向点制御手法を用いる場合,  $i$  番目のマイクロホンで収録する信号  $x_i(t)$  は以下のように表される.

$$x_i(t) = \frac{1}{d_{i0}} s\left(t - \frac{d_{i0}}{c}\right) = \frac{A}{d_{i0}} \sin\left\{2\pi f \left(t - \frac{d_{i0}}{c}\right)\right\} \quad (6)$$

但し,  $d_{i0} (= |\mathbf{r}_i - \mathbf{r}_0|)$  は音源から  $i$  番目のマイクロホンまでの距離,  $\mathbf{r}_i$  は  $i$  番目のマイクロホンの位置ベクトルを表す. 従って, 再生音場において指向点制御手法によって合成点

\* Numerical Analysis of Near 3D Sound Field Reproduction Using Wave Field Synthesis. by T. Kimura, Y. Yamakata, and M. Katsumoto (National Institute of Information and Communications Technology)

Table 1 Parametric conditions

Source amplitude ( $A$ )	1
Source frequency ( $f$ )	125, 250, 500, 1000, 2000, 4000, 8000, 16000 Hz
Source position vector ( $\mathbf{r}_0$ )	$(0, 0, 0)^T$ $(0.3, 0, 0)^T$ $(0, 0.3, 0)^T$ $(0, 0, 0.3)^T$
Sound velocity ( $c$ )	340 m/s
Total number of control points ( $M$ )	162
Radius of control points ( $r$ )	0.4 m
Total number of synthesis points ( $N$ )	162
Radius of synthesis points ( $R$ )	0.8 m
Normal unit vector ( $\mathbf{n}_i$ )	$\mathbf{r}_i/ \mathbf{r}_i $
Neighbor distance ( $\Delta d_i$ )	0.002 m
Directivity ( $D_i$ )	Omnidirectional, Unidirectional, Shotgun

$\mathbf{r}$  で合成される音圧  $p(\mathbf{r}, f, t)$  は  $x_i(t)$  から以下のように表される。

$$p(\mathbf{r}, f, t) = \sum_{i=1}^M \frac{D_i}{d_i} x_i \left( t - \frac{d_i}{c} \right) = \sum_{i=1}^M \frac{D_i A}{d_i d_{i0}} \sin \left\{ 2\pi f \left( t - \frac{d_i + d_{i0}}{c} \right) \right\} \quad (7)$$

但し,  $M$  はスピーカの総数,  $d_i (= |\mathbf{r} - \mathbf{r}_i|)$  は  $i$  番目のスピーカから合成点までの距離,  $D_i$  は  $i$  番目のスピーカの指向特性を表す。

計算機シミュレーションに使用したパラメータ条件を Table 1 に示す。制御点間隔は最大 13 cm であり, 周波数が 1000 Hz の場合の波長 ( $= \frac{340 \text{ m}}{1000 \text{ Hz}} = 34 \text{ cm}$ ) の 1/2 以下である。従って, 1000 Hz 以下において波面を忠実に合成するための空間サンプリング定理を満たしている。また, 制御点  $\mathbf{r}_i$ , 合成点  $\mathbf{r}$  は 3 次元座標上で以下のように設定した。

$$\mathbf{r}_i = (r \cos \theta_i \cos \phi_i, r \sin \theta_i \cos \phi_i, r \sin \theta_i \phi_i)^T \quad (i = 1 \dots M) \quad (8)$$

$$\mathbf{r} = (R \cos \theta_j \cos \phi_j, R \sin \theta_j \cos \phi_j, R \sin \theta_j \phi_j)^T \quad (j = 1 \dots N) \quad (9)$$

但し,  $\theta_i, \phi_i$  は  $i$  番目の制御点の方位角と仰角,  $\theta_j, \phi_j$  は  $j$  番目の合成点の方位角と仰角を表す。一方, スピーカの指向特性  $D_i$  は以下の式によって設定した。

$$\text{(無指向性)} D_i = 1 \quad (10)$$

$$\text{(単一指向性)} D_i = \frac{1}{2} (1 + \cos \theta_{is}) \quad (11)$$

$$\text{(超指向性)} D_i = \begin{cases} \cos \theta_{is} & (|\theta_{is}| \leq 90^\circ) \\ 0 & (|\theta_{is}| > 90^\circ) \end{cases} \quad (12)$$

但し,  $\cos \theta_{is} = \frac{\mathbf{n}_i \cdot (\mathbf{r} - \mathbf{r}_i)}{|\mathbf{n}_i| |\mathbf{r} - \mathbf{r}_i|}$  である。

さらに, 合成点  $\mathbf{r}$  における音圧  $p_0(\mathbf{r}, f, t)$ ,  $p(\mathbf{r}, f, t)$  に加えて, 音の到来方向を評価するために音響インテンシティを算出した。音響インテンシティは Fig. 3 に示すようなクロススペクトル法により算出した [9]。但し,  $I_x(\mathbf{r}, f)$ ,  $I_y(\mathbf{r}, f)$ ,  $I_z(\mathbf{r}, f)$  は音響インテンシティの  $x, y, z$  成分,  $p(\mathbf{r}_x^+, f, t)$ ,  $p(\mathbf{r}_x^-, f, t)$ ,  $p(\mathbf{r}_y^+, f, t)$ ,  $p(\mathbf{r}_y^-, f, t)$ ,  $p(\mathbf{r}_z^+, f, t)$ ,  $p(\mathbf{r}_z^-, f, t)$  は 6 点 ( $\mathbf{r}_x^+, \mathbf{r}_x^-, \mathbf{r}_y^+, \mathbf{r}_y^-, \mathbf{r}_z^+, \mathbf{r}_z^-$ ) における音圧である。6 点の位置ベクトルは以下の式のように設定した。

$$\mathbf{r}_x^+ = \mathbf{r} + (\Delta, 0, 0)^T, \mathbf{r}_x^- = \mathbf{r} - (\Delta, 0, 0)^T \quad (13)$$

$$\mathbf{r}_y^+ = \mathbf{r} + (0, \Delta, 0)^T, \mathbf{r}_y^- = \mathbf{r} - (0, \Delta, 0)^T \quad (14)$$

$$\mathbf{r}_z^+ = \mathbf{r} + (0, 0, \Delta)^T, \mathbf{r}_z^- = \mathbf{r} - (0, 0, \Delta)^T \quad (15)$$

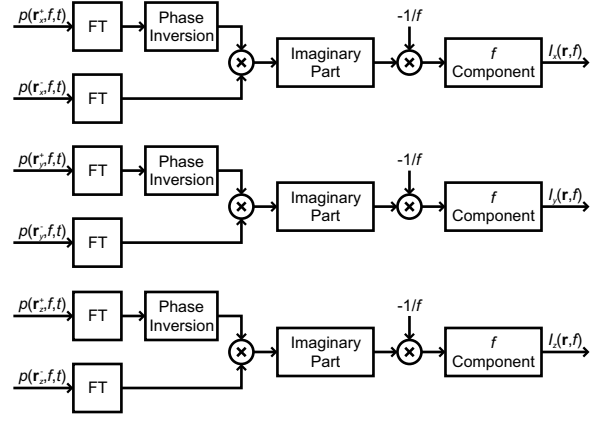


Fig. 3 Block diagram of the calculation of sound intensity (Reproduced sound field)

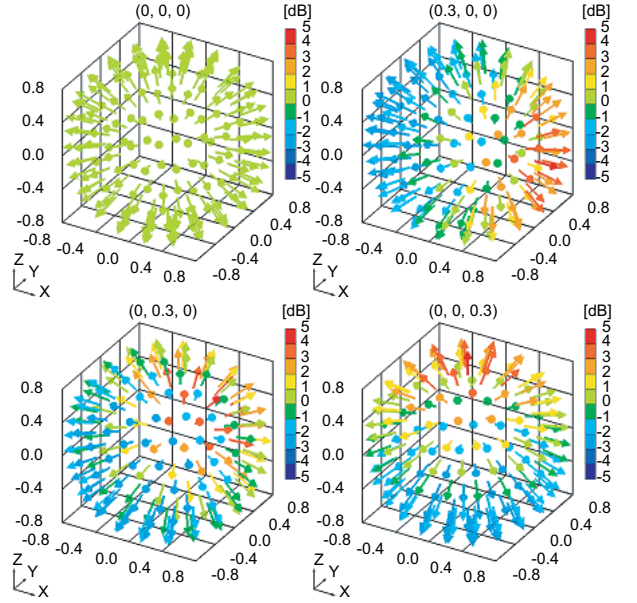


Fig. 4 RMSs of sound pressure and sound intensity vectors in the original sound field ( $f=1000 \text{ Hz}$ )

但し,  $\Delta$  は 0.001 m とした。得られた音響インテンシティベクトル  $\mathbf{I}(\mathbf{r}, f) = \{I_x(\mathbf{r}, f), I_y(\mathbf{r}, f), I_z(\mathbf{r}, f)\}^T$  を評価に用いた。また, 原音場における音響インテンシティベクトル  $\mathbf{I}_0(\mathbf{r}, f) = \{I_{x0}(\mathbf{r}, f), I_{y0}(\mathbf{r}, f), I_{z0}(\mathbf{r}, f)\}^T$  も同様に算出した。

## 2.2 シミュレーション結果

シミュレーション結果の一例として, 1000 Hz の場合の音圧と音響インテンシティベクトルの計算結果を Figs. 4-8 に示す。但し, 図中の矢印の色は以下の式で定義される原音場及び再生音場における平均二乗音圧  $p_0(\mathbf{r}, f)$ ,  $p(\mathbf{r}, f)$  である。

$$p_0(\mathbf{r}, f) = \sqrt{\int_0^1 \{p_0(\mathbf{r}, f, t)\}^2 dt} \quad (16)$$

$$p(\mathbf{r}, f) = \sqrt{\int_0^1 \{p(\mathbf{r}, f, t)\}^2 dt} \quad (17)$$

また,  $\mathbf{r}$  全体で平均二乗音圧の値を正規化し, デシベル単位に換算してから  $p_0(\mathbf{r}, f)$  と  $p(\mathbf{r}, f)$  をプロットしている。

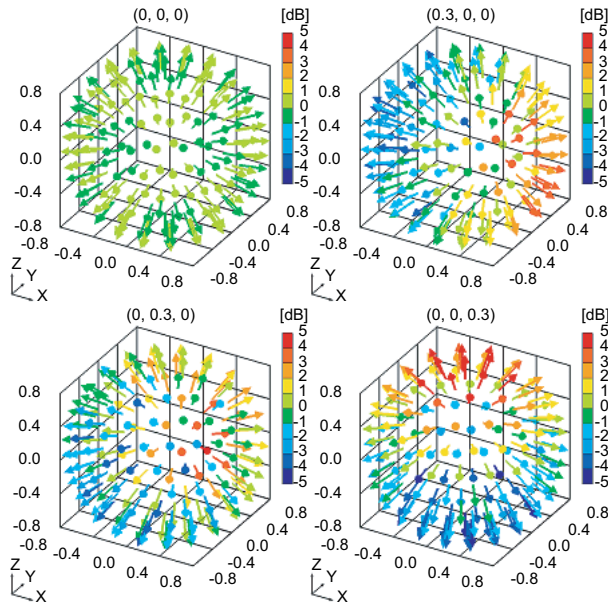


Fig. 5 RMSs of sound pressure and sound intensity vectors in the dipole control method ( $f=1000$  Hz)

全ての手法において、音響インテンシティの方向は原音場のそれに非常に近い。ゆえに、ある一点の聴取点に立ち止まって音を聞く場合、どの手法でも聴取者は音がどこで鳴っているかを判別できると考えられる。

しかし、無指向性スピーカを用いた指向点制御手法の場合は音圧分布が原音場と異なっている。それに対して、その他の手法では音圧分布は原音場とほぼ同じである。従って、スピーカアレイの回りを動き回って音を聞く場合、無指向性スピーカを用いた指向点制御手法を除く全ての手法において聴取者は臨場感を感じることができると考えられる。

さらに、全ての周波数において定量的な評価を行うために、2種類の指標を用いることにする。一つは以下の式で定義される平均二乗音圧の SNR である。これは再生音場における音圧分布が原音場とどれくらい異なるかを表す指標で、値が大きいくほど性能が良いと言える。

$$\text{SNR}(f)[\text{dB}] = 10 \log_{10} \frac{\sum_{\mathbf{r}} \{p_0(\mathbf{r}, f)\}^2}{\sum_{\mathbf{r}} \{p(\mathbf{r}, f) - p_0(\mathbf{r}, f)\}^2} \quad (18)$$

但し、 $p_0(\mathbf{r}, f)$  と  $p(\mathbf{r}, f)$  を  $\mathbf{r}$  全体で正規化を行ってから計算を行っている。もう一つは以下の式で定義されるインテンシティ方向誤差である。これは再生音場における音の到来方向が原音場とどれくらい異なるのかを表す指標で、0 度に近いほど性能が良いと言える。

$$\theta(f)[\text{degrees}] = \sqrt{\frac{1}{N} \sum_{\mathbf{r}} \left[ \cos^{-1} \left\{ \frac{\mathbf{I}(\mathbf{r}, f) \cdot \mathbf{I}_0(\mathbf{r}, f)}{\|\mathbf{I}(\mathbf{r}, f)\| \|\mathbf{I}_0(\mathbf{r}, f)\|} \right\} \right]^2} \quad (19)$$

但し、 $\mathbf{I}_0(\mathbf{r}, f)$ ,  $\mathbf{I}(\mathbf{r}, f)$  は原音場及び再生音場の音響インテンシティベクトル、 $N (= 162)$  は合成点の総数である。

計算機シミュレーションにおける提案手法ごとの SNR 及びインテンシティ方向誤差の結果を Figs. 9-10 に示す。

音源周波数が 2000 Hz 以上の場合、全ての手法において SNR は常に 20 dB 以下で、インテンシティ方向誤差は 20 度以上にまでなっている。これは計算機シミュレーションに用いた制御点の間隔のために 2000 Hz 以上では空間サンプリング定理が満たされていないからであると考えられる。

無指向性のスピーカを用いた指向点制御手法においては音源が中心以外のところでは 1000 Hz 以下においてインテ

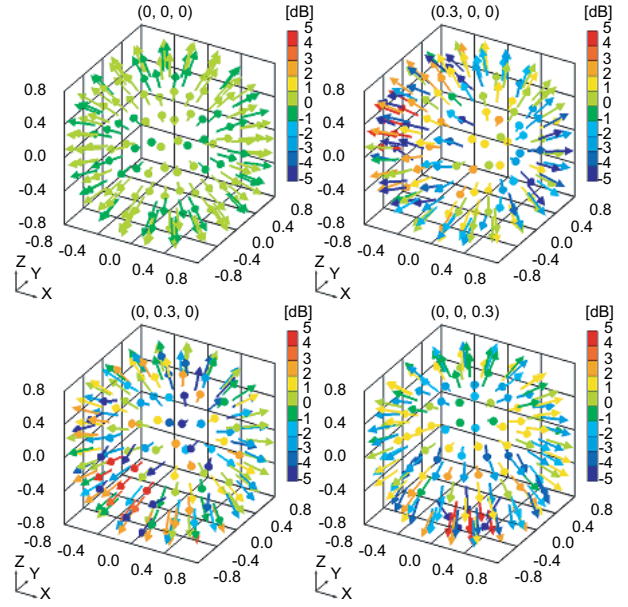


Fig. 6 RMSs of sound pressure and sound intensity vectors in the omnidirectional point control method ( $f=1000$  Hz)

ンシティ方向誤差は 12.5 度以下であるけれども、SNR は 12 dB 以下である。従って、この手法によってシステムを構築した場合、音響インテンシティが再現されているために聴取者は音がスピーカアレイ内のどこで鳴っているかを判別することはできるが、音圧分布が再現されていないために聴取者がスピーカアレイの回りを動き回ったときに臨場感を感じなくなると考えられる。

それに対して、超指向性のスピーカを用いた指向点制御手法では 1000 Hz 以下において SNR は 14.3 dB 以上で、インテンシティ方向誤差は 12.0 度以下である。同じように、単一指向性のスピーカを用いた指向点制御手法では 1000 Hz 以下において SNR は 15.0 dB 以上で、インテンシティ方向誤差は 8.4 度以下である。ゆえに、これらの手法によってシステムを構築した場合、音圧分布や音響インテンシティが十分に再現されているために、聴取者は音がスピーカアレイ内のどこで鳴っているかを判別することができるし、スピーカアレイの回りを動き回ったときでも臨場感を感じるようになると思われる。

さらに、双極子制御手法の場合、1000 Hz 以下において SNR は常に 21.3 dB 以上で、インテンシティ方向誤差は常に 4.3 度以下である。ゆえに、この手法は指向点制御手法に比べて 2 倍のマイクロホンとスピーカを必要とするが、この手法によってシステムを構築した場合、音圧分布や音響インテンシティが非常に良く再現されているために、聴取者はスピーカアレイの回りを動き回ったときにより臨場感を感じるようになると思われる。

### 3 まとめ

本報告では立体テレビや立体遠隔通信会議などの超臨場感コミュニケーションを実現するために、先に提案した 2 種類の波面合成法を用いた近接三次元音場再生技術 (双極子制御手法, 指向点制御手法) を計算機シミュレーションによって検討したところ、双極子制御手法は良い性能が得られ、指向点制御手法はスピーカに指向特性を設ければ良い性能が得られることが分かった。

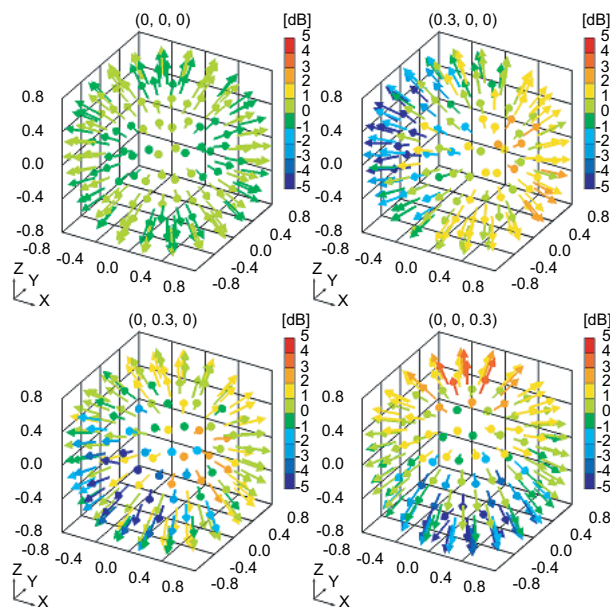


Fig. 7 RMSs of sound pressure and sound intensity vectors in the unidirectional point control method ( $f=1000$  Hz)

今後は今回提案した手法に基づいて実際にマイクロホンアレイ及びスピーカアレイを製作し、実環境における提案手法の波面合成の性能や聴取実験による提案手法の有効性を検討していく予定である。

## 参考文献

- [1] M. Camras: “Approach to recreating a sound field”, J. Acoust. Soc. Am., **43**, 6, pp. 1425–1431 (1968).
- [2] A. J. Berkhout, D. de Vries and P. Vogel: “Acoustic control by wave field synthesis”, J. Acoust. Soc. Am., **93**, 5, pp. 2764–2778 (1993).
- [3] 木村, 寛: “波面合成法による立体音場再生におけるマイクロホン及びスピーカの指向特性による波面の合成精度への影響”, 日本バーチャルリアリティ学会論文誌, **12**, 2, pp. 191–198 (2007).
- [4] J. Blauert: “Spatial Hearing”, pp. 372–392, MIT Press, Cambridge, Mass, revised edition (1997).
- [5] S. Takane, Y. Suzuki, T. Miyajima and T. Sone: “A new theory for high definition virtual acoustic display named ADVISE”, Acoust. Sci. & Tech., **24**, 5, pp. 276–283 (2003).
- [6] M. R. Schroeder, D. Gottlob and K. F. Siebrasse: “Comparative study of european concert halls: Correlation of subjective preference with geometric and acoustic parameters”, J. Acoust. Soc. Am., **56**, 4, pp. 1195–1201 (1974).
- [7] J. Bauck and D. H. Cooper: “Generalized transaural stereo and applications”, **44**, 9, pp. 683–705 (1996).
- [8] 木村, 山肩, 勝本: “波面合成法を用いた近接三次元音場再生の理論的研究”, 応用音響研究会技術報告 (submitted), 電子情報通信学会 (2008).
- [9] 橘, 矢野: “環境騒音・建築音響の測定”, pp. 20–23, コロナ社, 東京 (2004).

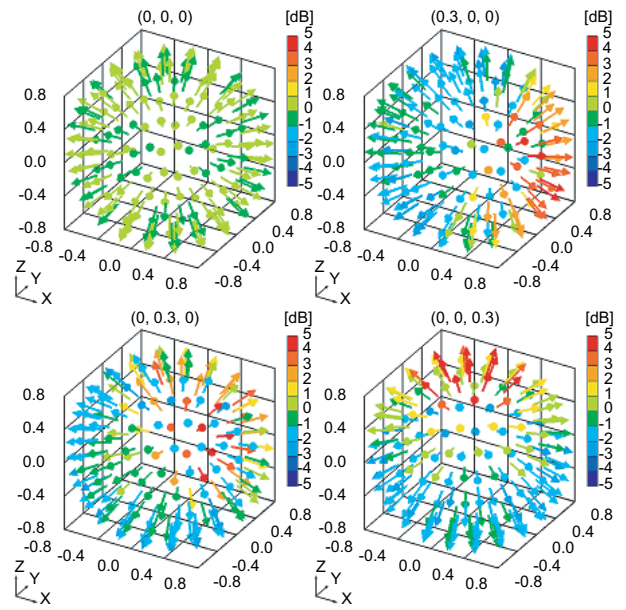


Fig. 8 RMSs of sound pressure and sound intensity vectors in the shotgun point control method ( $f=1000$  Hz)

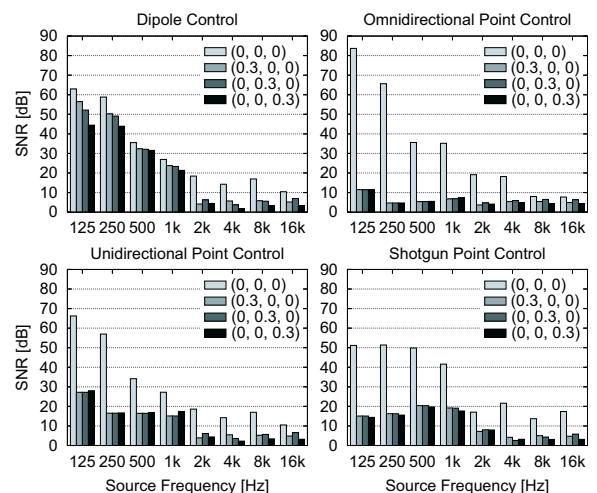


Fig. 9 SNRs in proposed methods.

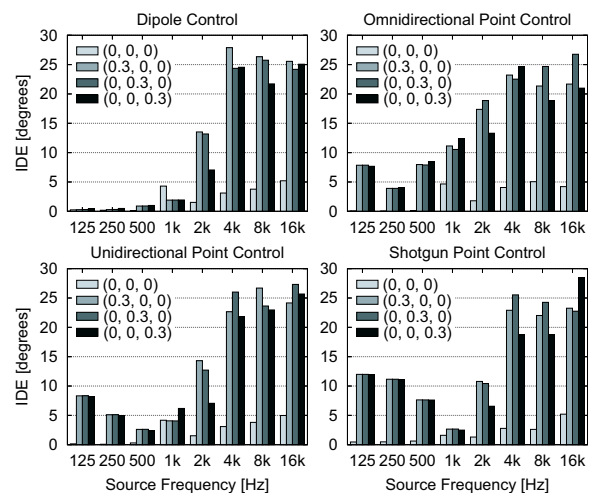


Fig. 10 Intensity direction errors in proposed methods.