

解説と展望

60 年が教えるもの*

山崎芳男 (早稲田大学)**

43.38. -p, Si, Vk

まえがき

音響学会も誕生して60年、ここに至るまでのみちのりは必ずしも平坦なものではなく、多くの先輩方のご苦勞は並大抵のものではなかったと承っているだけに感慨もまたひとしおである。

学会が50周年を迎えた際に学会誌の総目次づくりを担当させていただいた。60周年を期に学会誌がデータベース化されネットワークやCD-ROMを介して簡単に過去の会誌に接することのできる環境が整備されたことはまことに喜ばしい次第である。創刊当時の学会誌や研究発表は驚くほど内容が新鮮であり、毎年必ず研究室全員で読ませていただいている論文も幾つかある。

オールラウンドプレーヤ

限られた技術環境の下ですばらしい研究が行われていたということ、会員が“音”の専門家で音に対する研究を幅広くにこなされていた点、いや場合によっては音以外の分野でも意欲的に研究を進めておられた方々のお名前を多数目にする点に更に感銘を受ける。

現在でも音響学会には専門分野で活躍されると同時に幅広く音全般に目を向けて研究されている会員が多く、これはとかく専門が細分化される傾向の強い昨今、本学会の素晴らしい特徴と言えるのではないかと。

研究発表にしても一会場で行われていた。会員が増え研究が盛んになり、研究会を中心に各専門分野の研究が行われ、春秋の研究発表会も同一会場での研究発表が不可能となり専門ごとに並列発表が行われるようになった。

現在の音響学会の会員に専門とは尋ねれば躊躇なく音声、電気音響、超音波、建築音響等々答え

が返ってくるに違いない。このこと自体は歓迎すべきことではあろうが、研究発表会でも何か工夫して全員参加の発表形式が一部でもとれないものかと考えたりもする。

時空を超えて

先駆者達が限られた条件の下で行った素晴らしい実験、測定装置なども数多く掲載されている。当時の限られた条件の下で考えられ行われた実験や測定をヒントに現在の技術と組み合わせるといえば時空を超えた刺激に触発された新たな研究ができるのではないかと。例えばRayleigh板を使った粒子速度の直接測定、残念ながら私はRayleigh板を使った測定の経験はないので現在の技術で新「Rayleigh板」を作ってみたいし、また不安定で風などに弱いという欠点を現在のサーボ技術とレーザを組み合わせると克服してみたくもなる。「専門分野」の枠を超えた研究の重要性を痛感させられ始めた研究も幾つかある。遮音等室内音響評価への情報理論の適用もその一例である。

静かで暗い夜を取り戻そう

ところで幾つかの不便はあったものの、かつての東ヨーロッパを訪れてほっとした経験をお持ちの方も多いのではないかと。そこには意識して残したのではないにしても、私たちが忘れ去ってしまった落ち着きを垣間見ることができた。特に夜、明るく騒然とした都会の夜になれてしまっている我々は当初こそその暗さと静けさに戸惑うものの、必要最小限の光と静寂からなる落ち着きはなんとも快適で心休まるものであった。

どこで行われているか外からではほとんど分からない控えめなホールや劇場でコンサートやオペラを何度も堪能したものである。そこへあの激変、あの暗くて静かな夜がどうなってしまったか、いささか心配でもある。

* Things learned from passed 60 years.

** Yoshio Yamasaki (Waseda University, Tokyo, 169)

日本、特に都会は騒々しく夜は眩しいほど明るい、そこで聞こえるよう見えるようにとますますやかましく眩しくなるという悪循環を繰り返している。そのうえ音や光で厚化粧しようとする試みすらある。

室内は確かに冷暖房の設備等で快適になった面もあるが、空調騒音や換気扇の音、隣接家屋との遮音等音環境は多の問題を抱えている。先日も市場に出入りするトラックに電気自動車を試験的に導入してみたところ、静かすぎて人がよけてくれないのでトラックらしい音のする装置を考えて欲しいという笑えぬ話すら現実にあった。

東欧ばかりではなく地球上にいたるところで静かな街、暗い夜が急速に失われつつある。実際自然豊かな地域から日本にみえた方々の優れていた視覚や聴覚能力がほんの数か月の内に我々都会人並みになってしまうそうである。本来人間は生きていくために優れた感覚、視覚や嗅覚などを持っていたはずである。

井伏鱒二氏が関東大震災前には荻窪あたりでも耳にすることのできた東京港の船の汽笛が震災後ぱったり聞こえなくなったと何かに書かれていたのを読んだ記憶がある。恐らく震災後音の伝搬経路への家屋等の建設による遮音と騒音レベルの上昇によるマスキングに起因するものであろうが、東京ですら数十年前までは静かであったことは確かである。

音の大きさに敏感になる

音環境の悪化に伴いとりわけ大きな音に鈍感になってしまっているのではないか。例えばオーディオ、音質にうるさい人は多いが難聴になるほど大きな音で聴く人も多く音量には篤くほど無頓着である。少なくともクラシックやジャズは高忠実度再生という観点からも耳の位置近傍の音圧レベルを原音場に合わせて聴くのが基本であるはずである。

録音した街のざわめきなど再生音量を実空間のそれと合わせると突然リアルに聞えてくるものである。我々はこのような身近なところから音圧レベルにもっと神経質になって良いのではないか。絶対音感を持つ人は少なくないが、音を専門にしている者として“絶対音圧感”を養うべく努力している。

“文明”人の耳に頼ってよいものか

最近 MPEG (Moving Picture Encoding Expert Group) や MD (Mini Disk) など聴覚や視覚の特性を積極的に利用した音や映像の高効率符号化方法が数多く提案され一部はすでに実用に供され始めている。これらの規格・標準化は当然日・米・欧等のいわゆる先進国主導で行われている。劣悪な音環境の下知らず知らずのうちに聴覚や視覚が衰えてしまった我々の耳や目を使って高効率符号化を評価したり、方式を決定してしまっただけのよいものであろうか。新しいシステムこそ人間が本来持っていた優れた感性を満足する仕様でなければならぬ。

“文明”という名のもとに劣化してしまった聴覚や視覚能力をどこまで取り戻すことができるかは疑問であるが、できる限り静かな街そして暗い夜、人間的な生活を取り戻す努力を始められないものであろうか。音の専門家の努力により文明の本来の聴覚・視覚能力を十分満足する符号化方法などが実現できればと期待している。

高速1ビット処理

時間に伴い振幅が連続して変化する信号をデジタル化するには、時間方向の離散化を行う標本化と振幅方向の離散化を行う標本化の操作が必要である。このうち標本化については標本化定理を満足しているすなわち信号の帯域幅が標本化周波数(標本間隔の逆数)の1/2以下である限り信号の情報は失われないが、量子化に伴う量子化雑音は不可避である。

実用に供されているデジタルオーディオシステムでは CD (Compact Disk), DAT (Digital Audio Taperecorder) や衛星放送等の広帯域音響信号の符号化には数10kHzで標本化周波数と12~20ビットの量子化ビット数が使われている。一般に標本化周波数が帯域を決め、量子化ビット数がダイナミックレンジを決定すると捉えられがちであり、両者は密接な関係を持っており個々に論じられるものではない。

量子化雑音は入力信号の振幅が大きい場合には入力信号と無相関な量子化ステップ幅 Δ に一樣に分布する電力 $\Delta^2/12$ の周波数成分の平坦な白色性の雑音となる。一方、入力信号の振幅が小さ

い場合には量子化雑音はその総電力は $\Delta^2/12$ と変わらないものの、入力信号と相関の強い聴感的に不快な歪となってしまふ。

ところが、量子化ステップ幅 Δ あるいはその整数倍に一様分布するディザと呼ばれる一種の雑音である確率変数を量子化に先立ち入力信号に加算し、量子化器出力から同じ確率変数をデジタル減算すると、量子化雑音は入力振幅の如何にかかわらず大振幅入力のとくと同様に入力信号と無相関な電力 $\Delta^2/12$ の白色性の雑音とすることができる¹⁾。従って標準化周波数を高くすれば信号帯域内の量子化雑音は減少し、原理的には1ビット量子化で100dBのダイナミックレンジを確保することも可能であるが、標準化周波数は3 GHz、1ビットであるから伝送容量も1秒当たり30億ビットにも達してしまい実現不可能である。

情報理論の先駆者であるシャノンは1940年代後半に信号の平均情報量が伝送容量よりも小さければ無歪伝送可能な符号化方法が存在することを明らかにしている。これは標準化周波数と量子化ビット数が個々にではなく、むしろその積である情報量が重要な意味を持っていることを示している。しかし現実のシステムの構築にあたって標準化周波数と量子化ビット数の決定は一筋縄でいくものではないし、現行の標準化周波数や量子化ビット数が必ずしも人間の聴覚特性に馴染むものではないことは明らかである。

著者らは十数年来、標準化周波数と量子化ビット数の組み合わせの選択が困難であるのならば、いっその片方の極限である高速標準化1ビット量子化で信号の情報量と人間の聴覚特性に見合った符号化方式が可能なのではないかと様々な検討を加え、量子化雑音のスペクトル分布を制御することにより高速1ビット符号化が可能であることを明らかにした²⁾。

現行の48kHz標準化16ビット量子化と等しい伝送容量である768kHz標準化1ビット量子化で100kHz程度までの信号の伝送が可能ハードウェアシステムを実現し音響計測やスタジオ録音等実用に供している。

ところで前述のように聴覚特性を利用した各種の高効率符号化が実用に供され始めているが、その出発点が現行のCD、DATと同じ方式におかれているのには疑問がある。ここで述べた高速1ビ

ット符号化方式などより広帯域の情報源符号化を行えば、伝送容量を節約したうえ現行のシステムよりも質の高い符号化も可能なはずである。

また、音場の能動制御では処理による遅延が問題となるが本方式では遅延のほとんどない高速処理が期待できる。1ビットであるので乗算が単純なゲート回路で実現できるのでLMS (Least Mean Square) 等アナログ信号との乗算が著しく簡素化できるという特徴を持っている。

バーチャルリアリティ (VR) と音

人間にとって当然視覚は重要な情報摂取手段である、しかし目で捉えられる範囲は前方に限られ、それほど広いものではない。従って人間は元来後ろからの情報摂取は多くは聴覚、気配というものに頼っている。実際聴覚は後ろに対して驚くほど鋭い識別能力を持っている。特に後ろからの音場の伝送を考えた場合情報が大切であるということに常に認識する必要がある。

実際これは簡単な定位の実験をすれば明らかである。すなわち眼をつぶって前方に人が立って喋った場合定位は曖昧であるが、後方で喋ってもらえば非常に的確に位置や距離を判断できる。

いずれにしても人間はたった二つの耳で様々な方向から来る音を聞き分けている。信じられないことであり、また1本のマイクロホンがあれば原理的にもあり得ないことであるが、人間は片耳でも慣れればかなり正確な定位が可能である。これは永い経験で方向別の伝送特性を正確に記憶しているからであろう。

人間の聴覚能力は驚くべきものである。最近バーチャルリアリティという言葉が頻りに耳にするようになったが、音の分野では早くからVRを志向していた。

蓄音機の目指すところはVR

レコード再生装置は古くは蓄音機と呼ばれ文字どおり音を蓄えることを目的としていた。昭和初期は単にポータブルといえば蓄音機のことであったように、1950年代はハイファイ (Hi-Fi) がレコード再生装置のあるいは高級ラジオの呼称として定着していた。Hi-FiはHigh-Fidelity高忠実度の略であって、何に対して忠実かは議論はあろうが、演奏会場の音をそのまま自宅に持ち込もうと

いう願望の込められた的確な呼称である。

1958年に45/45ステレオレコードが登場すると立体音響時代を迎えHi-Fiステレオ,ステレオが音響再生装置の一般名称になったことは皆様よく御存知のとおりである。ステレオの登場と共に臨場感という言葉が盛んに使われ始めた。この臨場感こそVRを最も的確に捉えた日本語表現と言えるのではないか。レコードの歴史は正にバーチャルリアリティの追求であった。

CDの恩返し

1982年にCD (Compact Disk) が誕生し、レコードはまたたくまにデジタル時代を迎えることとなる。CDの誕生の陰には中島平太郎元会長をはじめ多くの日本音響学会会員の並々ならぬ貢献があった。音の記録媒体として登場したCDが今やCD-ROMやビデオCDとして広くコンピュータや映像の分野でなくてはならぬ記録媒体となり、マルチメディアの中枢に位置している。

これはコンピュータや映像の技術を導入して誕生した音のCDが立派に成長して多くの人々に恩返ししている姿である。音に携わるものとして少しばかり誇りに感じてよいのではないか。

DVDはデジタルビデオディスクではない

日本音響学会が60周年を迎える今年1996年にDVD (Digital Versatile Disk, デジタル万能ディスク) が誕生しようとしている。DVDは万能化したCDの次世代メディアとして多方面からの様々な要求と提案を踏まえ登場しようとしている文字どおりのデジタル万能ディスクの総称である。

DVDがしばしばデジタルビデオディスクと訳されているのを目にするが、DVDはあくまでもデジタル万能ディスクであり、デジタルビデオディスクはDVD-ROM, DVD-RAM, DVD-Audio, DVDファミリーの一つでDVD-DあるいはDVD-Videoと記述されるべきものである。いずれにしてもDVD登場による音の新たな発展を期待したいものである。

電話は定着したVR

電話もまたTelephoneが示すとおりテレイグジスタンス (Tele-existence) を志向したシステム

であることは間違いない。電話の通話品質は決して高いものではないが、電話はそれをVRと意識するかいなかは別として多くの人々に自然に空間移動の道具として利用されている。

最近新幹線内や自動車運転中の電話が社会問題化しているが、これはマルチメディア時代に高品質とは言い難いシングルメディアがいと簡単に人間を現実空間とは全く次元の異なる世界に引き込んでいく証でもある。

パイノーラルとステレオ

ところで人間の耳が二つあることに着目した電話回線を使った遠距離立体音響再生、すなわち音場の移動に関する研究はアメリカのベル研究所を中心に1920年代からすでに始められていた。1933年4月27日ワシントンとフィラデルフィア間を電話線で結んでストコフスキー指揮フィラデルフィア交響楽団の演奏などを使った有名な3チャンネル立体音響の実験が音響物理学者らによって行われた。この実験によりステレオ再生やバーチャルリアリティに通じる貴重な数々の成果が得られている³⁾⁻⁷⁾。

日本においても伊藤毅、牧田康雄らにより立体再生の詳細な理論検討や実験が行われている^{8),9)}。当時の立体音響再生こそ音空間の移動技術すなわちVRそのものと言える。

放送における立体音響再生の歴史も古くアメリカでは日本で放送が開始された1925年にコネチカット州ニューヘブーンにおいてAM放送2局による実験放送が行われた記録がある。

日本では1952年12月5, 6, 7の3日間通常放送の終了後ラジオNHK東京第1, 第2放送を使ってステレオの試験放送が行われた。翌年から全国ネットの実験が何度か行われ1954年4月から日曜日の定時番組に「立体音楽堂」が登場した。更にFENを含めた3チャンネルステレオの実験すら行われている。

また、1954年には東京の民放3局ラジオ東京、文化放送、ニッポン放送による3次元ステレオの実験が行われた。その後、東京、大阪、福岡、札幌などで民間放送によるステレオの定時番組も定着した。東京地区では1958年9月から文化放送とニッポン放送によるイブニングステレオは、平日の夕刻、毎日15分ずつ放送されていた。家庭用

音響再生装置にはAMラジオが2組組み込まれるという時代が数年続いた。

FM放送では1958年頃からクロスビー（FM-FM）方式のステレオ放送が実験され、1988年アメリカでは現行のパイロットトーン方式がFCCの標準方式として決定され、日本では1963年12月24日日本放送が開始された。FMステレオ放送の本格化と共にステレオ放送の中心はFMに移行した。

1波によるAMステレオ放送についてもアメリカでは古くから研究が行われていた。日本でも1962年にTBSに実験局の免許があり実験放送を行った。その後紆余曲折があり1991年AMステレオ放送が開始された。今となっては2波によるAMステレオ放送をご存じの方はほとんどないのではないかと。

テレビジョンの音声多重放送も1960年代から研究が行われ1982年12月7日から本放送が開始され、現在ではほとんどの局がステレオ化されている。また、1984年5月12日から直接衛星放送DBS（Direct Broadcasting Satellite）の実験放送が開始された。DBSの音声放送はAB2モードがあるが、両モードともデジタル放送であるがBモードでは2チャンネルAモードでは4チャンネルステレオ放送が可能である。

音空間の伝送

ダミーヘッドホン收音・ヘッドホン受聴あるいはトランスオーラル系によるスピーカ再生はいずれも原音場の空間情報は伝えることができる。しかしこれらは空間でなく2点のみの制御であり人の移動、頭の回転には対応できない。

現行のステレオシステムではたとえチャンネル数を増やしても正確な3次元音空間の伝送は到底不可能である。正確に伝送するには伊勢史郎らが提案するキルヒホフ積分に基づく方法しかない。

しかし膨大な数の收音伝送再生系を必要とし現段階では実用化は困難である。波面合成の理論に基づき数十点で收音・伝送し、同じ波面合成理論により数十個のスピーカにより再生する方法が検討されている。実現可能な数の收音・伝送再生系で実用上十分な3次元音空間の伝送が可能である。音源が片側に偏在する遠隔会議などの場合に

は制御する面が1面ですむので特に有効である。

更に遠隔会議や講演のように音源の数が一つあるいは少数の場合には近接4点法で音源位置を実時間で算出することにより、音源の数だけの伝送路と音源位置情報だけで空間音情報の收音・伝送が可能となる¹²⁾。近接4点法とは著者らが開発した同一平面上にない近接して配置したマイクロホンで收音し、その四つの信号の時間構造の違いに着目し地震の震源探査や3角測量と同様な原理で直接音及び反射音の位置を把握する音空間測定法である¹³⁾。

モノラルでVR実現

視覚のVRは一般にHMDや偏光眼鏡等を使って2眼視で行われているが、1987年9月のテレビジョン学会誌に、木内雄二氏のテレビジョンを片目で見れば自然で臨場感があるという大変興味深い寄稿があった¹⁴⁾。実際、実験をしてみると分かるが、ある大きさのTV画面や写真を片目で見ると非常にリアルである。これは1台のカメラで撮ったものを両目で見るから無理があるのであって、片目で見れば非常に自然であるという理にかなったものである。

耳についても前述のように片耳でもかなりの情報を得ることができる。案外割り切って通常の画面とモノラル集音ソースにより片目片耳で実用上十分なVRが実現できるかも知れない。

モノラルでなくとも現行のステレオのスピーカ配置は少なくとも両耳の間隔くらいに近づけた方が受聴範囲も広がり自然な臨場感が得られるのではないかと。電機大学の浜田晴夫らは小型スピーカを接して配置するSD（Stereo Dipole）を提案されている¹⁵⁾。

情報の産直

映画やテレビの映像には監督やプロデューサなどの制作者の意思が反映されており、それが価値を持っているのに対して、VRネットワークによる空間伝送ではむしろ監督やプロデューサなどの意思ではなく自分の意思で見たいものを見、聴きたいものを聴くことに意味があると言えるのではないかと。

いわば情報の産地直送である。例えばオペラハウスの何か所かにオペラグラスのようなビデオカ

メラとパイノーラル收音装置を設置すれば、オペラハウスの好きな席でオペラグラスを通して鑑賞するのとほぼ同等の体験をすることは現在の技術でも十分可能である。

ユネスコは世界遺産条約に基づいて、現在94か国411か所を遺産リストに指定している。日本でも姫路城と法隆寺地域の仏教建造物が文化遺産に、屋久島と白神山地が自然遺産に1993年登録された。遺産指定によって、特に自然遺産は人が押し寄せて破壊されるといった新たな懸念も生じている。そこで遺産の自然や景観を損なわない範囲で、カメラ、マイクロホン、温度、風の全方向センサ等を設置することにより電話や衛星回線を通していながらにして、遺跡や動植物の観察が可能となる。現在ユネスコや各国の大学、研究機関と協力してVRによる世界遺産の多数か所で同時体験可能なシステムの構築の検討を始めている。

世界遺産に限らず、例えばシャンゼリゼや銀座通り、世界各地の海岸やスキー場、幹線道路や鉄道の駅等を結ぶネットワークを構築すれば、自分の目と耳で的確な情報を把握することが可能となる。

む す び

音響学会が誕生して60年、入会してからの30年、私の周辺で最も変化したのはコンピュータ環境だろう。ハードウェアでは冷暖房完備の巨大な部屋を占拠していたと同じ以上の能力を持つコンピュータが持ち運びできる時代になった。ソフトウェアでは1960年代半ばのCooleyとTukeyによるFFTのアルゴリズムは衝撃的であった。FFTは正にコロンブスの卵的な発想で当時のコンピュータの能力には荷の重かった音響信号の調和解析が十分実用的な時間で可能になり信号処理は飛躍的に発展したのは周知のとおりである。FFTは分析点数を素数の積(一般に最も小さな素数である2のべき乗が使われている)に限定することにより計算を高速化する手法であるが、FFTがあまりにも便利な道具であるがゆえ、その制約を忘れてFFT分析では無理な処理にも使われている例を散見する。

最近のコンピュータの処理能力の飛躍的

によりウィナーが1950年代に提唱した調和解析の原点とも言える一般調和解析ですら、時間はかかるもののパーソナルコンピュータで実行することが可能となった¹⁸⁾。一般調和解析は観測窓内で原変形から残差成分が最小となる正弦波を抽出し、残差成分に対して同様の操作を繰り返すという単純明快な解析手法であり、信号処理の新たな飛躍を感じさせるものがある。

文 献

- 1) 山崎芳男, "広帯域音響信号の量子化への大振幅ディザの適用," 音響学会誌 39, 452-462 (1983).
- 2) 山崎芳男, 太田弘毅, 西川明成, 野間政利, 飯塚 秀, "広帯域音響信号の高速標本化1ビット処理," 信学技報 EA93-102 (1991).
- 3) H. Fletcher, "Basic requirement," Symposium on Wire Transmission of Music and Reproduction in Auditory Perspective, E. E. 53, 9-11 (1934).
- 4) J. C. Steinberg and W. Snow, "Physical factors," Symposium on Wire Transmission of Music and Reproduction in Auditory Perspective, E. E. 53, 12-17 (1934).
- 5) E. C. Wempe and A. L. Thuras, "Loudspeakers and microphones," Symposium on Wire Transmission of Music and Reproduction in Auditory Perspective, E. E. 53, 17-24 (1931).
- 6) E. O. Scriven, "Amplifiers," Symposium on Wire Transmission of Music and Reproduction in Auditory Perspective, E. E. 53, 25-28 (1934).
- 7) H. F. Afel, R. W. Cjesnut and R. F. Mills, "Transmission lines," Symposium on Wire Transmission of Music and Reproduction in Auditory Perspective, E. E. 53, 216-219 (1934).
- 8) 伊藤 毅, 音響工学原論(下)(コロナ社, 東京, 1957).
- 9) Y. Makita, "On the direction localization of sound in a stereophonic sound field," Eur. Broadcast. Union Rev. 37A, 102-108 (1962).
- 10) 伊勢史郎, "広範囲の音場再現についての研究(1),(2)" 音講論集, 479-482 (1993. 10).
- 11) A. J. Berkhout, "Acoustic control by wave field-synthesis," J. Acoust. Soc. Am. 93, 979-995 (1988).
- 12) 工藤修二, 及川靖広, 山崎芳男, 片桐竜二, "近接4点法により求められた反射音位置に基づく波面合成," 音講論集, 803-804 (1996. 3).
- 13) Y. Yamasaki and T. Itow, "Measurement of spatial information of sound fields by closely located four point microphone method," J. Acoust. Soc. Jpn. (E) 10, 101-109 (1989).
- 14) 木内雄二, "テレビ画像の立体感を強める一方法," テレビ誌 41, 834-835 (1994).
- 15) 浜田晴夫, 得能宏則, "新しい音場再生方式(Stereo Dipole)に関する研究," 信学技報 EA96-2 (1996).
- 16) J. W. Cooley and H. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," Math. Comput. 19, 297 (1965).
- 17) N. Winer, The Fourier and Certain of Its Applications (Dover, New York, 1958).
- 18) 牛山 聡, 東山三樹夫, 飯塚昌弘, 平田能睦, "一般調和解析による波形分析," 信学技報 EA93-103, 39-44 (1994).