

インタフェースとしての音を考える

山崎 芳男

早稲田大学理工学総合研究センター教授

奥の深い音

「百聞は一見にしかず」というわれるように、音は映像に比べて従属的なものとら

えられがちです。しかし、人間にとって音は映像や文字情報に優るとも劣らぬ重要なコミュニケーション手段であり、インタフェースの道具です。音の場合、耳だけでは

なく口というインタフェース手段があることが大きな特徴です。

ところで、新学期の講義では奇妙な光景に遭遇します。何人かの学生、特に留学生が天井を向いて講義を聞いているのです。私は今ステージ上で話していますが、目を閉じて聞いていただくとハース効果、先行音定位といって、自分の耳に真っ先に到達するもっとも近いスピーカの位置を音源と感ずます。あとからくる音がある程度大きくても、最初に音のやってくる位置を音源と感ずます。大学の大きな講義室では、多くの場合、天井に小さなスピーカが分散配置されており、ワイヤレスマイクを使って講義するので、当然、最初に近くの天井スピーカから音が聞こえることとなります。したがって、礼儀正しい留学生は天井をみながら先生の声を聞くことになるわけです。先生が教室のなかを移動しても、音は移動しません。

普段、意識していませんが、目でとらえられる範囲は前方にかぎられ、それほど広いものではありません。人間は元来、うしろからの情報の多くは聴覚、気配というものに頼っていました。人間はたった2つの耳でさまざまな音源の方向や距離を巧みに聞きわけています。ちょっと信じられないことですが、人間は片耳をふさいでも、しばらくするとかなり正確な音源位置の特定が可能となります。これは、経験によってさまざまな方向や距離からくる音の特徴を正確に記憶するからです。

音のバ - チャルリアリティ

バーチャルリアリティという言葉が頻繁に耳にするようになりましたが、音の分野

では古くから蓄音機、ステレオと、音環境の忠実な伝送を志向してきました。電話も立派なバーチャルリアリティです。新幹線や電車で使用を自粛するアナウンスがたびたび行われるように、隣りの人と会話をしているときには周りに気を配りながら会話をしますが、電話をかけた途端に別の世界に没入するので大声をだしてしまうのです。もちろん、電話の音質はよくありません。特に、日本の携帯電話の音質は、耳をおおわんばかりで社会問題化しています。それですら没入するので、電話は立派なバーチャルリアリティといえます。

実際、電話回線を使って、昭和初期にはワシントンとフィラデルフィアの間で、当時のかなり立派な拡声装置を使ってオーケストラの3次元伝送の実験を行っています。日本でも戦後間もないころ、ラジオ受信機がプアな時代に広がりをもった音を楽しもうと、AM2波を使ったステレオ放送が行われていました。

音のないナイアガラの滝

今(1998年11月末から12月にかけて)京都の国際会議場で世界遺産会議が開催されています。私どもは、ユネスコの“New technology for culture”というプロジェクトに協力して、世界遺産を含む自然や文化財をあるがままにどこにいても体験できるように記録・伝送するシステムの研究開発を進めています。今回はデジタルVTRに記録した4方向の映像とともに、独自の高速1bit方式でマルチチャンネル録音した3次元音場記録を紹介します。

図1は京都の会議でも紹介しているナイアガラの滝をデジタルVTRで記録した映

像です。映像はとてもきれいです、音がないとそれこそ別世界の出来事です。ここでは小型のスピーカしか用意できませんでしたので、あの滝の音量をそのままというわけにはいきませんが、4方向からの音を加えると、ナイアガラの滝がぐっと身近な現実のものとなります。音はこのように普段あまり意識しませんが、とても奥の深いものなのです。

青空間の記録と伝送

音の情報は左右の耳からはいつてくるので、自分の耳の位置にマイクロホンをつけたり、自分そっくりの人形、あるいは単純化して人間の顔の特徴を少し誇張したものの耳に相当する位置にマイクロホンをつけて、同じ音量でヘッドホンで再生すると、その場にいるのと同じ音空間を体験することができます。ただし、これはあくまでも耳の位置、2点の伝送にすぎません。頭を動かしたり動さ回することは不可能です。複数の人の場合、同じシステムを複数用意しなければなりません。

3次元音空間の記録と伝送に関しては、



図1 ナイアガラの滝

理論的な裏づけがあります。キルヒホッフの積分公式を使えば、3次元空間の正確な記録・伝送が可能です。キルヒホッフの積分公式とは、われわれを囲む閉曲面、この会場の場合、天井・壁・床といった面上の音圧と法線方向、面に直角な方向の空気の粒子速度を目的の音場のそれと同じにすれば、所望の3次元音場が再現できるというものです。動き回ることも、多人数で同時に楽しむことも可能です。

このキルヒホッフの積分に基づく音響のバーチャルリアリティは、残念ながら今のところ実現は不可能といわざるを得ません。なぜなら、人間の耳で聞くことのできる音の周波数は20Hzから20kHz程度、子どもには30kHz程度まで聞こえます。20kHzは音速が340m/sなので波長は1.7cm、したがって大人の聞くことのできる20kHz帯域のバーチャルリアリティを実現するには、標本化定理から制御する点は少なくとも8.5mm間隔にしなければなりません。8.5mm間隔にマイクロホンやスピーカ制御系を用意しようとする、その数はこの会場ではおそらく1億以上になり、残念ながら実現不可能です。

しかし、われわれはこれに何とか挑戦しようと努力しています。ひとつは、多数の超音波素子を使って超音波をAM変調して映像とともにスクリーンにあてる、いわば映像と音のプロジェクタです。スクリーンの反射によ

り音圧が上がり生ずる非線形現象を使ってAM超音波を検渡して音にするわけです。

もうひとつは、1cm間隔でボイスコイルをプリントした数十cm角のフレキシブル基盤を振動板として、5mm角の多数のネオジウム磁石と組合わせたマイクロセル型スピーカです。この方式により1cm間隔のスピーカを初めて実現しました。先ほど述べたように制御チャネルが膨大になる

るので、部屋全体をスピーカにするのはまだ不可能ですが、1m x 2m程度のスピーカはすでに完成しています。

この3月のバーチャルリアリティ学会の総会で、オペラ歌手の秋葉京子さんとピアニストの梅沢直子さんにお越しいただいてこのスピーカと実時間畳み込み装置を使ったバーチャルオペラコンサートを開きました(図2)。そのバーチャルオペラコンサートをここで再現してみましょ。学会で実験したときは、本物の肉声とピアニストでしたが、今回は、当日録音した音を使わせていただきます。

拡声装置を切ってみましょ。これがこの会場自体の響きです。秋葉さんにアベマリアを歌っていただきましょ。秋葉さんの声は舞台においたこのスピーカ、梅沢さんのピアノはこのスピーカからだします。

次にこれが「東京駅丸の内のコンコー



図2 オペラ歌手の秋葉京子さん(中央に立っている女性)とピアニストの梅沢直子さん(左下に座っている女性)

ス」の響きです。4方向からの響きを東京駅で測定したインパルス応答と、私の声をこの実時間畳み込み装置で畳み込み、4隅においた4つのスピーカから再生しています。次は伊豆の天城トンネルです。トンネル特有の響きを感じられますね。これは高知の鍾乳洞「龍河洞」の響きです。これは愛知の芸術文化会館のコンサートホールの響きです。これはパリ郊外ロワイヨールの修道院の響きです。

たった4つのスピーカでも驚くほど響きが変わります。私の声は響きによって話し方までかわってしまいます。秋葉さんの歌も今日は録音なのでかわりませんでした。バーチャルオペラコンサートでは音場により興味深い変化がありました。

新しい音響空間の創出

この実時間畳み込み装置は8チャンネル

分、数秒のインパルス応答との実時間畳み込みが可能です。十数年前に研究を始めて現在は3代目ですが、その演算量はスーパーコンピュータ以上です。

このシステムは実際の建物の設計にも使っています。早稲田大学で最近700人数室をひとつ、350人数室を3つつくることになりました。それらの教室の設計に、この研究成果を反映させ、先生の声が口元から自然に聞こえるように、音響設計しました。

例えば、語学ヒアリングの入試を実施しようとしたとき、各教室で、また教室の各席で等しい条件で明瞭に聞こえることが重要です。新しくできた校舎だけが素晴らしい音環境では困るわけです。どの教室でも同じ条件にする必要があります。各席に小さなスピーカをおいて制御する方法もありますが、費用とメンテナンスを考えると現実的ではありません。

そこで提案したのは、教卓側の壁を全面このスピーカにして、平面波が伝播するような音場をつくる方法です。こうすると教室のどこでも同じ音が得られます。反対側の壁は完全吸音にする必要があります。

黒板やスクリーンをおく必要から全面は無理ですが、新築された14号館の700人数室には黒板の両側に40個ずつこの平面スピーカを設置しました。また、講演卓にスピーカを埋め込み、パネルディスカッション用にパネリストの机にマイクロホンとスピーカを組み込んで、話している位置から音ができるように工夫しています。これでも動き回って講義する状態では先生の位置から聞こえないので、最近、チョッキ型の平面スピーカも試作してみました。

文明人の感性に頼るな

最近のコンサート会場では、立派なタワー型のスピーカを左右に積んで大音量で拡声しています。音量が大きすぎることで、どこで演奏しているかわからなくなることが難点です。バーチャルリアリティにおけるインタフェースでは3次元空間を実現すると同時に、そこから音が自然に聞こえることが重要です。拡声などしないほうがよいのかもしれませんが、ある程度多くの人数で聞くにはやはり拡声は不可欠です。メガホンやトランジスタメガホンの原点にかえるような、小さな声や、楽器の音を補助するような観点からの3次元音場の制御も、今後の課題ではないかと思います。

音は周波数は低く、帯域も狭いので扱いやすいようですが、どうしてどうしてたいへん難しいものです。衣食足りて音を知るではないですが、バーチャルリアリティや符号化は、最初はどうしても映像やコンピュータ情報が中心になりがちです。

私は現在、高能率符号化MPEG Audioの標準化作業に携わっています。その会合でよく“文明人”の感性に頼るのではなく人間が本来もっていた優れた感性を満足する符号化をしなければいけないと主張しています。新しい規格をきめる場合、先進国が中心にならざるを得ません。われわれには眼鏡もありますし、拡声器も補聴器もあるため、耳や目がスポイルされてしまっています。テレビの音を離れたところにあるステレオで聞いても、最初の1～2分は不自然に感じますが、すぐ慣れてしまいます。多分、この会場で私の声は拡声器を通じて聞こえ、目をつぶれば私の口の位置から聞

こえませんが、何ら不自然はお感じになっていないでしょう。本来、人間は目や耳がよくなければ生きていけなかったのです。新しいインタフェースや規格もまたわれわ

れの感覚だけではなく、本来、人間のもっていた鋭い感性を満足するものであってほしいと考えています。
