

音響樽を用いた音場共有システムの実現と室内音響指標による評価*

☆吉田飛里, 北川雄一, 渡邊祐子(東京電機大学),
伊勢史郎(東京電機大学/JST, CREST)

1 はじめに

我々は没入型聴覚ディスプレイ装置「音響樽」を用いた音場共有システム [1] の実現を目指している。このシステムは複数の音響樽をインターネットで接続することにより、遠隔に位置する演奏者がコンサートホール等の同じ音響空間を感じながら、かつ相手の存在を感じながらアンサンブル演奏を可能とするシステムである。本稿では2つの音響樽を用いて音場共有システムを構築し、室内音響指標によりその評価を行ったので報告する。

2 音場共有システム

2.1 システム構成

2者間の音場共有システムの構成を Fig. 1 に示す。音場共有システムは、図中の (1) の原音場における音響特性を境界音場制御の原理に基づき図中 (5)(6) の音響樽内に実時間で再現することにより、遠隔に位置する人に (2) のような仮想空間を感じさせることができるシステムである。図中の (1) は演奏者 A, B が共有する原音場である。まず、この原音場内の演奏者を想定した位置に BoSC マイクロホン (C_{80} フラワーレン分子構造のフレームの節点にマイクロホンを取り付けた 80 ch マイクロホンアレイ) と演奏者が発する音を模擬した音源を設置する。この音源と BoSC マイクロホンを用いて2種類の伝達関数を測定する。ひとつは自らの音源からの残響の伝達特性を再現するための伝達関数 $[W_{j,AA}]$ と $[W_{j,BB}]$ ($\in \mathbb{C}^{1 \times 80}$)、もうひとつは相手の音源からの伝達特性を再現するための伝達関数 $[W_{j,AB}]$ と $[W_{j,BA}]$ ($\in \mathbb{C}^{1 \times 80}$) である ($j = 1, \dots, 80$ は BoSC マイクロホンの番号)。これらの特性を独立した再生音場である音響樽 A, B 内に再現し、図中の (2) のような仮想的な共有音場を演奏者に知覚させることを考える。図中 (3) の再生音場である音響樽ではあらかじめ、それぞれの音響樽内の 96 ch スピーカアレイの各チャンネルから BoSC マイクロホンおよび演奏者の発した音を収録するためのひとつのマイ

クロホン (以下、楽音用マイクロホン) までの伝達関数 $[G_{iq,A}]$, $[G_{iq,B}]$ ($\in \mathbb{C}^{96 \times 81}$) を計測し、その逆システム $[H_{qi,A}]$, $[H_{qi,B}]$ ($\in \mathbb{C}^{81 \times 96}$) を算出しておく。ここで、 $q = 1, \dots, 81$ は制御点番号であり、 $1, \dots, 80$ は j と同じく BoSC マイクロホン、 81 は楽音用マイクロホンに対応する。 i はスピーカ番号である。図中 (4) のデータベースでは、原音場や逆システムなどのデータあらかじめ蓄積されている。図中 (5)(6) の再生音場では、2つの音響樽 A, B それぞれの楽音用マイクロホンの出力に対し、 $[W_{j,AA}]$ および $[W_{j,BB}]$ の残響成分のみを再現するフィルタ (以下、残響生成用フィルタ) $[Q_{i,AA}]$, $[Q_{i,BB}]$ ($\in \mathbb{C}^{1 \times 96}$) をリアルタイムに畳み込み、それぞれの音響樽内の各スピーカから再生する。これにより原音場の音響特性が再現され、音響樽内の演奏者はあたかもその場で音を発したかのような残響を感じることで音場シミュレータを構築することができる [2]。次に、2つの音場シミュレータシステムをインターネットに接続し、それぞれの楽音用マイクロホンの出力信号を相互に伝送する。送られてきた相手の音に対し、 $[W_{j,AB}]$ および $[W_{j,BA}]$ を再現するフィルタ (以下、伝達生成用フィルタ) $[T_{i,AB}]$ および $[T_{i,BA}]$ をリアルタイムに畳み込み、音響樽内で再生する。これにより演奏者は、同一音場に相手の存在を感じることができ、音場共有が実現される。

2.2 音場再現フィルタの設計法

次に前述した2種類のフィルタの設計法について説明する。ここでは特に音響樽 A に対する音場再現フィルタについて述べる。まず、残響生成用フィルタ $[Q_{i,AA}]$ により再現する信号 $[W_{q,AA}]$ を式 (1) とする。

$$[W_{q,AA}] = [W_{j,AA} e^{j\omega\tau_r} O_1] (\in \mathbb{C}^{1 \times 81}) \quad (1)$$

ここで、 τ_r は計測した伝達関数から再現を必要としない直接音を取り除く時間長である。この時、残響生成用フィルタ $[Q_{i,AA}]$ は境界音場制御の原理に基づき式 (2) で設計される。

$$[Q_{i,AA}] = [W_{q,AA}] [H_{qi,A}] e^{j\omega\tau_q} \quad (2)$$

*Development of a Sound-Field Sharing System and its Evaluation on Room Acoustics. by Hisato YOSHIDA, Yuichi KITAGAWA, Yuko WATANABE, Shiro ISE

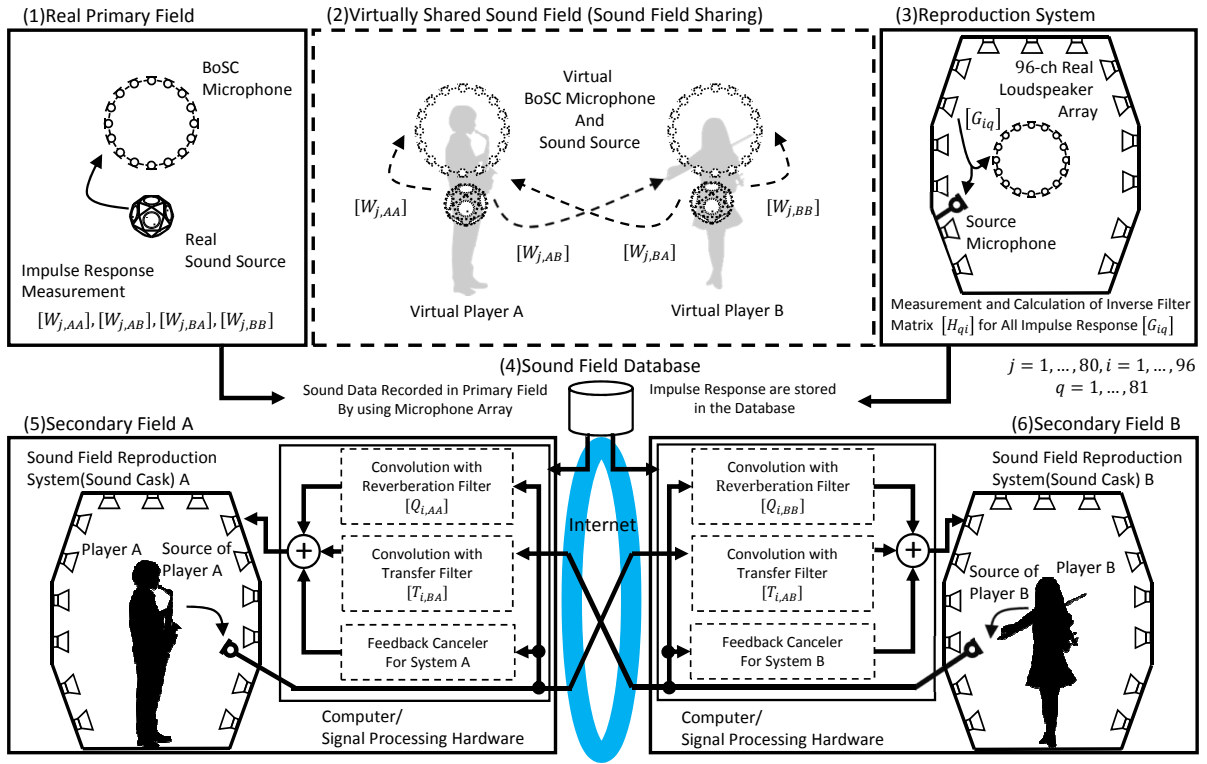


Fig. 1 Sound-Field Sharing System for Two Players

ここで、 τ_q はシステムを実現する際に生じる遅延を制御する時間シフトであり、式 (3) になれば残響生成用フィルタ $[Q_{i,AA}]$ は原音場の反射音の時間特性を再現することができる。

$$\tau_q = \tau_h + \tau_c + \tau_s - \tau_r \quad (3)$$

但し、 τ_h は逆システムの因果性を保つための遅延時間、 τ_c は実時間畳み込み処理に起因する遅延時間、 τ_s は音の入出力装置に起因する遅延時間である。この時間シフト τ_q により逆システムが一部欠落するが、逆システムの遅延が約 42 ms の場合、 τ_q が 30 ms 以内であれば音場の再現精度に影響を与えない [4]。以上により設計された残響生成用フィルタにより、音響樽 A 内において演奏者が自ら発した音に対する原音場の残響音を BoSC マイクロホンの位置に再現する。また、楽音用マイクロホンの位置には無音を再現することによって、スピーカからのフィードバックは抑制される。

次に、伝達生成用フィルタ $[T_{i,BA}]$ の設計法について述べる。このフィルタで再現する信号 $[W_{q,BA}]$ を式 (4) とする。

$$[W_{q,BA}] = [W_{j,BA} e^{j\omega\tau_d} O_1] (\in \mathbb{C}^{1 \times 81}) \quad (4)$$

ここで、 τ_d は $[W_{j,BA}]$ における直接音の到来時刻である。この時、伝達生成用フィルタ $[T_{i,BA}]$

は境界音場制御の原理に基づき式 (5) で設計される。

$$[T_{i,BA}] = [W_{q,BA}] [H_{qi,A}] e^{j\omega\tau_t} \quad (5)$$

ここで、 τ_t はシステムに存在する遅延を考慮し、直接音の到来時間を再現するための時間シフトであり式 (6) で決定される。

$$\tau_t = \tau_h + \tau_c + \tau_s + \tau_i - \tau_d \quad (6)$$

ここで、 τ_i はインターネットによる伝送遅延である。この伝達生成用フィルタにより、音響樽 B 内にいる相手の音源からの伝達特性が音響樽 A 内の BoSC マイクロホンの位置に再現されると同時に、音響樽 B を介して帰還するフィードバックが抑制される。

3 システムの実現

3.1 原音場の測定

原音場として空席時の残響時間が 1.5 s であるコンサートホールのステージ上を選び、音場の測定を行った。音源は 12 面体スピーカ (Bruel & Kjaer Type4292) を使用し、それをステージの中央から下手側に 50 cm の位置に高さ 90 cm で設置した。BoSC マイクロホンは、音源の位置から水平方向に外側 30° 、35 cm 離れた位置に高さ 120 cm、音源の方向を正面として設置した。さら

に上手側にも同じスピーカと BoSC マイクロホンを設置した。この配置は、ステージ上でのアンサンブルにおける典型的な位置と、演奏者と演奏される楽器の位置関係を参考に決定した。この配置において、ステージ下手側を演奏者 A、上手側を演奏者 B 用として、伝達関数 $[W_{j,AA}]$, $[W_{j,AB}]$, $[W_{j,BA}]$, $[W_{j,BB}]$ を測定した。測定には TSP 信号を用い、サンプリング周波数 48 kHz、量子化ビット数 24 bit とした。また、 $[W_{j,AA}]$, $[W_{j,BB}]$ からは測定に使用した音源の過渡特性を考慮し $\tau_r = 25 \text{ ms}$ とし直接音を除去した。

3.2 システムの条件

2 台の音響樽を用い音場共有システムの構築した。楽音用マイクロホンとして小型の無指向性マイクロホン (DPA 4060-BM) をそれぞれの音響樽内の扉の上部に設置した。今回使用したシステムにおいて、入出力に起因する遅延時間 τ_s は計測により 4.4 ms であった。信号の伝送にはインターネットを用いず直接接続したため、 $\tau_i = 0$ である。逆システム $[H_{qi,A}]$, $[H_{qi,B}]$ は正則化パラメータ法により設計し、正則化パラメータはオクターブバンド毎に最適化した値を用いた [5]。逆システムのフィルタ長は 4096 点であり、遅延時間 τ_h は約 42 ms であった。実時間畳み込みには FPGA を用いた低遅延畳み込み装置 [6] および低遅延畳み込みアルゴリズムによる畳み込み装置 [7] を用いた。これにより、実時間畳み込み処理に起因する遅延時間 τ_c を実用上 0 ms とすることができた。以上から、 $\tau_q = 21.4 \text{ ms}$ とし残響生成用フィルタを設計した。伝達生成用フィルタにおける時間シフト τ_t は、再現する音場の直接音到来時刻 τ_d がシステムの遅延時間と比べて非常に小さいため、式 (6) とすると再現精度が劣化することが予想される。そこで計算により音場再現精度と τ_t の関係性を評価し、音場再現精度に影響を与えない最長のシフト時間であった 30 ms とした。

4 室内音響指標による評価実験

4.1 実験方法

本章では実現した音場共有システムの評価実験を行う。BoSC マイクロホンを音響樽 B に、音源用スピーカを音響樽 A、B 内に設置し、インパルス応答を測定する。音源に最も近い 31 ch のマイクロホンにおけるインパルス応答を室内音響指標 (残響時間 RT_{60} , エコータイムパターン、

反射音エネルギー) を用いて解析し原音場と比較することで評価する。反射音エネルギーは、インパルス応答を用い式 (7)~(11) で算出した。

$$L_{er,self} = 10 \log_{10} \frac{\int_0^{100 \text{ ms}} p_{self}(t)^2 dt}{\int_0^{10 \text{ ms}} p_{self}(t)^2 dt} \quad (7)$$

$$L_{rev,self} = 10 \log_{10} \frac{\int_{10 \text{ ms}}^{\infty} p_{self}(t)^2 dt}{\int_0^{10 \text{ ms}} p_{self}(t)^2 dt} \quad (8)$$

$$L_{dir,cross} = 10 \log_{10} \frac{\int_{t'}^{t'+10 \text{ ms}} p_{cross}(t)^2 dt}{\int_0^{10 \text{ ms}} p_{self}(t)^2 dt} \quad (9)$$

$$L_{er,cross} = 10 \log_{10} \frac{\int_{t'+10 \text{ ms}}^{100 \text{ ms}} p_{cross}(t)^2 dt}{\int_0^{10 \text{ ms}} p_{self}(t)^2 dt} \quad (10)$$

$$L_{rev,cross} = 10 \log_{10} \frac{\int_{100 \text{ ms}}^{\infty} p_{cross}(t)^2 dt}{\int_0^{10 \text{ ms}} p_{self}(t)^2 dt} \quad (11)$$

但し、 $p_{self}(t)$ および $p_{cross}(t)$ はそれぞれ $W_{31,BB}$ および $W_{31,AB}$ の逆フーリエ変換であり、 $p_{self}(t)$ の直接音の到来時刻は $t = 0$ 、 $p_{cross}(t)$ の直接音の到来時刻は $t = t'$ である。

4.2 実験結果

残響時間 RT_{60} の解析結果を Fig. 2 に示す。 $p_{self}(t)$, $p_{cross}(t)$ の両方において原音場との誤差が全周波数帯域において 0.1 s であり良好な対応が見られた。

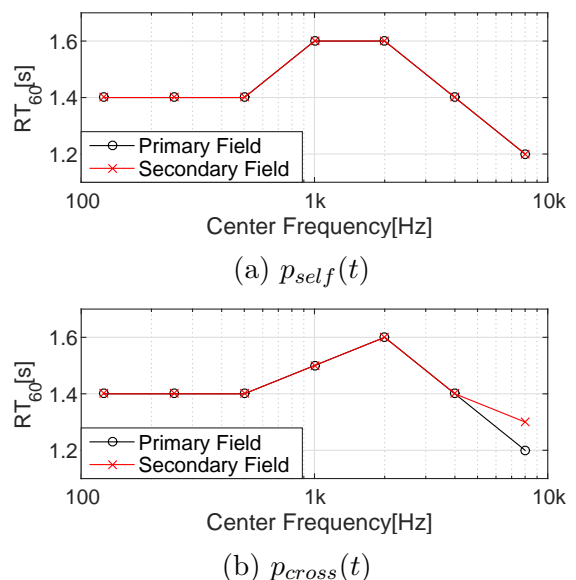


Fig. 2 Reverberation Time

次に反射音エネルギーの結果を Fig. 3 に示す。全ての結果において、再現音場における 500 Hz 以下の周波数帯域のエネルギーが約 10 dB 高く、

1 kHz の周波数帯域のエネルギーが約 10 dB 低い事が分かる。この誤差は、音響樽内の楽音用マイクロホンの収録によって生じていると考えられる。そこで、楽音用マイクロホンの入力に対し $L_{rev,self}$, $L_{rev,cross}$ の周波数特性を補正するフィルタを付加した [8]。その結果、 $L_{rev,self}$, $L_{rev,cross}$ だけでなく、その他の反射音エネルギーも概ね良好に再現することができた。

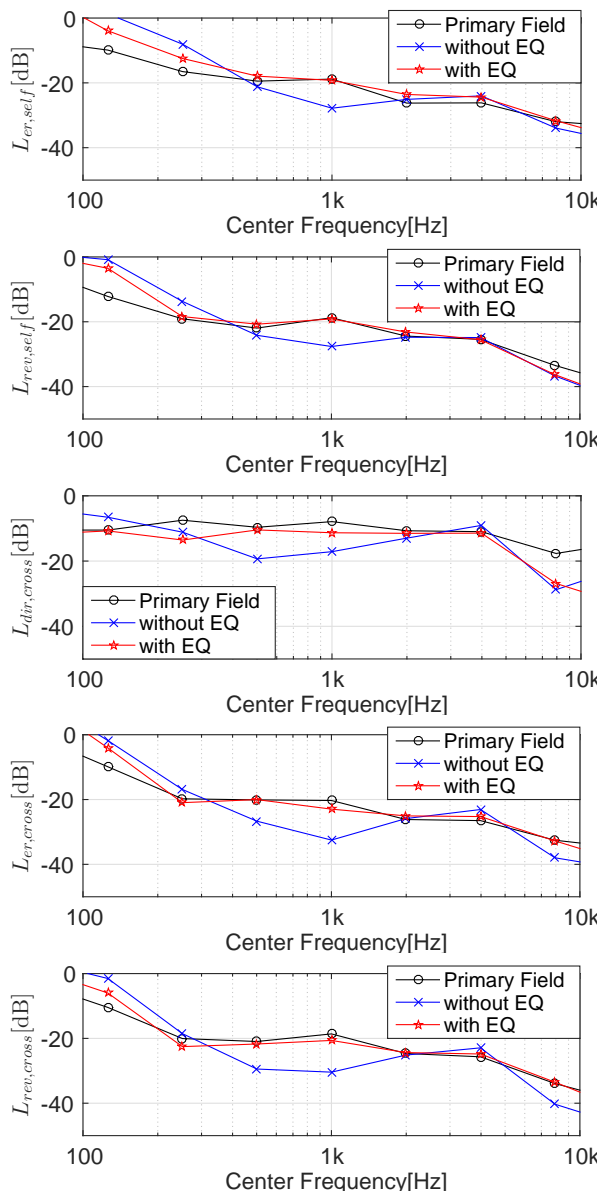


Fig. 3 Reflection Sound Energy

次に、インパルス応答の 1 ms 毎の RMS 値をプロットしたエコータイムパターンを Fig. 4 に示す。反射音の到来時刻は概ね原音場と同じであった。特に $p_{self}(t)$ は約 70 ms, $p_{cross}(t)$ は約 110 ms に現れる反射音の到来時刻が良好に再現されている。一方で、特に $p_{self}(t)$ の再現音場における直接音の減衰が原音場と比べて遅いこと

が分かる。これは音響樽内のエンクロージャや吸音材で防ぎきれない低域の反射音によるものであると考えられる。 $p_{cross}(t)$ の直接音到来時刻の遅延は約 16 ms であり、原音場との誤差は約 13 ms であった。

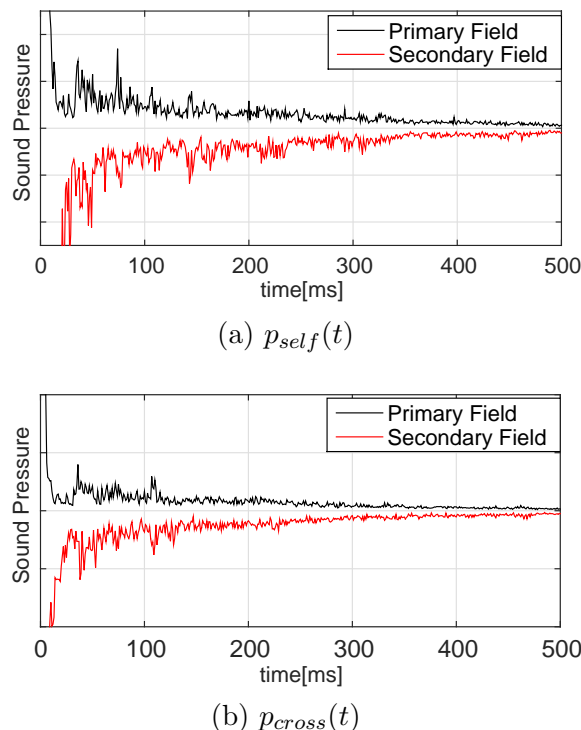


Fig. 4 Echo Time Pattern

5 おわりに

本稿では音場共有システムを実現し、その評価を室内音響指標によって行った。その結果、残響時間や反射音構造を概ね再現することができる事が分かった。反射音の周波数特性に関しては周波数特性補正フィルタを付加することにより良好な対応が見られた。

参考文献

- [1] 伊勢, 音講論 (秋), 3-5-13, 2011.
- [2] 渡邊他, TVRSJ Vol.20 No.1 pp.45-53, 2015.
- [3] 河野他, 音講論 (春), 1-10-4, 2015.
- [4] 吉田他, 音講論 (秋), 1-Q-28, 2014.
- [5] 伊勢, 音講論 (秋), 1-1-2, 2014.
- [6] 吉田他, 音講論 (秋), 1-6-4, 2015.
- [7] 北川他, 音講論 (春), 1-3-12, 2016.
- [8] 渡邊他, 音講論 (春), 1-8-14, 2015.