

## デジタルオーディオの技術的可能性

山崎 芳男

早稲田大学理工学総合研究センター

### 1. まえがき

デジタルの起源は人類が数という概念を使い始めたころにあり、のろしや旗による通信という形で情報のデジタル伝送も古くから行われてきた。17世紀に数値解析が発展してデジタル信号処理の基礎が出来上がり、19世紀にはモールス符号によるデジタル通信、電信が実用化された。

しかしデジタル信号処理技術が今日の隆盛を究めるに至った原動力は、1940、50年代のC. E. Shannonの業績に代表される情報理論の発展と電子計算機の実用化と1970年代以降の半導体技術の飛躍的発展にあるといえる。

デジタルオーディオの実用化の現在までの道のりには大きな山が幾つかあった。第一は、1960年代の電話中継回線のデジタル化、放送局やレコード制作用のデジタルテープレコーダなど莫大な投資をしてアナログ伝送路の隘路をデジタル化することにより改善した「黎明期」である。

第二は、手軽に良い音が楽しめるようになったCD、MD、パーソナルコンピュータ、家庭用デジタルVTRなどLSI化と大量生産による身近な分野へのデジタル技術の「浸透期」である。

以前はデジタルにより改善はされるが膨大な帯域、高価な機器を必要としていた。しかるに、最近では有限な電波の有効利用を目的に、いわば「省資源」の観点から、デジタル化による携帯電話や放送の多チャンネル化が始まっている。

20世紀後半に登場し、当初は畏敬の念をもって迎えられたデジタル技術も、省資源や安く丈夫だからといった観点から導入されるようになり新たな時代を迎えようとしている。

### 2. シャノンの美しい符号化定理

シャノンは信号の単位時間当たりのエントロピー(Entropy)すなわち平均情報量より大きな伝送容量を持つ伝送系を使えば、あいまい度を限りなく小さくする符号化法が存在するという美しい定理を導いている。このシャノンの情報理論の第1定理はアナログ信号の無ひずみ伝送の可能性を示唆している。アナログ信号が無限の情報量をもつならば、伝送容量が有限なデジタル伝送路ではひずみのない伝送は不可能である。しかるに、われわれが情報伝達に用いる信号の情報量は一般に有限である。そもそも情報が意味を持つということはその情報量が有限の値であることに他ならない。適確な符号化を施せば、伝送容量の有限な現実の機器を使用してもデジタル化により劣化のない伝送・処理が期待できることになる。

#### 2.1 アナログかデジタルか

一般に信号の伝送路には必ず雑音が存在する。これは伝送路の伝送容量が有限であることを示している。従ってアナログ信号をそのまま伝送すると信号には多かれ少なかれ雑音、誤差が加わり劣化する。アナログでは伝送過程で加わった雑音やひずみは佗復不可能であるが、デジタルでは数字が読めさえすればもとどおりに清書することが可能である。さらにたとえば数字の合計をメモしておけばどれか一つが読めなくとも修復することが可能となる。これがデジタル信号処理の大きな特長、誤り訂正機能である。

このようにデジタル伝送では伝送容量の有限な現実の機器を使用しても劣化のない伝送・処理が

期待できる。従って信号処理へのデジタル信号処理導入の採否は、アナログ処理による信号劣化と、デジタル化の際生じる信号劣化との比較により決定すべきものといえる。

また、一般にアナログでは質の高い処理を行うには熟練を要するのに村し、デジタル信号処理では誰が何時行なっても同じ結果が期待できるという利点もある。

## 2.2 デジタル時計ではなく水晶時計？

よくアナログ録音のLPレコードを「針のついたアナログ時計」に、デジタル録音を「数字表示のデジタル時計」に例えた説明を見受ける。しかしLPであろうがCDであろうが再生される音は同じように空気の気圧変化として耳で捉えられている。むしろデジタル録音されたCDは水晶共振素子をデジタル分周して時計の針を動かしている水晶時計に例えるべきである。

振り子時計も振り子がギアを動かしており、音の陥時計もテンブが時を刻んでおりデジタル的である。砂時計だって粒子が駆動源だし、日時計も光子に頼るデジタル装置である。

こうして見ると時計は案外昔からデジタル的でありアナログ時計に相当する時計を探すのは難しいくらいである。強いてアナログ時計を探すなら電源周波数に同期したモーターで駆動された電気時計くらいであろう。

デジタル表示の時計はさしずめMIDIデータによる音楽といったところであろうか。

## 2.3 アナログより身近な存在

20世紀は量子力学の時代ともいわれている。前の例ばかりでなく意外に純粋なアナログ信号、アナログ現象を探すのは難しい。われわれは技術の流れに従い半ば無意識、理論体系も技術開発もアナログで培った考え方、方法をデジタルに置き換えるというやり方で進んできた。しかしむしろ理論も設計技術もまずデジタルで構築し、アナログはデジタルの特別な場合、離散化を細かくした極限として捉える方が自然なのではないか。実際幅が0で高さ

が無限大で面積が1であるデルタ関数などというものを考えるより、幅も高さも1で面積1の単位標本の方がずっとわかりやすいし、積分は加算に、微分は引き算に置き換えた方がずっと親しみやすくなるであろう。

## 2.4 劣化しないという欠点

信号をデジタル化すると、たとえ雑音のある伝送路を使っても信号劣化のない伝送が可能である。これはデジタルの大きな特徴であるが、ソフトウェアに携わる側にある種の戸惑いがあったことも事実である。それはアナログでは多かれ少なかれ劣化した信号が聴取者の耳に届いていたが、デジタル化すると信号は劣化することなくそのまま届く。また、デジタルでは劣化のないコピーが可能である。法律的な観点からの議論も盛んに行われており、デジタル透かしなどの導入も検討されている。

## 3. 原点に戻ろう The simpler the better

### 3.1 アナログの良さをデジタルで

#### 高速1bit信号処理

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るが、標本化周波数と量子化ビット数の積である伝送容量、信号伝送速度が重要な意味を持っており個々に論じられるものではない。

量子化ステップ  $d$  に比較して入力信号の振幅が大きい場合あるいは適確なディザ導入されている場合には量子化雑音は入力信号と無相関な標本化周波数の  $1/2$  の帯域に一樣に分布する電力  $d^2/12$  の周波数成分の平坦な白色性の雑音となる。

したがって標本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、1ビット量子化で20kHz帯域で100dBのダイナミックレンジを確保することも原理的には可能である。しかしこの場合標本化周波数は3GHzにも達してしまい実現不可能である。

我々は十数年来、標本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その片方の極限である高速標本化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てきた。1974年には高次差分PCMにより標準化周波数が1/6以下のダイナミックレンジを拡大することが可能であることを明らかにし、1979年には量子化器を帰還ループの中に設けその順方向伝送特性 $H(s)$ と帰還特性 $H(z)$ を適当に選び、量子化雑音のスペクトルを制御することにより広帯域音響信号を扱う1ビット量子化が可能であることを明らかにした。

この方式はアナログのオペアンプや数個のトランジスタあるいは真空管と比較器で構成することができる。比較器以外はアナログ回路であり出力の1ビット信号はれっきとしたデジタル信号でありながら入力信号のスペクトルがそのまま存在するという単純なシステムである。通常、復調に必要なマルチビットのDA変換器は不要であり、原信号の帯域の成分がそのまま復調アナログ信号となる。

現行の48kHz標準化16bit量子化と等しい伝送容量である768kHz標準化1ビット量子化で直流から20kHzにわたり100dBのダイナミックレンジを確保したうへ100kHz程度までの信号の伝送が可能である。我々の研究室では1985年来この方式によってマルチチャンネル録音用ハードウェアを数台試作し音響計測やスタジオ録音等実用に供している。

### 3.2 単純明快な一般化調和解析

広く使われているフーリエ解析は、周期的で調和的な信号に対する基礎理論でFFTなど広く利用されているが、信号は観測時間外では観測時間内の波形が周期的に繰り返されると仮定しているので、観測区間によって多数の異なる周波数分析が抽出されるという欠点がある。

一般化調和解析は観測区間内で原波形から残差エネルギーが最小となる正弦波を抽出し、残差成分に同様の処理を繰り返すという単純明快な解析方法である。従って定常的でない僅かな周波数変動に対しても正確な周波数成分の抽出が可能であり、人間の優れた聴覚を模した処理が可能であるなどの特徴が存在する。多くの特徴を持つ一般化調和解析ではあるが、膨大な演算時間を必要とするので実

用の信号の分析に使われることはほとんどなかった。

1994年平田、東山らは効率的な一般調和解析の算出手法を提案し楽器音の分析などを試みている。

### 3.3 “文明人”の感性に頼らない

最近MPEG(Moving Picture Encoding Expert Group)など人間の聴覚や視覚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、既にビデオCDやDVD-D等の符号化に使われ始めている。これらの規格・標準化作業は当然いわゆる先進国主導で行われている。

眼鏡や拡声器が当たり前の生活をして、知らず知らずのうちに聴覚や視覚が衰えてしまった「文明人」の目や耳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つてよいものであろうか。地球上には目や耳の優れた人々がたくさんいるはずである。新しいシステムこそこれらの人の力を結集して人間が本来持っていた優れた感性を満足する仕様にしなければならない。

次世代オーディオが人々の英知と感性を結集して、優れた聴覚・視覚能力をも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。

### 3.4 前向きな高効率符号化

高効率符号化には信号の情報量に着目してその冗長性を除去することにより、無ひずみで伝送要領の節約を図る方法と、人間の感覚を利用してひずみを出来るだけ感知されがたく工夫した圧縮とに大別される。一般にデジタル化した信号を圧縮するという側面のみ捉えられがちであるが、高効率符号化はむしろ、伝送路は節約したうへ質の向上を図るという前向きな技術として捉えるべきものである。MPEGオーディオは、マスキングや臨界帯域により、耳で検知することが出来ない音を伝送の対象から除外することにより伝送量を節約する高効率符号化である。

著者らは前述の高速1bit音響信号処理を提案し、もっとも高効率符号化、時間軸圧伸・ピッチ変換等の信号処理についても検討を加えてきた。窓幅11

msで一般化調和解析を試みたところ100周波数成分を抽出し、それらを合成することにより聴感上、原音と違いのない音が得られた。

聴覚特性にはマスキングや臨界帯域がある。次に聴覚に着目し一般化調和解析により高速1bit信号区間毎にデータレートを変化させる不等長符号化を検討した。分析区間により違いはあるものの平均で約26本程度、伝送容量約107kbit/sで聴感上原音と変わらない高能率符号化が可能であることが確認された。

### 3.5 約束ごとは最小限に

デジタルオーディオでは標準化周波数や量子化特性に始まり各種の高能率符号化、伝送形式等少々約束事が多すぎるきらいがある。これがシステムの自由度を制約する結果につながっているのではないか。

標準化周波数や量子化特性、エンファシス特性などどんな方式で書かれているかを媒体ごとに記しておき、このメモがきに従って再生系を設定するようなシステムがあっても良いはずである。

アナログレコードでは演奏時間の短い場合には溝間隔を大きく取ることによりダイナミックレンジの広い記録が可能であった。しかるに現行CDでは30分収録しようが74分収録しようが伝送品質は同じである。自由度を与えさえすれば帯域やダイナミックレンジの広い信号を記録することも、マルチチャンネルのソースをつくることも、BGMを24時間1枚に収容したディスクを作ることさえ可能になる。

## 4. ネットワーク時代のオーディオ

### 4.1 点から空間

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、例えチャンネル数を増やしても正確な三次元空間音場の伝送は到底不

可能である。

正確に伝送するには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実現は不可能である。

我々は波面合成と近接4点法により得られた仮想音源分布に基づき数十の伝送系とスピーカにより再生する方法を検討している。

なお近接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4つの信号の時間構造の違いに着目し地震の震源探査や三角測量と同様な原理で直接音および反射音の位置を把握する音空間測定法である。

### 4.2 音のオブジェクト階層符号化

さらに独奏や弦楽四重奏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。再生は前述の方法と同様の方法、すなわち波面合成により現実的な数の制御系・スピーカで三次元音空間を再構築することが可能である。

### 4.3 音場の産直

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、それが価値をもっているのに対して、オーディオにおける音場の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場合もある。

いわば情報の産地直送である。例えばオペラハウスの何箇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とおして鑑賞するのとは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。

### 4.4 パソコン購入のきっかけは音楽

最近学生が大学によく来るようになった。講義が急に魅力的になったわけではない。大学では自由にインターネットが使えるからである。

可能である。

正確に伝送するにはキルヒホフ積分に基づく方法しかない。しかし膨大な数の収音伝送再生系を必要とし実現は不可能である。

我々は波面合成と近接4点法により得られた仮想音源分布に基づき数十の伝送系とスピーカにより再生する方法を検討している。

なお近接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収音し、その4つの信号の時間構造の違いに着目し地震の震源探査や三角測量と同様な原理で直接音および反射音の位置を把握する音空間測定法である。

#### 4.2 音のオブジェクト階層符号化

さらに独奏や弦楽四重奏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収音・伝送が可能となる。再生は前述の方法と同様の方法、すなわち波面合成により現実的な数の制御系・スピーカで三次元音空間を再構築することが可能である。

#### 4.3 音場の産直

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、それが価値をもっているのに対して、オーディオにおける音場の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場合もある。

いわば情報の産地直送である。例えばオペラハウスの何箇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収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とお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。

#### 4.4 パソコン購入のきっかけは音楽

最近学生が大学によく来るようになった。講義が急に魅力的になったわけではない。大学では自由にインターネットが使えるからである。

ネットワーク上に情報を登録しておくことにより、どこにいても利用できるばかりか自分の意志によってそれを他人に開放することも可能である。

むしろソフトはネットワークで供給され近い将来CDやDVDに同じ物を複製して売るといった形態はなくなるかもしれない。同じデジタルデータを何百万枚も複製するのは資源の無駄以外の何者でもないからである。ネットワークを介してアクセスするたびに課金すれば済むことである。

この分では早晚自動車も、場合によっては家も個人所有しない、会社もネットワーク上にといった時代が来るかもしれない。学校や大学の授業、学会、新聞や雑誌、全国規模の放送局なども近い将来大きく形を変えることは間違いないだろう。

#### 4.6 コンテンツは自然に出来上がる

彼らは決して情報を受け取るばかりではない。自分の周りの情報をごく自然に発信している。いや発信というよりはまわりの出来事を使われるかどうかはともかく、取り合えずネットワークに登録しておくのである。この積み重ねがネットワークの価値を高めることを本能的に知っているに違いない。

よく言われることであるがネットワーク上では大企業も個人も平等である。むしろ決済や著作権といった制約と無縁な無数の個の結集が途方もない力を生む現場を見せられている。すでに多くの中学生や小学生の力さえも結集されているはずである。

世の中で現在作り出されている膨大なマスメディアに加えて全人類の私信にいたるまでの天文学的な情報を収容することなど不可能だという指摘もあるが、現在の技術動向からみれば記憶容量は十分過ぎるくらい供給されよう。

#### 4.7 奥の深い音

百聞は一見にしかずなどと言われ、マルチメディアにおいて音は映像やコンピュータデータに比べて従属的なものと捕らえられがちであるが、実は音は映像や文字情報に優るとも劣らぬ重要な情報伝達手段なのである。

普段意識しないが、目で捉えられる範囲は前方

に限られ、それ程広いものではない。したが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。

人間はたった2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本のマイクロホンでは原理的に不可能であるが、人間は片耳でも慣れればかなり正確な音源位置の特定が可能である。これは永い経験でさまざまな方向や距離から来る音の特徴を正確に記憶しているからである。

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、音の分野では古くは蓄音機、ステレオと音環境の忠実な伝送を志向していた。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た。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、これは音だけによるコミュニケーションがいと簡単に人間を現実空間とはまったく次元の異なる世界に引き込んでいた証である。

衣食足りて...ではないが、マルチメディアには快適な音環境が不可欠である。

#### 5. むすび デジタルがとれて一人前

音響信号は実に10オクターブにおよぶ帯域を100dBにおよぶダイナミックレンジで扱うまだまだ未解決の問題の多い分野である。しかしながら周波数も比較的安く取り扱いやすい、あるいは技術はとうに完成しておりいわゆるハイテクとは無縁の存在のように見なされがちである。

現実にはまさに正反村、1960年代後半に実剛ヒされたレーザーをさっそく取り入れてCDを作り上げてしまったり、理論技術段階であった誤り訂正のアルゴリズムをLSIに押し込んでCDやDATあるいは衛星放送などとデジタルシステムを家庭にまで持ち込んだりと何とも果敢な頼もしい新しい技術に対して貪欲な分野であり、ひょっとすると超伝導なども何らかの形でオーディオ分野に顔を出すかもしれない。

#### 文 献

- (1) C.E.Shannon, "A mathematical theory of communication," Bell Syst. Tech J., 27, (1948).
- (2) L.G.Roverts, "Picture coding using pseudo-random noise," IRE Trans. IT-8 (2), 145-154 (1962. 2).
- (3) N.S.Jayant and L.R.Rabiner, "The application of dither to the quantization of speech signals," B.S.T.J. 5 1(6), 129 1-1 104 (1972. 7-8).
- (4) D.E.L.Shorter and J.R.Chew, "Application of pulse code modulation to sound signal distribution in a broadcast network," Proc. Inst. Elec. Eng. 119 (10), 1442-1448 (1972).
- (5) 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 1. 9, 452-462 (1981).
- (6) 山崎芳男, 伊藤毅, "広帯域音響信号の高忠実度PCM記録および伝送について," テレビジョン学会, 録画研資 11-2 (1975).
- (7) 山崎芳男, 太田弘毅, 西川明成, 野間政利, 飯塚秀幸, "広帯域音響信号の高速標本化1bit処理," 信学技報, EA93-102 (1994).
- (8) 松谷康之, 内村国治, 岩田, "多段量子化雑音抑圧(MASH)方式16ビットCMOS A/D変換LSI," 信学会集積回路研資 ICD 87-52, 7-12 (1987).
- (9) 杉山道則, 大内康裕, 天田武, 山崎芳男, "一般調和解析による音響信号の分離," 音響講論, 607-608 (1996).
- (10) 林雅尚, 山崎芳男, 及川靖広, "一般調和解析を用いた高速1bit音響信号の時間軸圧伸・ピッチ返還," 音響講論, 603-604 (1996).
- (11) 伊勢史郎, "広範囲の音場再現についての研究(1) キルヒホフ積分公式に基づいて," 音響講論, 479-480 (1993).
- (12) 伊勢史郎, "広範囲の音場再現についての研究(2) 多チャンネルステレオフォニクス," 音響講論, 481-482 (1993).
- (11) 山崎芳男 "これからの音声符号化技術," JAS コンファレンス '90 予稿集, 145-148 (1990).