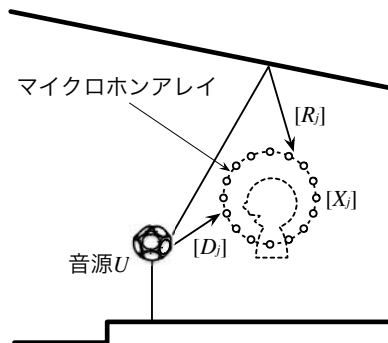


没入型聴覚ディスプレイ装置“音響樽”を用いた音場シミュレータの開発 —ステージ音場の再現とその評価—*

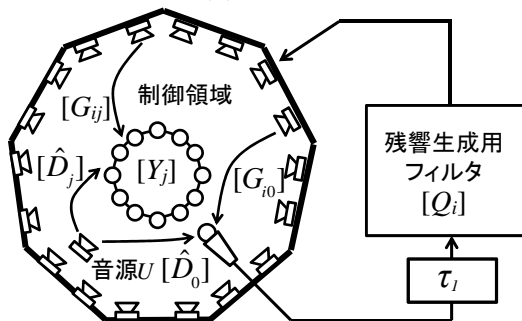
☆吉田飛里, 渡邊祐子, 池田雄介, 伊勢史郎 (東京電機大学/JST, CREST)

1 はじめに

境界音場制御の原理に基づいた音場再現システム (BoSC システム) が提案されており [1][2], 最近では内部に 96 個のスピーカを備えた没入型聴覚ディスプレイとして“音響樽”が開発されている [3]。音響樽は再生音場において受聴者が頭部を動かしても音場再現精度が保たれる特徴を有すること [4] から, 再生音場内部にて楽器を演奏した時に仮想的な室内音場を演奏者に感じさせるような音場シミュレータへの応用が期待される。そこで本研究では音響樽を用いた音場シミュレータを提案する [5]。このシステムは音響樽内の受聴者が自らの発する音 (音声や楽器音) に所望の音場の三次元的な残響を付加するシステムである。本稿では原音場としてホールの上の音場に着眼し, 音場シミュレータの再現精度を室内音響の物理指標を用いて評価する。



(a) 原音場



(b) 再生音場

Fig. 1 音場シミュレータの構成

2 音場シミュレータの構成

原音場の空間的な音響特性を再生音場に再現する音場シミュレータのシステム構成を図1に示す。原音場において音源 U が音を発した時, 80ch のマイクロホンアレイ (以下 BoSC マイクロホン) における信号 $[X_j] (\in \mathbb{C}^{1 \times 80})$ は次式で表される。

$$[X_j] = [F_j]U = [D_j + R_j]U \quad (1)$$

ここで $[F_j] (\in \mathbb{C}^{1 \times 80})$ は音源 U から BoSC マイクロホンにおける伝達特性, $[D_j] (\in \mathbb{C}^{1 \times 80})$ と $[R_j] (\in \mathbb{C}^{1 \times 80})$ はそれぞれ $[F_j]$ の直接音成分と反射音成分である。再生音場には自ら発する音を收音するための楽音用マイクロホンを設置する。ここで音響樽内の音源 U から制御点までの伝達特性を $[\hat{D}_j] (\in \mathbb{C}^{1 \times 80})$, 音響樽内の各スピーカから各制御点までの伝達特性を $[G_{ij}] (\in \mathbb{C}^{96 \times 80})$, 音源 U から楽音用マイクロホンまでの伝達特性を \hat{D}_0 , 音響樽内の各スピーカから楽音用マイクロホンまでの伝達特性を $[G_{i0}] (\in \mathbb{C}^{96 \times 1})$, 実時間処理に起因する遅延時間を τ_l とする。楽音用マイクロホンによって収録された信号は実時間畳み込み装置によって残響生成用フィルタ $[Q_i] (\in \mathbb{C}^{1 \times 96})$ と畳み込まれ, 音響樽内の各スピーカから出力される。この時, 音響樽内 BoSC マイクロホン出力 $[Y_j]$ は

$$[Y_j] = [\hat{D}_j + \hat{R}_j]U \quad (2)$$

ただし

$$[\hat{R}_j] = \frac{\hat{D}_0 [Q_i] [G_{ij}] e^{-j\omega\tau_l}}{1 - [Q_i] [G_{i0}] e^{-j\omega\tau_l}} \quad (3)$$

となる。ここで原音場の直接音成分 $[D_j]$ と再現音場の直接音成分 $[\hat{D}_j]$ が等しいとすれば, $[X_j] = [Y_j]$ とするには $[R_j] = [\hat{R}_j]$ となるような残響生成用フィルタ $[Q_i]$ を設計すればよい。ここで楽音用マイクロホンを音源 U に十分近づけ, 超指向性マイクロホンを使用することにより $\hat{D}_0 = 1$, $[G_{i0}] = 0$ が成り立つと仮定すれば, 式3から残響生成用フィルタ $[Q_i]$ は以下の式となる。

*Development of a Sound-Field Simulation System Based on Boundary-Surface Control Principle — Reproduction and evaluation of stage sound field —. by YOSHIDA, Hisato, WATANABE, Yuko, IKEDA, Yusuke, ISE, Shiro (Tokyo Denki University/ CREST, JST)

$$[Q_i] = [R'_j][H'_{ji}]e^{j\omega\tau_q} \quad (4)$$

$$\tau_q = \tau'_h - \tau_r + \tau_1 \quad (5)$$

ただし $[R'_j]$ は反射音成分 $[R_j]$ を τ_r 前方に時間シフトした信号, すなわち $[R'_j] = [R_j]e^{j\omega\tau_r}$, $[H'_{ji}]$ は因果性を保つために十分な時間 τ'_h で設計した逆システムである。したがって, 設計した逆システム $[R'_j][H'_{ji}]$ に対して, τ_q 前方に時間シフトする必要がある。

3 システムの実現

3.1 概要

音場シミュレータを以下の手順で実現する。

1. 再現する原音場の伝達特性 $[F_j]$ を BoSC マイクロホンにて測定する。(サンプリング周波数は 48kHz)
2. 実時間畳み込み装置に起因する遅延時間 τ_1 を測定する。
3. 原音場の伝達特性 $[F_j]$ の直接音成分と考えられる時間 τ_r までの範囲を 0 に置き換え $[R_j]$ として, 式 4 より残響生成用フィルタ $[Q_i]$ を設計する。
4. 音響樽内において楽音用マイクロホンで収録された信号に $[Q_i]$ を実時間で畳み込み, スピーカから出力する。

3.2 原音場の測定

原音場として空席時の残響時間が 1.5 秒の多目的ホールを選んだ。測定風景を図 2 に示す。BoSC



Fig. 2 原音場の測定風景

マイクロホンをステージ上の中央から上手方向に 2 m, ステージ縁から 2.5 m, 高さ 1.2 m の位置に設置した。音源として 12 面体スピーカ (Bruel & Kjaer Type4292) を BoSC マイクロホンの中心から水平方向に 0.35 m, 高さ 0.9 m の位置に設置し, BoSC マイクロホンへの伝達特性 $[F_j]$ を測定した。反射音成分の開始時刻 $\tau_r=25\text{ms}$ として $[F_j]$ の 0~25 ms の信号を 0 として $[R_j]$ とした。また再現精度の評価用に BoSC マイクロホンの中心位置に無指向性マイクロホン (RION NL-21) を設置し, 音源からの伝達特性を計測した。

4 音場シミュレータの再現精度の評価

4.1 計算による逆システムの評価

計測したインパルス応答から逆システムを算出し, 評価する。実時間処理に起因する遅延時間 τ_1 はオーディオ出力におけるバッファサイズなどにより変化するが, 遅延時間 τ_1 を小さくすると計算コストがかかるためチャンネル数や畳み込み時間長が短くなる可能性がある。しかし, τ_1 が大きくなると式 5 における τ_q も大きくなり, したがって設計した逆システム $[R'_j][H'_{ji}]$ を前方にシフトする時間 τ_q が増大し, 因果性が保てなくなるため音場再現精度が低下する可能性がある。そこで τ_q の変化による逆システムの再現精度を評価する。 τ'_h がそれぞれ 10, 21, 42ms (フィルタ長 N が 1024, 2048, 4096 点) の 3 種類の逆システム $[H_{jiN}]$ を計算し, τ_q を変えたときの音場再現精度を計算した。音場再現精度は次式により求めた。

$$SNR = 10 \log_{10} \frac{\sum_{j=1}^{80} |r_j|^2}{\sum_{j=1}^{80} |r_j - \hat{r}_j|^2} \quad (6)$$

ただし $[r_j]$ は原音場における反射音すなわち $[R_j]$ の逆フーリエ変換, $[\hat{r}_j]$ は再生音場における反射音であり, $[\hat{R}_j] = [Q_i][G_{ij}]$ の逆フーリエ変換である。時間シフト τ_q と音場再現精度 SNR の関係を図 3 に示す。フィルタ長 N が長い方が SNR は高く, $\tau_q = 0$ において $\tau'_h = 42\text{ms}$ ($N = 4096$) のとき 17.3 dB, 21ms (2048) のとき 13.7 dB, 10ms (1024) のとき 7.0 dB の SNR となった。また $\tau'_h = 42\text{ms}$ のとき τ_q が 30 ms を超えると減衰を始める。すなわち $N=4096$ で計算した逆システムを 3000 点以上切り詰めると再現精度が低下する。同様に $\tau'_h = 21\text{ms}$ のとき τ_q が 10 ms を超えると減衰を始める。すなわち $N=2048$ で計算

した逆システムを 1000 点以上切り詰めると再現精度が低下する。これらの結果から再現精度を維持するためには元々の逆システムの時間信号の -10 ms 以降を逆システムに含めるべきであることがわかる。

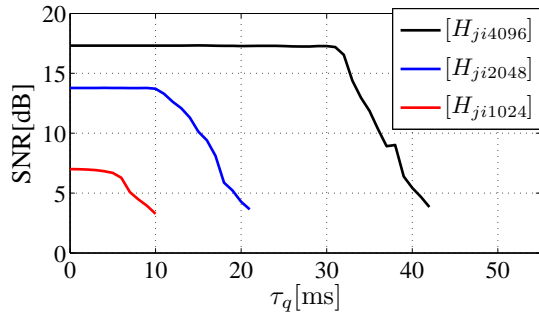


Fig. 3 τ_q と SNR の関係

4.2 計算によるシステムの評価

音場シミュレータの再現精度を測定したインパルス応答から算出する。音源から評価用マイクロホンまでのインパルス応答を直接音成分、残響生成用フィルタと音響樽内の再生用スピーカから評価用マイクロホンまでのインパルス応答の畳み込み信号を残響成分として、これらを加算した信号を評価点におけるシミュレーション音場の応答 $p(t)$ とする。実時間処理に起因する遅延時間 τ_1 は 11ms と仮定し、遅延を付加した。逆システムは 4.1 節の結果から長さ 4096 点とする。

原音場のインパルス応答、算出したインパルス応答を 1ms 毎に平均化したエネルギー波形を図 4(a)(b) に示す。原音場において直接音のピークから 60ms (図中の点線) の付近に再生音場でもほぼ同様の時刻に初期反射音を確認できる。図 5(a) に 1 オクターブ帯域毎 (中心周波数 125~8000Hz) の残響時間を示す。図中、実線は原音場、破線は再生音場における結果である。図 5(a) より原音場と再生音場の残響時間の誤差は 0.1s 以内であり良好な対応をしていることがわかる。

またホールステージ上で演奏家が自身の楽器から直接受けるエネルギーに対するホールの反射音のエネルギーの相対値として表す指標 [6] として、初期反射音エネルギー L_{er} と残響音エネルギー L_{rev} を以下のように定義する [7]。

$$L_{er} = 10 \log_{10} \frac{\int_{25ms}^{100ms} p_N(t)^2 dt}{\int_0^{10ms} p_N(t)^2 dt} \quad (7)$$

$$L_{rev} = 10 \log_{10} \frac{\int_{100ms}^{\infty} p_N(t)^2 dt}{\int_0^{10ms} p_N(t)^2 dt} \quad (8)$$

ただし $p_N(t)$ は $p(t)$ に中心周波数 N の 1 オクターブ帯域フィルタをかけた信号である。図 5(b) に初期反射音エネルギー、図 5(c) に残響音エネルギーを示す。初期反射音エネルギーと残響音エネルギーのどちらにおいても周波数特性は再生音場は原音場に近いが、1kHz 以下の周波数帯域において誤差が増えることが確認できる。

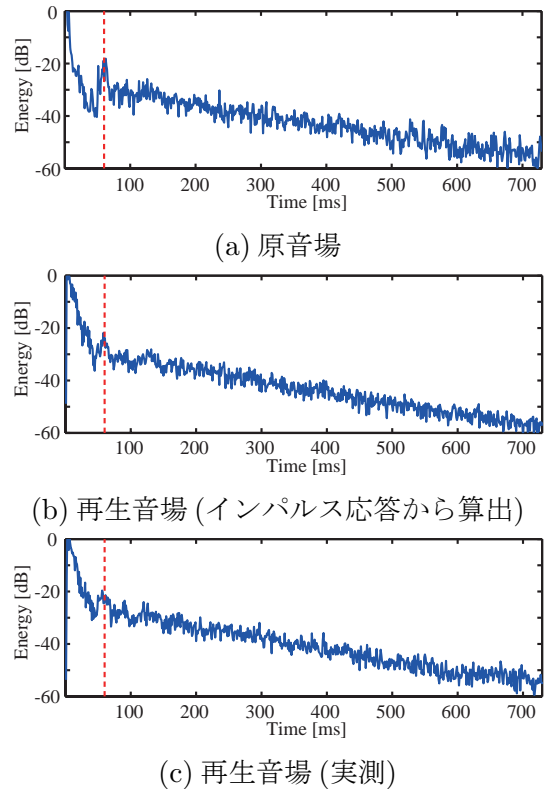


Fig. 4 インパルス応答のエネルギー波形

4.3 実測によるシステムの評価

音場シミュレータにもちいる楽音用マイクロホンとして超単一指向性マイクロホン (sennheiser MKH416) を音源から 50cm の位置に設置した。原音場における計測で用いたものと同じ音源を音響樽内に設置して TSP を出力し、楽音用マイクロホンで収録された信号に残響生成用フィルタ $[Q_i]$ を実時間で畳み込み、スピーカから出力した信号を評価用マイクロホンで測定した。実時間畳み込みは Steinberg Nuendo 6.0 Reverence により行った。実時間処理に起因する遅延時間 τ_1 は測定により 11ms であった。

測定したインパルス応答のエネルギー波形を図 4(c) に示す。初期反射音のタイミングは原音場やシミュレーション音場と比べ、突出した波形として初期反射音を確認することができなかった。これは音響樽内においてスピーカのエンクロージャによる高域の反射や内部の吸音材では

防ぎきれない低域の反射が直接音の初期反射音よりも前に多く含まれるためである。次に実測した残響時間、初期反射音エネルギー、残響音エネルギーを図6に示す。残響時間については前述の計算による再生音場の結果と同様に良好な対応が見られる。初期反射音エネルギー、残響音エネルギーに関して、実測の再生音場の周波数特性は原音場の周波数特性と傾向としては一致しているが、計算による再生音場に比べると全体に誤差が大きく、特に高域において誤差が増加する。計測中に高域におけるハウリングが生じることもあったため、実際にはフィードバックの影響は無視できないと考えられる。

5 まとめ

没入型聴覚ディスプレイ“音響樽”の一つの応用として音場シミュレータを実時間畳み込み装置を用いて実現し、その基本的な性能をステージ音場を再現することによって評価した。まず逆システムの評価においてはシステムの遅延を考慮した残響生成用フィルタの設計法の定量的に評価した結果、因果性を保つために十分な時間遅延で設計した逆システムを用いる事で再現精度を保ちながら遅延を制御することができる事を確認した。また計測したインパルス応答から求めた再生音場と実時間畳み込み装置を用いて実測した再生音場を室内音響物理指標を用いて比較した。残響時間についてはどの周波数帯域においても良好な対応が見られた。反射音エネルギーに関してはインパルス応答から求めた再生音場に比べて実システムの誤差は大きくなる事がわかった。これらの誤差は再生用スピーカから楽音用マイクロホンまでのフィードバック、および音響樽内のスピーカのエンクロージャや床からの反射が直接音以外の成分を多く含んでいることなどに起因すると考えられる。

参考文献

- [1] 伊勢, 音学誌, 53(9), 706-713 (1997.9)
- [2] 伊勢, 音学誌, 67(11), 532-537 (2011.11)
- [3] 伊勢他, 計測と制御, 51(12), 1110-1115 (2012.12)
- [4] 山下他, 音講論, 1-10-17, (2011.9)
- [5] 渡邊, 音講論, 1-Q-27, 2014
- [6] Gade, Acoustica, 69, 193-203(1989.11)
- [7] Ueno *et al.*, Acoust. Sci. & Tech. 24(3), 130-138 (2003.5)

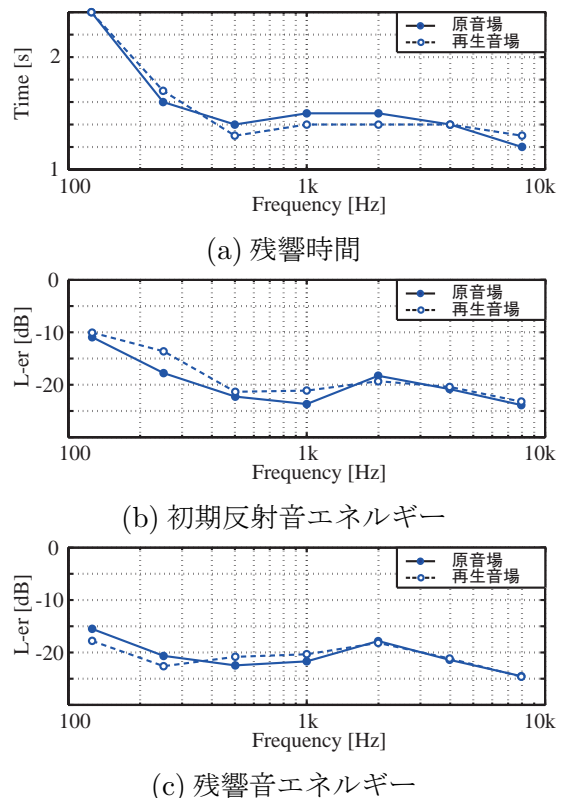


Fig. 5 インパルス応答から算出した結果

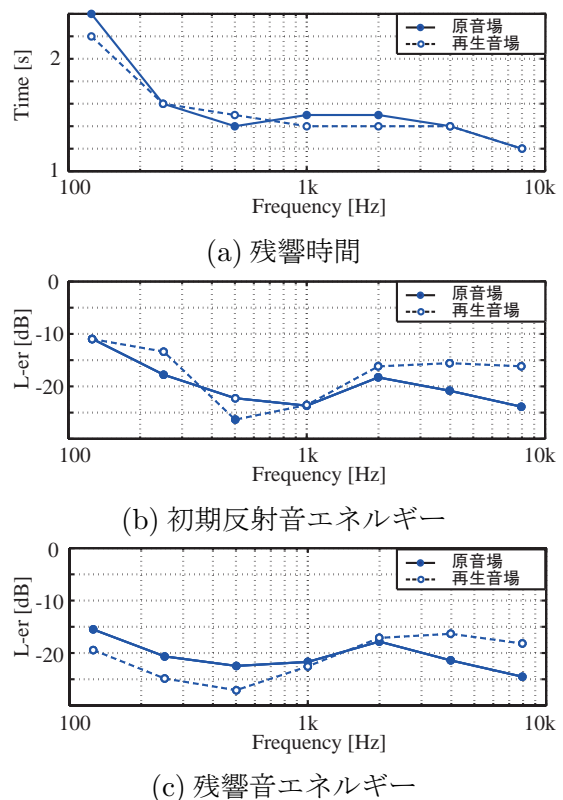


Fig. 6 実測による結果