

没入型聴覚ディスプレイ“音響樽”がもたらす ヒトとオトの新しい関係

伊勢史郎
東京電機大学

1. はじめに

音はヒトが古代から築き上げてきた大切な道具である。しかし、音という道具の使い方は近年のインターネットや携帯電話・スマートフォンなどの急速な普及により大きく変わりつつある。特にコミュニケーション活動において音が介在する余地は日増しに減り、大量の文字情報が瞬時に遠方へ届く情報通信技術が我々の生活の中で不可欠の要素となった。利便性・合理性・効率性を追求した結果、我々のコミュニケーションは多様な形態で行うことが可能となったが、本来必要な身体的なコミュニケーションまで切り捨てられる傾向があり、そのため社会的な不満や不安は少しも解消されていないように思われる。新しい技術や文化は社会を活性化させるために必要であるが、一方でヒトの社会が豊かさを維持可能な新しい技術とはどのようなものであるか。没入型聴覚ディスプレイ“音響樽”をその一つの解として、ヒトとオトの新たな関係を探る。

2. 現代の情報通信技術に関する問題意識

2.1 コミュニケーション形態の急激な変化

ヒトがやりとりする情報には意味化が容易な情報と、意味化が困難な情報が含まれる。文字、絵、音声、ジェスチャー、表情などそれぞれの表現方法において言語・非言語情報の含まれる割合は異なり、例えば文字の中には言語情報が多く含まれ、ジェスチャーや表情の中には非言語情報が多く含まれる。インターネットでやりとりされる情報は主に文字情報であり、言語的に表現可能な意味内容は容易に伝送することが可能となるが、多くの感情は文字情報に含めることが難しく、伝送されないままコミュニケーションが行われることが多い。文字も存在しない古代の世界ではヒトは空間を共有し、音によって意思を伝えあっていたが、伝えあう意味内容以外に、共に同じ空間を過ごしたという記憶がヒトとヒトの信頼関係を築いていた。10 年以上かけて築いてきた音という道具を使用する技能は我々の DNA にすでに刻まれており、約 1 万年前から始まった文字の使用、約 1000 年前から始まった印刷技術の使用、約 100 年前から始まった電話の使用、そしてここ 10 年の携帯電話やインターネットの使用は、ヒトの身体に刻まれた自然なコミュニケーション形態を崩壊する勢いで技術革新が進んでいる。

2.2 言語と非言語

一般にコミュニケーションでは、意味を伝達する以外に感情を共有することが重要である。通常的生活空間で行われる音声によるコミュニケーションでは文字化可能な情報以外に、音声の抑揚・トーンなどに感情の情報が含まれ、さらに 3 次元波面の中に含まれる発話者の身体情報が高いリアリティを生成することにより、受話者に緊張感とともに能動的に話を聞く態度が現れる。逆に文字化可能な言語情報のみの伝達では感情を伝えることが難しく、意味発生の主体が曖昧になるため、意味を理解するための能動的な態度が生じにくくなる。携帯電話やインターネットによるコミュニケーションでは明確に指示された意味を素早く正確に伝送することが目標となり、感情の伝送は取り残されたまま様々な社会的な活動が行われることになる。現在、一向に進展しない社会問題の根底には感情を共有する機会不足が原因となっているのではないだろうか。

2.3 感情共有を可能とする音場再生技術

以上のような問題意識に対する解決の糸口として我々は情報環境とヒトとのインターフェースに着眼する。すなわちヒトが情報空間において音という道具を十分に使いこなすことができるような音響インターフェースを構築することを目標とする。話者あるいは楽器演奏者が身体を用いて発した音の 3 次元波面をありのままの状態に収録し、再生するためには 3 次元音場を空間的に収録し、再生する技術が必要となる。そのための理論基盤として「境界音場制御 (BoSC) の原理」を用いる。BoSC の原理はキルヒホッフ-ヘルムホルツ積分方程式の音場制御理論としての新しい解釈の上に、逆システム理論を適用することにより、従来技術的に不可能と言われてきた 3D 波面の完全な収録・再生が可能であることを示したものである。1993 年に基礎理論が提案されたものであるが、近年のマルチチャンネル信号処理技術の発展、電子技術の高性能化により実現可能であることが実証されつつある。

3. 音場再現システム

3.1 BoSC 原理による音場再現

図 1 のように、ある空間に境界面 S の領域 V の音場 (原音場) と、それとは別の空間に領域 V と合同となる領域 V' の音場 (再生音場) を想定する。境界音場制御の原理に基づけば、原音場においてある領域を囲む境界面 S 上の音圧と粒

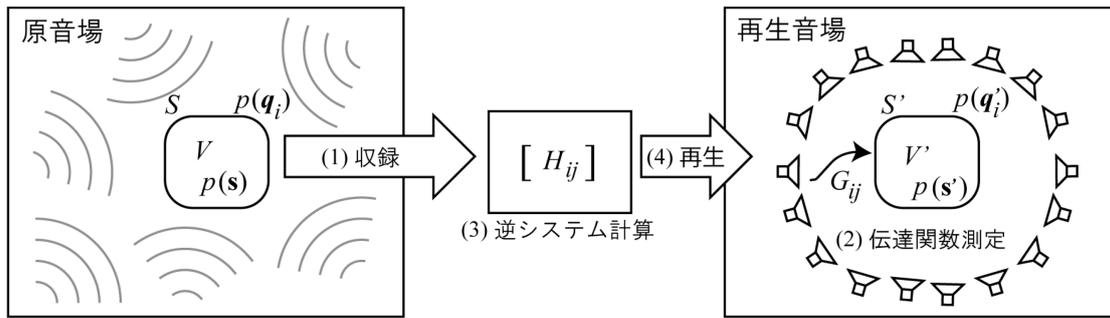


図1 BoSC 原理による音場再現

子速度(音圧勾配)を計測し、それらが再生音場の境界面 S' 上において相対的に同じ位置で再生されたとき、原音場における領域 V 内の音場は再生音場の領域 V' 内に完全に再現される[1]。ここで原音場における境界面 S 上の音圧および粒子速度は M 個のマイクロホンで計測した音圧信号により再現可能と仮定し、その位置座標を \mathbf{q}_j ($j = 1 \dots M$)とする。同様に再生音場に設置するマイクロホンの位置座標を \mathbf{q}'_j とする。原音場でのマイクロホン出力信号から得られる逆システムの入力信号ベクトルを $[X_j] (\in \mathbb{C}^{1 \times M})$ 、再生音場における L 個のスピーカからマイクロホンへの伝達関数マトリクスを $[G_{ij}] (\in \mathbb{C}^{L \times M})$ 、逆システムの伝達関数マトリクスを $[H_{ji}] (\in \mathbb{C}^{M \times L})$ 、再生音場におけるマイクロホンからの出力信号ベクトルを $[Y_j] (\in \mathbb{C}^{1 \times M})$ とすると次式が成り立つ。

$$[Y_j] = [X_j][H_{ji}][G_{ij}] \quad (1)$$

ただし、 $X_j = P(\mathbf{q}_j)$ 、 $Y_j = P(\mathbf{q}'_j)$ である。ここで式(1)が成立するためには $[Y_j] = [X_j]$ となる $[H_{ji}]$ を求めればよい。

3.2 逆システムの設計法

マイクロホンよりもスピーカの数が多い場合は逆システムを時間領域で求めることにより FIR システムとして設計できるが、本システムのように多チャンネルシステムの逆システムを時間領域で求めることは困難である。周波数領域で求める場合には式(1)を解くことにより逆システムを求めることができる。しかし、 $M < L$ の場合には逆行列を一意に求めることができないため、最小ノルム解により求める方法が提案されている[2]。最小ノルム解を与えるムーアペンローズ (MP) 一般逆行列は二次音源からの出力を最小化するため、比較的安定した逆システムの設計が期待できる。しかし、チャンネル数が増えれば条件数が過度に大きくなる可能性が増え、想定した時間範囲で収束する逆システムを設計することが困難となる。そこで正則化一般逆行列

$$[H_{ji}] = \left([G_{ij}]^\dagger [G_{ij}] + \beta I_M \right)^{-1} [G_{ij}]^\dagger \quad (9)$$

を用いる。ただし $[\cdot]^\dagger$ は行列の共役転置、 β は正則化パラ

メータ、 I_M は M 次元単位行列である。正則化パラメータを加えることにより、行列の対角成分が大きくなるためその逆行列から安定した FIR フィルタを設計することが可能となる[3,4]。

4. システムの実現

4.1 收音システム

一人の受聴者の頭部を取り囲む大きさを想定し、また堅固な力学的な構造で支えるフレーム構造として、直径約45cmのC80 フラーレン分子構造の形状のマイクロホンフレームを開発した。フレームの節の部分80か所に小型無指向性マイクロホン (DPA4060) を取り付けられた (図2)。



図2 BoSC マイクロホン

4.2 再生システム

受聴者を取り囲む閉じた境界面に高い精度で音圧波面を生成するためには、できるだけ多くの方向から波面を供給可能な音響装置の構成が必要となる。またスピーカを取り付けるためのフレームとして、より堅固な力学的構造が安全面からも重要となる。そのためには建築物に取り付けるのが容易であるが、多くの人に音場を体験してもらうために分解、運搬、組立が可能なスピーカフレームが望ましい。また精度の高い逆システムを設計するためには壁面やスピーカのエンクロージャなどによる音響的な境界条件の影響が懸念される。すなわちモードが小さく、かつ偏りが少ない境界条件をもつ音場再生室が必要となる。このよう



図3 音響樽

に多数のスピーカを支える堅固な力学構造、分解、運搬、組立の容易さ、壁面のモードの偏りを考慮して、天井を含む床以外の壁面に 96 個のスピーカを取り付け、スピーカ以外の壁面をポリウレタン（120mm 厚）で覆った平面の断面が 9 角形となる樽型形状の音場再生室を開発した（図 3）。

5. 3D 音場による実在感の心理・生理評価

BoSC システムによる再生音場の評価について“話者（発音体）の実在感”に着目した心理・生理実験として、(1)話者の微細動作を再現することの効果、(2)接近する移動音源に対する生理反応、(3)動作音・非動作音に対するミラーシステムの活動（脳波計測）の面から心理・生理実験を行った。(1)については心理的評価に加え、自律神経系の活動およびホルモン分泌を客観的指標として用い、実在感に及ぼす微細動作の効果について確認した[5]。(2)では主観評価とともに移動音源の接近に伴う自律神経系の活動の時系列変化を指標として、実在感に及ぼす空間再現性効果を確認した。その結果、空間再現性が高い場合には人の接近音が提示されることによって交感神経系の活動が高まり、再現音場においても実空間と同様の生理反応が確認された。これはパーソナルスペースの侵害が生じている場合に観測されており、仮想空間でも音によってパーソナルスペースが侵害される可能性を示す[6]。(3)ではチャンネル数および逆システムを操作することで音場の再現性を操作し検証した結果、音場再現性が高い場合にはミラーシステムの活動が活発になることを確認した(図6, 7)[7]。これらの結果から、これまでに構築した再生音場において、きわめて高い話者（発音体）の実在感が得られていること、また話者（発音体）の実在感を間接的・客観的に評価できることが示された。

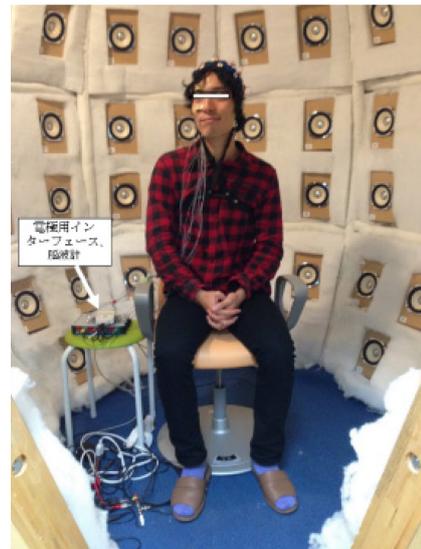


図6 音響樽内での脳波測定実験の様子

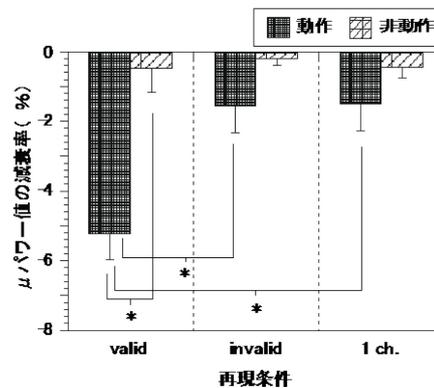


図7 μ波抑制の結果

6. 音場共有システム

6.1 音場共有システムの基本構成

2つの音響樽をネットワークで接続することにより遠隔に位置する複数のヒトが同一の音響空間を感じながら、すなわち音場共有しながらコミュニケーションをとることが可能となる[8]。例えばコンサートホールなどにおいて計測したインパルス応答と逆システムを畳み込み、図8のように音響樽においてリアルタイムで再生することにより、アンサンブル演奏が可能となる。

6.2 遅延の低減

オーケストラなどにおいて演奏者が例えば 10m 程度離れると 34ms（音速 340 m/s の場合）の遅延が生じるため、指揮者は不可欠となる。指揮者がいない2名のアンサンブル演奏では遅延の影響が深刻となるが、20ms 以内であれば演奏に影響はないため[9]、本システムでは遅延を 20ms 以下に抑えることを目標とする。音場共有システムの実現において、遅延は通信システム、逆システム、オーディオ入出力システムの3つの要因により生じるが、オーディオ入出力装置による遅延は技術的に解決可能である。そこで低遅延で 96 ch 畳み込みが可能な FPGA ボードの開発を行っ

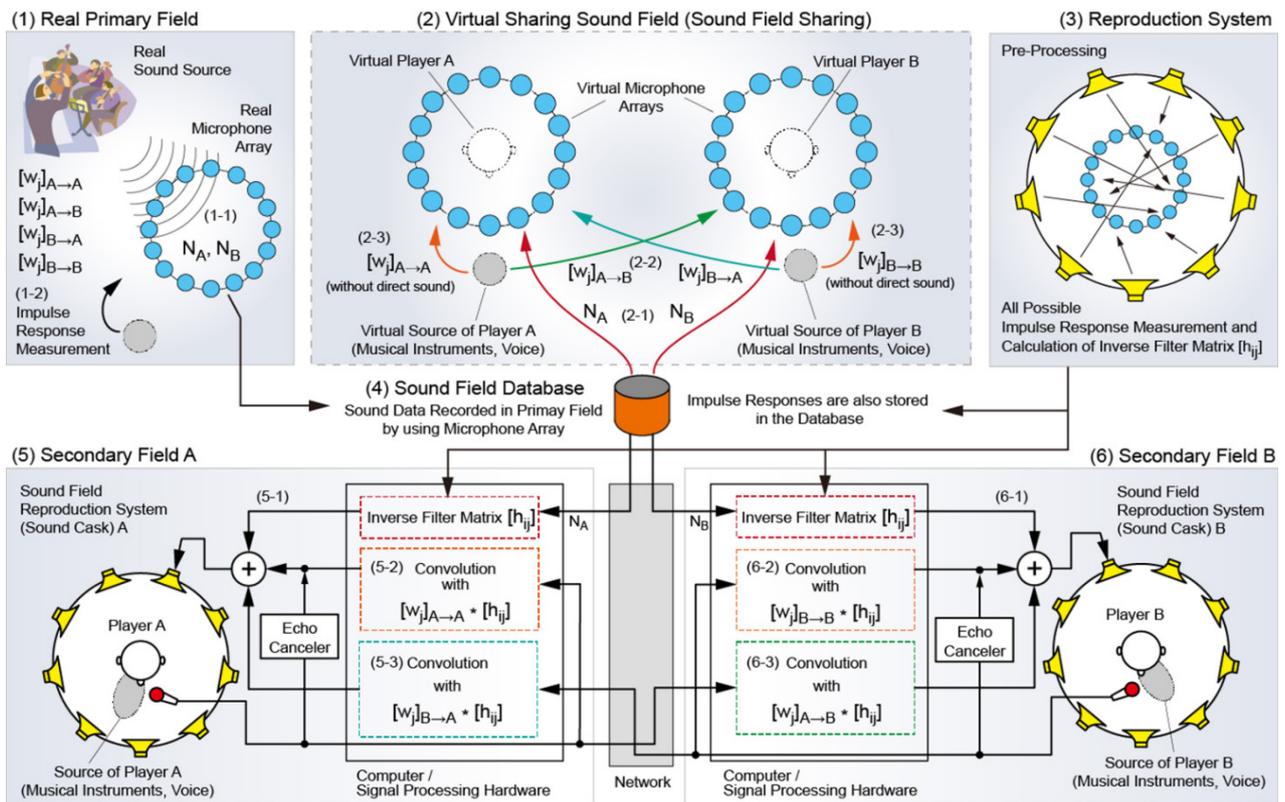


図8 聴空間共有システムの全体構成

た. システム伝達関数の最初の 20 ms を開発した FPGA によって畳み込み, 20ms 以降のシステム伝達関数を PC によって畳み込み, 両者を加算することにより 4 秒以上のインパルス応答を遅延 0.02 ms と実用上無視できるレベルの遅延で演算することが可能となった[10].

6.3 フィードバックの低減

音場共有システムには自分が発した音に残響が付加されたフィードバック経路と通信相手の空間を介するフィードバック経路が存在する. 通常の通信システムでは信号のレベルでフィードバックキャンセルを行うが, 本システムではフィードバック信号に含まれる伝達関数の信号長は長く, またチャンネル数も多いためハードウェアとして実現することが難しい. そこで楽音を收音するためのマイクロホンの位置で再生音が無音となるように逆システムを設計する手法を採用する. 実験の結果, 全帯域で 20 dB 程度, 特にハウリングが生じやすいピークにおいては 30 dB 程度のフィードバックの抑圧を確認した. また, フィードバックキャンセラを実装した音場シミュレータを建築音響指標により評価したところ, 残響時間および残響音のエネルギー比に再現性能に十分な改善がみられた. これにより楽音收音マイクを音響樽内壁面に設置することが可能になり, 演奏環境が改善された[11].

7. むすび

コミュニケーションの基盤を形成する感情共有を情報空間内で実現するためのインターフェースは, 現代の情報

通信技術におけるごちない社会関係を健全化するために必要な仕組みである. 没入型聴覚ディスプレイ“音響樽”はその解決策の一つとして有用であるが, 今後コストおよび実用面での改良が必要である.

参考文献

- [1] 伊勢, “キルヒホッフ-ヘルムホルツ積分方程式と逆システム理論に基づく音場制御の原理”, 音響学会誌, Vol. 53, pp. 706-713, 1997.
- [2] 神沼他, “周波数領域における最小ノルム解を利用した多チャンネル音場再現システムにおける逆フィルタの設計”, 音響学会誌, Vol. 57, pp. 175-183, 2001.
- [3] 李他, “正則化パラメータに着目した多チャンネル逆システムの最適化設計法”, 音響学会誌 69(6), 276-284, 2013.
- [4] 伊勢, “没入型聴覚ディスプレイ装置“音響樽”における逆システム設計法の検討”, 音講論, 2014.09.
- [5] 大石他, “話者の無意識な微細運動が聴き手の自律神経活動に及ぼす効果”, 第 37 回日本神経科学大会, 2015.09.
- [6] M. Kobayashi et. al., “The Effects of Spatialized Sounds on the Sense of Presence in Auditory Virtual Environments: A Psychological and Physiological Study”, Presence, Vol. 24, No. 2, Spring 2015, 163-174.
- [7] 小林他, “ミラーニューロンシステムの活動計測による 3 次元音場再現システムの定量的評価”, VR 学会論文誌, Vol.21(1), 2016.
- [8] 伊勢, 聴空間共有を実現する音響樽の構想”, 音講論, 2011.09.
- [9] 長尾他 “音の遅延条件がアンサンブル演奏に与える影響に関する検討”, 音講論, pp.997-998, 2012.03.
- [10] 吉田他, “FPGA を用いた低遅延畳み込み演算の実現と音場共有システムへの応用”, 音講論, 2015.09.
- [11] 河野他, “没入型聴覚ディスプレイ”音響樽”を用いた音場シミュレータにおけるフィードバックの抑圧”, 音講論, 2015.03.

A new relationship between human and sound provided by the immersive auditory display “Sound Cask”, Shiro Ise, Tokyo Denki University