

高臨場感空間音響技術の最新動向と将来展望

Trends in High-realistic Spatial Sound Technologies and Prospects for the Future

鈴木陽一 トレビーニョ ホルヘ 坂本修一

Abstract

本稿では、まず音空間技術の歴史を概観する。それを踏まえ、現在の超高臨場感三次元聴覚ディスプレイ技術を到来方向再現、バイノーラル音響、音場再現・合成と、それらのハイブリッド法に分類し、最先端技術動向を示す。次に、将来の超高臨場感空間音響技術の発展の鍵としてアクティブリスニングとインタラクティブ性を提示する。更に、それらの実現に向けた今後の三次元聴覚ディスプレイ及び、それと両輪をなす音空間情報取得技術の動向と課題を示し、インタラクティブ性実現の鍵となるコンピュータグラフィックス相当の空間音響技術（シリコンミュージックホール）実現への期待を述べ、これらがもたらす将来の可能性について若干の考察を試みる。

キーワード：高臨場感音響、聴覚ディスプレイ、音場合成、頭部伝達関数、三次元音響

1. 空間音響技術の発展

人間にとっての音の重要な役割に空間情報の伝達がある。視覚と異なり聴覚は、全方位の空間情報を知覚することが可能である。また、本小特集行場次朗氏の解説にもあるように、臨場感の重要な決定要因でもある。

本稿では、超高臨場感実現の上で映像技術とともに極めて重要な役割を果たしている空間音響技術について、これまでの発展の歩みと、現在の動向を述べ、更に将来の展望を試みたい。ではまず今に至る歩みを見ていこう。

立体音響技術が長い歴史を持つのも、音情報が豊かな空間性を有するゆえであろう。電話と電気蓄音機の発明から僅か 10 数年後の 1881 年、パリで国際電気会議とともに開催された国際電気博覧会において、電話 2 回線を用いてオペラ座の舞台公演の音を両耳で聞く公開実験が

行われ、大変評判を呼んだと言う。1920 年代には英国や米国で AM 放送 2 波を用いた 2ch ステレオ音響の実証的試みが行われている。1933 年には AT&T ベル研究所がワシントンとフィラデルフィアを電話線で結び、3ch ステレオ音響によるフィラデルフィア管弦楽団のライブ再生実験を行った。この実験の結果の解析を受けて、ステレオ音響は 2ch が標準となっていく。1949 年には磁気テープを用いたステレオ録音機が発表され、直後に商品化されている。1950 年代には定常的なステレオ放送が実現、1958 年にはステレオ LP レコードも発売された。日本でも NHK による AM 2 波を用いた実験放送が始まったのは、ステレオレコード発売に先立つ 1952 年であった。

1960 年代から 1970 年代初頭に掛けては、現在の高臨場感空間音響技術につながる技術が次々と芽生えた。

一つは多チャンネル化である。1960 年代にはゲッチンゲン大学第 3 物理学研究所において上半球に 65 のスピーカを均等配置したアレーが作られた⁽¹⁾。アナログ処理による 4ch 方式が提案されたのもこの頃である。その後、1990 年前後には日本のゼネコン各社が多チャンネルスピーカアレーを用いて、室形状から室内音響を可聴化するシステムを相次いで開発した。現在、研究用ではフランス IRCAM が 347ch⁽²⁾、東北大学では 157ch（図 1）⁽³⁾等、超多チャンネルスピーカアレーによる研究が各地で進められている。また、欧米では 100ch を優に超え

鈴木陽一 正員：フェロー 東北大学電気通信研究所人間情報システム研究部門
E-mail suzuki.yoiti@tohoku.ac.jp
トレビーニョ ホルヘ 正員 東北大学電気通信研究所人間情報システム研究部門
E-mail jorge@ais.riec.tohoku.ac.jp
坂本修一 正員：シニア会員 東北大学電気通信研究所人間情報システム研究部門
E-mail saka@ais.riec.tohoku.ac.jp
Yōiti SUZUKI, Fellow, Jorge TREVIÑO, Member, and Shuichi SAKAMOTO, Senior Member (Research Institute of Electrical Communication, Tohoku University, Sendai-shi, 980-8577 Japan).
電子情報通信学会誌 Vol.101 No.8 pp.786-792 2018 年 8 月
©電子情報通信学会 2018



図1 東北大学電気通信研究所に設置された157chスピーカアレイ 次数5次(2014年まで世界最高)の高次アンビソニックスによる音場再現, 合成が可能。

るスピーカアレイを備えた映画館が増えている。

第2は, 現在の超高臨場感音空間技術体系の中核を成すアーキテクチャの源泉となる二つの技術の誕生である。

バイノーラル音響と呼ばれる技術体系の源泉は, 1960年代にドイツを中心として形成された, 人間の音空間知覚を音源から両耳までの伝達関数, すなわち, 室伝達関数と頭部伝達関数で説明する概念に求められよう。これは更に, 音源と聴取者両耳の間の伝達関数を用いて両耳音圧を再現, 合成することにより音空間情報を提示するという概念を生み出したのである。この技術は1980年の神戸大学における実証⁽⁴⁾を経て, 次章で述べるように, 現在の三次元聴覚ディスプレイの中核的アーキテクチャの一つとなっている。

更に現在, ある領域内の音場そのものを波動性まで含めて再現あるいは合成する技術の研究開発と体系化が急速に進んでいる。その源泉は1972年のGerzonによる球面調和解析を用いた提案と考えられる⁽⁵⁾。それを1世紀ほど遡る1859年, ヘルムホルツが単一周波数の音について積分定理を定式化した。キルヒホッフはこれをキルヒホッフの積分定理として一般化し, この定理を用いてホイヘンス・フレネルの原理を理論的に導出した⁽⁶⁾。Gerzonの提案当時には, キルヒホッフの積分定理との関連は意識されていなかったが, 2000年代になると, この定理に基づく方程式(キルヒホッフ・ヘルムホルツ積分方程式, 以下KHIE)を用いた音場解析, 合成技術の一定式化と理解されるようになり, 高次アンビソニックス(HOA: high-order Ambisonics)⁽⁷⁾として発展する。今では, 1990年代に提案された波面合成法(WFS: wave field synthesis^{(8), (9)}), 境界音場制御(BoSC: boundary surface control⁽¹⁰⁾)等とともに, KHIEを基盤とするアーキテクチャとして体系化されている。

次章では, 以上のような経緯を経て成立した現在の主要な超高臨場感音空間技術動向について, 三次元聴覚ディスプレイ技術を中心に概説する。また3.では将来を見据えた技術及び応用の展望を行う。

2. 超高臨場感音空間技術の動向

前章で述べたように空間音響技術は1960~1970年代に生まれた技術が核となって進化してきた。現在の音空間情報提示技術, すなわち三次元聴覚ディスプレイ技術(3D auditory display, virtual auditory display)は表1に示すように, 多チャネルスピーカアレイを用いて音の到来方向に着目しその再現を目指した到来方向再現, 両耳音圧の再現・合成を規範とするバイノーラル音響, 波動としての音場を作り上げる音場再現・合成へと進化的発展を遂げ, 近年は更にそれらの利点を生かし得るハイブリッド技術も次々と生まれている。以下, この表に示す分類に従って概説していく。

2.1 到来方向再現

この技術の基本的な考え方は音の到来方向を再現することである。古典的には, 複数の指向性マイクロホンをあからじめ指定した方向に向けて収音し, 再生にあたっては聴取者から見てそれらのマイクロホンと同方向にあるスピーカから再生する。

その最も基本的な方式が通常の2chステレオ音響であり, その自然な拡張法は正面や後方, 上方など別の方向にスピーカを増やすことである⁽¹¹⁾。テレビ放送では, 4ch, 5.1chを経て, 現在はSHDTV用として22.2ch方式が実用化されている⁽¹²⁾。映画用としては5.1ch方式が家庭でも広く用いられ, 更にチャンネル数の多い7.1ch方式なども用いられている。

従来の到来方向再現技術では, 音響エンジニアが各チャンネルから提示される音をあらかじめミクシングしておき, 聴取時はそれをそのまま加工せずに用いている。しかし, 各種のスピーカを混ぜて用いたり, 将来の技術発展に備えるための柔軟性が求められたりするようになってきた。そこで, 音源ごとに音信号を準備しておいて, 再生時には, 再生システムのスピーカ配置に合わせて出力信号を生成する方式, すなわち, オブジェクトオーディオと呼ばれる技術が注目されつつある。

オブジェクトオーディオでは, 各オブジェクトの音信号を複数のスピーカから異なるレベルで出力すること(パニング)によって音の到来方向の設定を行う。一例として, VBAP(ベクトル形振幅パニング, vector-based amplitude panning)では, 各音源について, 設定したい方向付近の最大三つのスピーカ間のパニングにより到来方向を定めている⁽¹³⁾。映画館で用いられているDolby Atmosでは, 背景音には従来からのサラウンド

表1 高臨場感三次元聴覚ディスプレイ技術の分類と主な実現法

分類	基本概念	提案技術		特徴・実現例
到来方向再現	音の到来方向の再現、合成を目指す技術	(2ch) ステレオ音響		聴取者の左右前方に置いた2スピーカにより再生。
		サラウンド音響	5.1ch サラウンド	通常のステレオに前方中央、後方左右と、低域1chのスピーカを加えた方式。
			22.2ch サラウンド	スーパーハイビジョン(8K SHDTV)用として提案された、水平上下、3層から成る三次元サラウンド音響技術。低域は通常のステレオ配置2ch。
			Dolby Atmos	最大128ch。背景音は9.1chにより上下感も含めて表現する。音源定位位置が明確なオブジェクトは最大118個配置可能。
		オブジェクトオーディオ	ベクトル形振幅パニング(VBAP)	聴取者を取り囲むスピーカアレーを用い、各音源について、設定したい方向付近の最大三つのスピーカ間のパニングにより到来方向を定めて提示する。
バイノーラル音響	両耳における音圧の再現、合成を目指す技術	バイノーラル録音	ダミーヘッド録音	ダミーヘッドの左右外耳に仕込んだマイクロホンを用いて録音。通常は、特別な処理は行わず、そのままヘッドホン、イヤホンで再生する。
			マイクロホンアレー収音 (SENZI, VAH)	個人の全方位頭部伝達関数を用い、多チャンネルマイクロホンアレー出力からバイノーラル信号を生成。SENZIは録音後のアクティブリスニングに対応可能。
		バイノーラル合成	伝達関数合成	音源から聴取者両耳までの伝達関数(室伝達関数と頭部伝達関数)を用いてバイノーラル信号を合成し、所望の音源定位を実現する。
			レンダリング	
音場再現・合成	波動性を含む音場そのものの再現、合成を目指す技術	境界音場制御 (BoSC: boundary surface control)		キルヒホッフ・ヘルムホルツ積分方程式(KHIE)を離散化し、收音境界の外側に配置した多チャンネルスピーカアレーにより收音境界の音圧と音圧傾度を再現、合成する。
		波面合成 (WFS: wave field synthesis)		レイリー第2種積分を用い、壁面等に配置した多チャンネルラウドスピーカアレーにより半無限空間の音場を再現、合成する。欧米では映画館への導入が進んでいる。
		高次アンビソニックス (HOA: high-order Ambisonics)		聴取者位置の音場を球面調和解析に基づいて部分波の合成によって近似する。次数が高いほど部分波の数が増え性能が向上する。收音系と再生系は独立に設計できる。
ハイブリッド法	上記3種の方法の利点を生かすべく、複数を合わせて使う技術	バーチャルサラウンド (virtual surround)		サラウンド音響の各スピーカの位置から両耳までの頭部伝達関数を準備し、これを各チャンネルの音に畳み込むことによって音空間を再生する。
		バイノーラル高次アンビソニックス (binaural HOA)		頭部伝達関数の空間依存性を球面調和解析し、これを用いてHOAで收音した音場情報を部分波ごとにバーチャルサラウンドと同様の処理を行いレンダリングする。
		仮想球モデル聴覚ディスプレイ (ADVISE)		聴取者を囲む境界を設定し、その境界にKHIEを適用しレンダリングする。例:境界が球状なら室特性はHOAで求め、聴取者にはバイノーラル合成する構成が可能。
		周波数分割形併用レンダリング		低域ではHOA等の音場再現・合成を用い、聴覚の位相感度が低い高域ではVBAPなどパニングによるオブジェクト音響レンダリングを用いる混合法。

音響を用い、音源の明確な音についてはオブジェクトオーディオを用いている⁽¹⁴⁾。

2.2 バイノーラル音響

音空間知覚の最も基本的な情報は両耳への入力音である。ここで、ある音源から耳に到来する音の周波数スペクトルは、聴取者の耳介、頭、胴による散乱や共振等により、音源の位置に応じて変化する。人間はそのような周波数スペクトルの音源位置依存性を総合的に利用して音源の位置、特に方向を知覚していると考えられる⁽¹⁾。したがって、それを考慮した音を両耳に与えることにより三次元音空間知覚の操作が可能となる⁽⁴⁾。これが、バイノーラル音響の基本原理である。

バイノーラル音響の古典的实现法はダミーヘッドの外耳に仕込んだマイクロホンを用いて録音した音をそのま

まヘッドホンで(もちろんイヤホンでも)聞くことである。しかし、上述の耳入力信号の方向依存性は耳介や頭の形状に強く依存するため、個人性が大きい。

バイノーラル音響の音像定位精度、臨場感を高めるには、個人性を反映した信号処理を行う必要がある。そこで、自由音場に置いた音源から両耳までの音伝搬特性を伝達関数として与え、信号処理に用いる。これが頭部伝達関数(HRTF: head-related transfer function)である。

頭部伝達関数は従来、主に物理的計測により求めていたが、現在では耳介、頭、胴の形状を用いて数値計算を行うことにより高精度な算出が可能になっている⁽¹⁵⁾。

頭部伝達関数に加え、部屋の音伝搬特性も伝達関数として表し(室伝達関数, RTF: room transfer function)、頭部伝達関数と縦続接続する形で反映した音をヘッドホ

ンで提示するのがバイノーラルディスプレイである。クロストーク打ち消し処理を行えばスピーカを用いることも可能である（トランスオーラル法）。

近年では、個々人の頭部伝達関数を全方向についてちゅう密に与えた上で多チャンネルマイクロホンアレーによって記録した信号に反映させる方法が実用化を迎え、体系化も進んでいる⁽¹⁶⁾。我々も人頭大の球状アレーを收音系として用いる信号処理法 SENZI を提案、252ch の実時間システムを構築している⁽¹⁷⁾。ドイツでは平面上にランダムに配置したシステム VAH (virtual artificial head) が提案されている⁽¹⁸⁾。

2.3 音場再現・合成

到来方向再現では、表現できる空間性の密度に限界がある。バイノーラル音響では個人性が強い頭部伝達関数を聴取者ごとに調整する必要がある。これに対し、波動としての音場を作り上げる音場再現・合成では、音場が再現、あるいは合成されている領域に身を置くことにより音の空間性が物理的に自然に形作られ、したがって自然に空間性が知覚できることになる。

音場は波動方程式の解であり、その空間的な特性はヘルムホルツ方程式によって与えられる。現在の音場再現・合成法では、前章で述べた KHIE を何らかの方法で近似して与えている⁽¹⁹⁾。KHIE によれば、ある領域内の音圧はその領域境界の音圧と音圧傾度により完全に記述できる。（逆に、領域内の音圧分布を与えても境界条件は唯一には定まらないことに注意が必要である。）

KHIE を最も素直に応用するには、聴取者周辺にモノポール音源とダイポール音源を密に配置すればよい。これが BoSC（境界音場制御）の背景にある基本的な考え方である⁽¹⁰⁾。実システムとして用いるには、マイクロホンアレーとスピーカアレー間の整合フィルタ（matching filter）を与える必要がある。

一方、KHIE の境界を半径無限大の半球と考え、その境界平面の片側に聴取者がおり、もう片方には音源があると考え。これが WFS（波面合成法）の出発点である⁽⁹⁾。系の対称性から平面境界では KHIE のダイポール音源を考える必要がなくなり、レイリー第2種積分となる。そのため、WFS では通常のスピーカをそのまま用いることができる。

WFS のように半無限領域内の音場が再生あるいは合成できることは多くの利用者用システムとしては有用である。しかし、利用者がいない空間にも音場を形成するという無駄が生ずる。これに対し、与えられた数のスピーカで聴取者付近に音場を再生あるいは合成することを考えるのが HOA の大きな特徴である。HOA では、聴取者の位置における音場を部分波の合成によって近似する^{(5), (7)}。部分波の数を増すにつれて、音場の精度が高まり、ある精度以上の領域が広がっていく。HOA で

は、部分波ごとにスピーカアレーの入力信号を与える。この手続きは、音場の球面調和解析結果の各モードに対しスピーカアレー入力信号を与えるフィルタを定めることを意味するため、モード整合（mode matching）と呼ばれる。なお、HOA には、收音系と再生系を独立に設計できるという大きな利点がある。

2.4 ハイブリッド形三次元聴覚ディスプレイ

上記3種のアーキテクチャは、それぞれに利点、欠点を有する。そのため、近年、それぞれの「いいところ取り」を狙ったハイブリッド技術の提案が盛んに行われている。

一例がバーチャルサラウンド（virtual surround）である。この方法では、多チャンネルサラウンドの音をバイノーラルディスプレイで与える方法であり、5ch なり 22ch なりのスピーカ位置から聴取者の両耳までの頭部伝達関数を反映することによって音空間を再生する⁽²⁰⁾。この方法ではスピーカ数×2（両耳）の頭部伝達関数だけを使えばよいという利点がある。

これと似た方法がバイノーラル高次アンビソニックス（binaural HOA）である。頭部伝達関数の空間依存性を球面調和解析し、これを用いて HOA で收音した部分波ごとにバーチャルサラウンドと同様の処理を行う⁽²¹⁾。この方法の利点は、比較的少数の位置の頭部伝達関数を用いて、距離、音源の大きさ、残響等を含めた任意の音場をヘッドホンによりレンダリングできることである。

音空間情報の再生、合成では、音源の方向や距離の定位のみならず、音源の広がりや残響など部屋の音響によって決まる情報も極めて重要である。例えばバイノーラルディスプレイを用いた場合、再生については SENZI 等の適切な收音法を用いれば頭部伝達関数の個人性対応も含め、質の高い空間情報の提示が可能である。しかし、（ちょうどCG 動画像と同じように）人工的に合成したバーチャル音空間のレンダリングでは、部屋の特性を精密に与えることが簡単ではない。例えばある部屋の伝達関数として与えることを考えると、数秒にも及ぶことがある残響時間の倍程度の長さのインパルス応答の形であらゆる点について求め、その中から、室空間における音源と聴取者の絶対位置に対応したものを選んで積み込む必要があることになる。ADVISE（仮想球モデル聴覚ディスプレイ, acoustic display based on virtual sphere model）では、聴取者を囲む境界で空間を二分し、その境界について KHIE を適用することにより、その問題を解決している⁽²⁹⁾。

スピーカアレーによる音場再生精度を向上させるため、低周波数領域と高周波数領域に異なる方法を用いる手法がある（周波数分割形併用レンダリング）。波長の短い高域で高精度再生を目指す極めて多くのスピーカが必要になる。しかし 1.5 kHz 程度以上の周波数では聴

覚の位相に対する感度が大幅に低下するので、高域は旧来のパニングを用いるという割り切りを行ってスピーカ数の増加を回避する。一例を挙げれば⁽²²⁾、低域では HOA を用い、高域では VBAP を用いる。

3. 超高臨場感音空間技術の将来像

最後に、今後の三次元音響技術が目指すべき方向を少し考えてみたい。

前節までに述べたように、音響技術の高臨場感化は映像技術に先行して進んできた。しかし、時代は変わった。本小特集高田英明氏の解説にあるように、超高臨場感映像技術の高度化は目覚ましい。先進エレクトロニクス技術の駆使という意味では、音よりも映像の方が先行している時代になったと言える。

このような時代、音響技術は映像との相乗効果の実現という視点から更なる進歩が必須である。近年、古くから知られる腹話術効果とは逆に、通過反発事象や音誘導性視覚運動のように、視知覚が音情報により変容する例が報告されている⁽²³⁾。臨場感にも音と映像との間で強い相互作用が見られる。これら、音と映像が持つ、かつて考えられていたよりもはるかに強い双方向の相互作用は、是非、相乗効果の形で生かしていかなばならない。

このような超高臨場感音空間技術の実現と、社会実装の鍵は次の2点と考える。一つは、音空間知覚が前庭感覚や自己運動感覚など、聴取者の姿勢や動きの情報とのマルチモーダル感覚情報処理としての性格を持つこと（アクティブリスニング性⁽²⁴⁾）を考慮すること。そしてもう1点はバーチャル音空間とユーザのインタラクティブ性を考慮することである。

3.1 アクティブリスニングを考慮した音空間技術

音空間知覚は両耳への2チャンネル入力の情報処理に基づいているだけではない。我々は能動的に動きつつ自分が置かれた環境を理解しようとする。そこで行われる音空間知覚は前庭感覚情報、自己運動情報、視覚情報をも入力とするマルチモーダル感覚情報処理過程であると考えべきである。例えば、聴取者の動きに応じて両耳入力を与えることにより、音空間知覚の精度が大きく改善し、臨場感も強まることが明らかになっている⁽²⁵⁾。我々はこのような聴取をアクティブリスニングと呼び、その知覚過程の更なる解明と、その知覚過程に根ざしたシステム開発が、今後の高臨場感、高精細音空間システムを考えていく上で重要な要件と考える。

聴取者の動きに対応するには、二つの方法がある。一つは、バイノーラルディスプレイで意識されるようになってきた聴取者のトラッキング技術の導入である。パーソナル応用では、そのようなバイノーラルディスプレイが、システム規模、性能の両面から、しばらくは優

位性を持つだろう。かつてはトラックと組み合わせるのは特別のことであったが、現在では HMD と一体化することなどにより、解決ができるようになったことも見過ごせない。

もう一つは、ある領域で音場を実際に作ってしまう技術（音場再現・合成）である。高精度に音場が作られた領域に聴取者がいれば、当然、人間が知覚する音空間情報は自然に形成されることになる。したがって、再現された領域は元の音場と同じなので特に聴取者の動き等をシステム側で勘案しなくても聴取者の動きに対応した音情報が聴取者に提示されることになり、応用を考えていく上でも極めて柔軟性を高くできる。

この意味で、22.2ch 方式をはじめとした多チャンネルのスピーカアレーを用いたシステムでは、せっかく多くのチャンネルを使う以上、音の方向性にとどまらず波動としての音場合成の考え方を導入することが強く望まれる。その候補として最適と思われるのは HOA（高次アンビソニックス）である。HOA は音場再生・合成法の中ではシステム規模が比較的小さく、更に、他の多くの多チャンネル音空間技術とは異なり、收音系と再生系を独立に設計できるという大きな利点もある。したがって、例えば 64ch の球状アレーで收音した音を、再生系が 32

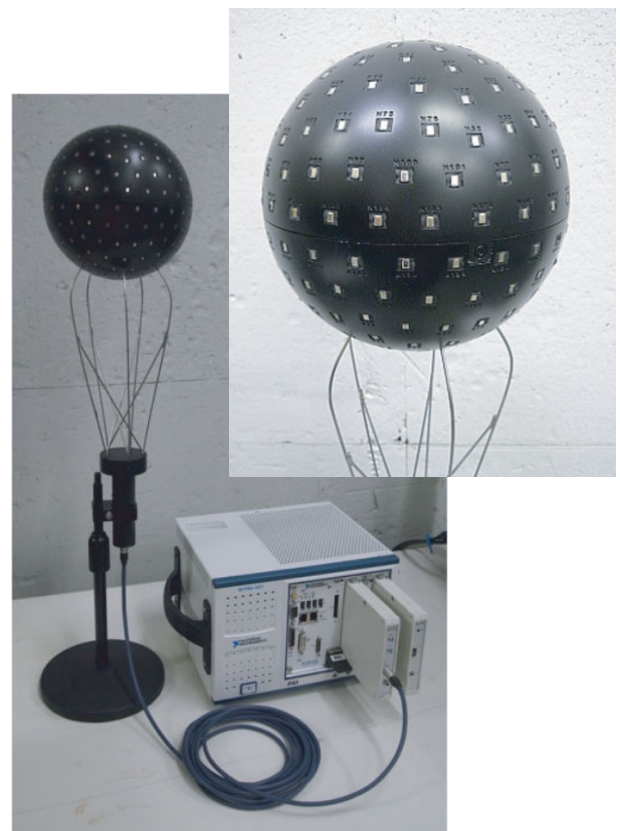


図2 実時間 SENZI システム 球は直径 17 cm（人頭大）の 252ch マイクロホンアレー。集音時でも録音後でも、聴取者の頭部の方向に応じた三次元音空間情報の提示が可能である⁽¹⁷⁾。

チャンネルでも、22チャンネルでも、それで得られる最善の精度で再生できる。固定したチャンネル数のくびきを解き放ち、様々な応用を特別の作り込みなしに実現できることにもなり、浸透が待たれる。

收音技術でもアクティブリスニングへの配慮が必要である。そのためには、音空間情報を一旦記録した後、聴取する際に聴取者の位置や動きに対応した音空間情報を合成できる技術が必要である。これが可能な收音法はほとんどなく、僅かに SENZI が、收音後、別の環境でも收音場所の音場情報を聴取者の顔の向きと回転運動に合わせて再生できる (図 2)。しかし水平移動への対応は今後の課題である。收音技術の研究は世界的に見ても多くはなく、今後の多様な発展が強く望まれる。

3.2 インタラクティブなバーチャル音空間創成技術

もう一つの方向は、VR など環境と相互作用するシステムに必須のインタラクティブ性の付与である。従来からの音空間システムは、ゲーム用などを除けば、実世界の音空間の「再生」が基本であり、視聴者の働き掛けにより音空間が変容するというようなインタラクティブ性を考慮していない。例えば、ある部屋で扉や窓を開け閉めすれば音場は変化し、聞こえてくる音は変化する。実時間の收音、再生であれば簡単に実現できるが、バーチャル空間であると簡単ではない。しかし、VR や AR など、環境とのインタラクティブ性が欠かせないシステムで用いる音響システムに必須の要件である。また、自由視聴点放送などでも、それに準じた技術が必要となる。

聴取者を取り囲む環境がインタラクティブに変化することを考えると、聴取者がいる空間 (部屋) に起因する音空間情報を実時間で表現する必要がある。しかし現在は、せいぜい初期反射音だけを簡単なシミュレーションにより与え、後部残響音はパラメトリックに合成するのが限界である⁽²⁶⁾。一方、実時間性が不要なら、ある程度の大きさの部屋であれば BEM や FDTD 等を用いて音場を高精度に計算することが既に可能になっている。将来、広い空間の音伝搬を実時間以上の速さで厳密に計算できれば、それはもうシミュレーションではなく、半導体素子の中に音場そのものがあるとみなすこともできる (シリコンミュージックホール⁽²⁷⁾)。また、人間と環境との相互作用という意味では、人間の位置や動きに応じた音環境提示に加え、人間を含めた環境全体をセンスした音環境レンダリングも重要だろう。

シリコンミュージックホールの音を実際に聴取するための研究も進んでいる⁽²⁸⁾。半導体中で計算した音場情報を、あたかもその音場にいるものとしてユーザに与えることができれば、アクティブリスニングにも自然な対応が可能となり、二重の意味で極めて高い臨場感を持った音空間の合成が可能となる⁽²⁹⁾。これは言わば、動画

像をインタラクティブにコンピュータグラフィックス (CG) で実時間合成していくことに対応する技術といえる。このような技術 (仮称: コンピュータ聴覚レンダリング, computer auditory rendering) が実現すれば、サイバー空間の実在性を大幅に向上できよう。コンピュータゲームの音の臨場感の圧倒的向上が実現するだけではなく、放送、サイバー空間上の共同作業、音楽聴取経験のありようの多様化など、極めて強いインパクトを持つに違いない。今後の更なるエレクトロニクス技術の進歩と、それを利した空間音響及びマルチモーダル信号処理法の研究の進展を心待ちにしたい。

謝辞 音空間技術史に関し有意義な意見交換を頂いた平原達也博士、安藤彰男博士に感謝する。なお 3. の論考の一部は科研費基盤研究 A (16H01736) による。

文 献

- (1) イェンス・ブラウエルト, 森本政之, 後藤敏幸, 空間音響, 鹿島出版会, 1986.
- (2) M. Noisternig, T. Carpentier, and O. Warusfel, "ESPRO 2.0—Implementation of a surrounding 350-loudspeaker array for 3D sound field reproduction," 4th Intl. Symp. on Ambisonics and Spherical Acoust. and 25th AES UK Conf., no. 14, March 2012.
- (3) T. Okamoto, B. FG Katz, M. Noisternig, Y. Iwaya, and Y. Suzuki, "Implementation of real-time room auralization using a surrounding 157 loudspeaker array," Principles and Applications on Spatial Hearing, pp. 373-382, world Scientific, March 2011.
- (4) M. Morimoto and Y. Ando, "On the simulation of sound localization," J. Acoust. Soc. Jpn (E), vol. 1, no. 3, pp. 167-174, May 1980.
- (5) M.A. Gerzon, "Periphony: With-height sound reproduction," J. Audio Eng. Soc., vol. 21, no. 1, pp. 2-10, Feb. 1973.
- (6) M. Born and E. Wolf, "The integral theorem of Kirchhoff," Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, 7th ed., 8.3.1, Cambridge Univ. Press., Cambridge, 1999.
- (7) M.A. Poletti, "Three-dimensional surround sound systems based on spherical harmonics," J. Audio Eng. Soc., vol. 53, no. 11, pp. 1004-1025, Nov. 2005.
- (8) A.J. Berkhout, D. de Vries, and P. Vogel, "Wave front synthesis: a new direction in electroacoustics," Audio Engineering Society Convention 93rd, no. F-2, San Francisco, USA, Oct. 1992.
- (9) A.J. Berkhout, D. de Vries, and P. Vogel, "Acoustic control by wave field synthesis," J. Acoust. Soc. Am., vol. 93, no. 5, pp. 2764-2778, May 1993.
- (10) 伊勢史郎, "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," 音響誌, vol. 53, no. 9, pp. 706-713, Sept. 1997.
- (11) M. Gerzon, "Panpot laws for multispeaker stereo," 92nd Conv. Audio Eng. Soc., no. 4PS2.04, March 1992.
- (12) K. Hamasaki, T. Nishiguchi, R. Okumura, Y. Nakayama, and A. Ando, "A 22.2 multichannel sound system for ultrahigh-definition TV (UHDTV)," J. SMPTE Motion Imaging, vol. 117, no. 3, pp. 40-49, April 2008.
- (13) V. Pulkki, "Compensating displacement of amplitude-panned virtual sources," 22th Int. Conf. Audio Eng. Soc., pp. 186-195, June 2002.
- (14) Dolby white paper, "Dolby atmos next-generation audio for cinema," Issue 3, 2014.
- (15) M. Otani and S. Ise, "Fast calculation system specialized for head-related transfer function based on boundary element method," J. Acoust. Soc. Am., vol. 119, no. 5, pp. 2589-2598, May 2006.
- (16) C.D. Salvador, S. Sakamoto, J. Trevino, and Y. Suzuki, "Design theory for binaural synthesis: Combining microphone array recordings and head-related transfer function datasets," Acoust. Sci. Technol., vol. 38,

- no. 2, pp. 51-62, March 2017.
- (17) S. Sakamoto, S. Hongo, T. Okamoto, Y. Iwaya, and Y. Suzuki, "Sound-space recording and binaural presentation system based on a 252-channel microphone array," *Acoustical Sci. Technol.*, vol. 36, no. 6, pp. 516-526, 2015.
- (18) E. Rasumow, M. Blau, S. Doclo, M. Hansen, S. van de Par, D. Püschel, and V. Mellert, "Least squares versus non-linear cost functions for a virtual artificial head," *J. Acoust. Soc. Am.*, vol. 133, no. 5, pp. 3525-3525, May 2013.
- (19) 安藤彰男, "高臨場感音響技術とその理論," *信学FR誌*, vol. 3, no. 4, pp. 33-46, Nov. 2010.
- (20) K. Matsui, Y. Nakayama, and K. Hamasaki, "Development of binaural headphone processor for 22.2 multichannel Sound," *Annual Tech. Conf. & Exhibit. SMPTE*, pp. 1-10, Oct. 2015.
- (21) M. Noisternig, A. Sontacchi, T. Musil, and R. Höldrich, "A 3D ambisonic based binaural sound reproduction system," *24th Int. Conf. Audio Eng. Soc.*, no. 1, June 2003.
- (22) E. Rasumow, M. Blau, M. Hansen, S. Doclo, S. van de Par, V. Mellert, and D. Püschel, "Robustness of virtual artificial head topologies with respect to microphone positioning," *Proc. Forum Acusticum*, pp. 2251-2256, July 2011.
- (23) 鈴木陽一, "空間と時間," *音と時間*, 難波精一郎(編), 9章, コロナ社, 東京, 2015.
- (24) Y. Suzuki, "Auditory displays and microphone arrays for active listening," *Key note address, 40th Conf. Spatial Audio, Tokyo, Japan*, Oct. 2010.
- (25) Y. Suzuki, T. Okamoto, J. Trevino, Z.-L. Cui, Y. Iwaya, S. Sakamoto, and M. Otani, "3D spatial sound systems compatible with human's active listening to realize rich high-level kansei information," *Interdisciplinary information sci.*, vol. 18, no. 2, 71-82, 2012.
- (26) V. Valimaki, J.D. Parker, L. Savioja, J.O. Smith, and J.S. Abel, "Fifty years of artificial reverberation," *IEEE Trans. Audio, Speech, and Language Proc.*, vol. 20, no. 5, pp. 1421-1448, July 2012.
- (27) 鈴木陽一, "息をのむ臨場感をリビングルームで," *ICT Dreams*, pp. 112-113, 電気通信工学振興会, 2007.

- (28) 飯野慎吾, 松下 翔, 土屋隆生, "フレネル・キルヒホッフ回折公式を用いたバイノーラル音空間レンダリングの高速化の検討," *信学技報*, EA2017-32, pp. 33-38, Aug. 2017.
- (29) S. Takane, Y. Suzuki, T. Miyajima, and T. Sone, "A new theory for high definition virtual acoustic display named ADVISE," *Acoust. sci. technol.*, vol. 24, no. 5, pp. 276-283, Nov. 2003.

(平成 30 年 3 月 8 日受付 平成 30 年 3 月 27 日最終受付)



すずき よういち
鈴木 陽一 (正員:フェロー)

1976 東北大・工・電気卒. 1981 工博. 1999 から同電気通信研究所教授. 聴覚とマルチモーダル知覚情報処理, 音信号処理研究に従事. 日本音響学会会長, URCF 会長, VR 学会理事等を歴任. 2004 米国音響学会フェロー, FIT 2005, 2013 船井ベストペーパー賞など各受賞.



トリ ビニョーホルヘ (正員)

2000 モンテレイ工科大・工・電気卒. 2014 東北大・博士 (情報科学). 2015 から同電気通信研究所助教. 物理音響, 音場収録・再現, 調和解析, 音信号処理に関する研究に従事. 日本音響学会粟屋潔学術奨励賞 (2014) 受賞. 2016-2017, 本会応用音響研究専門委員会幹事補佐.



さかもと しゅういち
坂本 修一 (正員:シニア会員)

1995 東北大・工・情報卒. 2004 博士 (工学). 2011 から同電気通信研究所准教授. 聴覚・多感覚情報処理の解明及びシステム応用に関する研究に従事. FIT 2013 船井ベストペーパー賞, 日本音響学会佐藤論文賞 (2017) など各受賞.