

## ハイレゾに迫る技術 「良い音」を伝える技術の革新

日本オーディオ協会理事・編集委員/NTT エレクトロニクス株式会社

遠藤 真

日本オーディオ協会では「ハイレゾ」を定義するとともに、ロゴマークを制定してハイレゾ普及を図るとともに、あらためて「良い音」とは何かという本質的な課題にも取り組んできました<sup>[1]</sup>。この取り組みでは、ハイレゾの定義を満たせば良い音になるわけではなく、良い音の追求はさらに続けられるべき、ということが強調されました。折りしも昨今は「ハイレゾ」の定義を満たさないまでも「良い音」を追求してハイレゾに迫る新しい技術がいくつか現れてきています。また「ハイレゾ」の元の語である「レゾリューション(分解能)」についても、周波数以外に「良い音」を求める方向、軸があります。本稿では、「良い音」とほぼ同義に使われてきた「音質」に関する技術史の一端から、ハイレゾに迫ろうとする新しい技術を位置づける観点を考えようとするものです。

### 1.はじめに

これまで音楽を楽しむには、音の収録と再生の段階があり、それらの中間にはレコード、テープ、CDなどの媒体に音楽を記録して流通させるという3段階がありました。そして、それぞれの段階において「良い音」を追求してきたのがオーディオの歴史だと言えます。さて、近年のオーディオにおいては、これらでは欠けているものがあるのにお気づきでしょうか？

### 2.音質と情報量

かつて音質については、(1)電話並み、(2)AMラジオ並み、(3)FMラジオ並み、という言い方がありました。これらに後に(4)CD並み、というのが加わり、(5)ハイレゾ、ということになるのですが、これらを音声の周波数帯域幅で表すと、およそ次のようになります。

- (1) 電話:約 3.4kHz
- (2) AMラジオ:約 7.5kHz
- (3) FMラジオ:約 15kHz
- (4) CD:約 22kHz
- (5) ハイレゾ:40kHz 以上

ここで(1)から(3)はアナログ技術です。(4)CDはデジタル技術が前提で、符号化方式は16ビットのLPCM(リニアPCM)です。アナログTVの音声は(3)FMラジオ並みの音質でした。さらに技術が進んだ現在の地上波デジタルTV音声は(4)CD並みを目指したのですが、それまでの(1)から(5)とは異なり、圧縮技術を施してビットレート(1秒当たりの情報量に相当)を絞っています。これは電波利用上の制約から、映像に多くのビットレートを割り当てるためです。その結果、情報量はCDの数分の一以下になっています。情報量はファイルにすればファイルサイズに相当し、通信ではビットレートに相当します。

たとえばAMラジオの音声を16ビットのLPCMでデジタル化したとすると、ビットレートはモノラル(1ch)で240kbps(kビット/秒)になります。一方、地上波デジタルTV音声のビットレートはステレオ(2ch)で同程度なので、チャンネルあたりの情報量はAMラジオの半分です。しかし、

地上波デジタル TV の音質は(2)AM ラジオ並み以下ではなく、より高音質の(3)FM ラジオ並みに聴こえます。これは前述したロッシェー圧縮符号化方式のおかげです。

このようにデジタル技術を使うと、ビットレートは必ずしも音質に比例しないということになります。もうお分かりいただけたかと思いますが、初めに述べた「音楽を楽しむため」に必要な(1)収録、(2)再生、(3)記録の他に必要な段階とは(4)通信です。音楽は媒体に記録されて流通するのと同様に、放送やネット配信などの通信を通じて流通しています。

これまで見たように電話からハイレゾに至る「良い音」の追求はもっぱら周波数帯域幅拡大を目指してきました。その一方で通信上のビットレートの制約をデジタル音声圧縮技術によって、少ない情報量で元の音に近い音質になるように工夫が重ねられて来た歴史があります。音質は通信で交わされる情報量が指標の一つにはなりますが、情報量だけでは音質を評価しにくい、もっと言えば周波数帯域幅だけでは「良い音」を評価できない段階に至っていると言えないでしょうか？

### 3. デジタル音声圧縮技術と通信

#### 3.1. デジタル音声圧縮技術の分類

デジタル音声圧縮技術は(1)ロッシェー圧縮符号化方式と(2)ロスレス圧縮符号化方式に 2 分されます。(1)は情報量を削って圧縮するものです。この一種の AAC という方式が前述した地上波デジタル TV 音声で使われています。(2)はファイル圧縮のように、圧縮しても情報が失われずに完全に元に戻せる方式で、ファイルサイズやビットレートは小さくできるものです。4K、8K 衛星デジタル放送の ARIB 規格にもなった国際標準の MPEG-4 ALS<sup>[2]</sup> <sup>[3]</sup>、ハイレゾ音楽で普及した FLAC<sup>[4]</sup>などがあります。ハイレゾの定義は原音のまま(非圧縮)の LPCM、1 ビット符号化あるいは圧縮してもロスレス圧縮符号化したものが前提です。

#### 3.2. ロッシェー圧縮符号化の観点と通信の観点

ロスレス圧縮符号化方式に対してロッシェー圧縮符号化方式はその名の通り情報の一部が失われているのですが、それぞれどのような観点で圧縮しているのでしょうか？まず電話です。アナログ時代からデジタルになって PHS、携帯と代を重ねてきました。初めは会話が主眼ですから、(1)言葉が聞き分けられること、(2)話している人の特徴が分かること、すなわち誰が話しているのか、会話で重要な言葉の抑揚まである程度伝わること、が求められました。その結果が 3.4kHz 帯域でした。デジタル化されても帯域はほぼそのままでした。ちなみにこれは LPCM14 ビット相当を非線形の 8 ビット PCM で符号化して、ビットレートは 64kbps でした<sup>[5]</sup>。その後電話では、より自然な音声求められて周波数帯域幅を倍の 7kHz にしたり、音声品質を維持したままで携帯を狭い電波帯域にしたり、会話だけではなく保留メロディーや BGM のような音楽もそこそこの音質で伝わるようにしたり、と圧縮技術が進んで来ています。

その一方で特に IP 電話になってからは、音質以外に以下のような通信特有の観点でも技術が論じられています。これらは送受信部や通信経路を含めて考えると、ロスレス、ロッシェーに拘わらず通信に共通です。

- (1)ビットレート
- (2)ノイズ、歪み
- (3)障害時の品質
- (4)遅延時間

### (5)相互接続性

(1)はお分かりと思います。(2)は通信経路でノイズや歪みが生じるほか、ロッシー圧縮符号化の副作用として生じるものがあります。(3)は通信が乱れたり途切れたりしても、不快な音が出ないことや通信が復帰したら音声も速やかに復帰することなどです。(4)は特に双方向の会話の場合などで円滑な会話のやりとりができるという観点で遅延が少ないことが求められます。(5)はいわば送信側と受信側の製造会社が異なっても問題なくつながるようになっているか、という観点です。(a)送信側と受信側が1対1で直接接続される場合、(b)1対1の接続でも、古くは電話の交換機のように、スイッチを介して相手を切り替えて接続する場合、(c)放送のように送信側が1つで受信側が複数の場合など、それぞれにおいて接続できる条件がどのようになっているか、という観点です。

さて、通信ではインターネット回線の場合、たとえ光ケーブルで接続されていてもIPパケットという単位でデータが欠けたり到着時間(遅延)が揺らいだりします。送信側と受信側が遠距離にあるとは限りません。Bluetooth ヘッドフォンや赤外線によるワイヤレス機器などのように極近いところでの無線接続も通信です。無線通信では経路上の電磁波や光などのノイズがあります。有線、無線を問わずこれらそれぞれの通信において(1)から(5)の観点で様々な工夫がされています。

### 3.3.音質評価の課題

電話など通信における音質(品質)評価は、ある決められたビットレートにおいて、大雑把に以下のように行われます。

- (1) 遅延、SN比などのように物理的に測定可能な量。
- (2) 言葉の明瞭度のように言葉が聞き取れるか、聞き分けられるかで判定するもの。
- (3) 人間の聴覚による比較によって音質を相対判定するもの。

ここで(1)は客観評価、(2)と(3)は主観評価と呼ばれます。音質評価には既に国際機関で定められているいくつかの手法があり、それぞれ音の比較で以下のような5段階の尺度で評点を出します。

カテゴリ	評点
・差が分からない	5
・差が分かるが気にならない	4
・差がわずかに気になる	3
・差が気になる	2
・差が非常に気になる	1

こうして、ある規模の評価者の結果を統計的に処理して、3.2など1から5の間の小数点の値が得られます。

前述した地上波デジタルTV音声で採用されたAACはこのような主観評価によって「CD並み」と評価されたものです。同様に今後現れる技術を「ハイレゾ並み」と呼ぶには、このような主観評価の観点も必要ではないかと思います。このように「ハイレゾ並み」という表現を突き詰めていくと、周波数帯域幅のような客観評価尺度と、「ハイレゾ」音との比較のような主観評価の条件をどのように決めていくかと言う課題に行き着くと思います。これはとりもなおさず「良い音」の追求の課題と本質的に同等だと思えます。ただし、ここで注意を促したいのは、電話の音声品

質が受話器、再生系の性能だけではなく、通信を経て最終的にユーザに届く音声の品質で評価されるということです。

オーディオ機器として見た場合、通信を考慮すると第 2 章に挙げた(1)電話並みから(5)ハイレゾまでの周波数帯域幅での分類では収まらず、通信上の制約に対する技術上の工夫を施して、ロスレスではないものの音質は「ハイレゾ並み」にした、という表現が自然に現れてくると思います。その際の音質の評価では、たとえば Bluetooth やネット配信など通信を経由する場合に通信込みでの評価手法が求められるのではないかと思います。

「ハイレゾ並み」はこれまでの分類の(4)CD 並みと(5)ハイレゾの間、すなわち「ハイレゾ未満」というわけではないということだと思います。「良い音」の追求と合わせて、現在は「ハイレゾ並み」としか表現できない技術の評価は今後検討されてしかるべきだと思います。一方メーカーの側には、どのような制約の下でどのようにして音質を「ハイレゾ並み」にできたのか、という点をよりアピールしていただきたいと思います。

#### 4.音の再現性と解像度(Resolution)

前章までは、電話からハイレゾ、「ハイレゾ並み」に至る技術の進展を、周波数帯域幅とビットレートの関係、通信上の制約に対するデジタル音声圧縮技術の観点から眺めて、音質の評価を考えようとしてきました。オーディオに限って言えば、「Hi-Fi」から「ハイレゾ」に至る技術革新が目指してきたものは音の再現性とそれに伴う臨場感の向上と言えます。「ハイレゾ」はその名の通り解像度を上げる流れです。以前、この音の解像度には以下のように 3 軸の方向があることを示しました<sup>[6]</sup>。

- (1)時間分解能=サンプリング周波数
- (2)音圧(dB)分解能/ダイナミックレンジ=ビット長
- (3)空間分解能=チャンネル数

現在の「ハイレゾ」の定義は(1)と(2)に関するものになっています。4K,8K 衛星デジタル放送の規格では(3)のチャンネル数は従来の最大 5.1ch から最大 22.2ch と大幅に拡大しています。上記の 3 軸では空間分解能をチャンネル数と等しいと置いています、22.2ch であっても人間の聴覚上検知できる方向の分解能は、チャンネル数あるいはスピーカ数で区切られる空間の角度より一般に細かいです。たとえばステレオであっても録音、再生環境が整えば左右のスピーカの間位置するオーケストラの各楽器の位置や、歌手の位置を指差せることからご理解いただけると思います。今後 5.1ch サラウンドや 22.2ch の良質なコンテンツに触れる機会が増えることで、音の再現性の議論が周波数帯域幅に留まらずに空間的な音場にまで広がり、(1)から(3)全てを含み、人間の聴覚特性を考慮した音場の音質評価法が一般的になることを期待したいと思います。また、それぞれの高解像度化に伴って通信技術がさらに発展することを期待しつつ、技術に対するユーザの理解も深まるようにアピールしていきたいと思います。

#### 5.まとめ

オーディオには収録、記録、通信、再生の 4 段階があります。「ハイレゾ」の定義はロスレスが前提になっており、客観評価ために(1)時間分解能=サンプリング周波数あるいは再生周波数帯域 と(2)音圧分解能/ダイナミックレンジ=ビット長を規定していますが、主観評価手法の規定には至っていません。

通信を切り口に考えるとき、通信上の制約からロスレスにできなくとも「ハイレゾ」に匹敵する音質を追及した技術革新が現れてきています。今後、これらの技術を評価するためには、ロスレス/ロッキーあるいはハイレゾであるか否かに拘わらず、通信を経た音も評価できる手法が求められると思います。それは客観評価と主観評価を含むものになるでしょうし、これはとりもなおさず「良い音」の評価が目指してきたところと本質的に同じだと思います。科学史的に見れば、技術の革新は、測定、評価技術の進展に拠るところが大きいのはご承知の通りです。「良い音」の評価手法の追求がオーディオのさらなる発展につながると思います。

また解像度の内、(3)空間分解能=チャンネル数については、前述した主観評価の観点を入れるということのほか、将来的には映像品質と合わせた臨場感の観点からも評価がほしいところです。今後解像度の3軸を追求するにあたって、通信上の制約がこれまで以上にクローズアップされて来ると思います。それらを克服するためにも「良い音」の本質に基づいたさらなる技術革新が必要となるでしょう。

最後に、新しい「ハイレゾ並み」の技術は「ハイレゾ未満」にあらず、ということをあらためて強調しておきたいと思います。

## 参考文献

- [1] 小谷野「ハイレゾ時代の良い音 - 今、オーディオ機器に求められるもの - 」JAS Journal 2017 Vol.57 No.1 (1月号) pp.14-pp.23
- [2] ARIB「デジタル放送における映像符号化、音声符号化及び多重化方式」標準規格 ARIB STD-B32 3.9 版 (2016.12.09)
- [3] 遠藤、原田、鎌本、守谷「MPEG-4 オーディオ・ロスレスとストリーミング配信について」JAS Journal 2012 Vol.52 No.1(1月号)pp.22-pp.29
- [4] <https://xiph.org/flac/>
- [5] ITU-T G.711 : Pulse code modulation (PCM) of voice frequencies
- [6] 遠藤「変わりつつあるテレビとハイレゾ配信=その展望と今後」 JAS Journal 2017 Vol.57 No.2 (3月号) pp.14-pp.23

## ■筆者プロフィール

遠藤真(えんどうまこと)

1978年千葉大学工学部卒業、同年日本電信電話公社(現NTT)入社。

2005年よりNTTエレクトロニクス株式会社(現職)。

2014年より日本オーディオ協会理事。

地上波デジタル放送用MPEG-2 HDTV符号化LSI開発等で2004年日本産業技術大賞内閣総理大臣賞(団体)、2006年前島賞、2007年文部科学大臣表彰 科学技術賞(開発部門)受賞。残響制御技術の開発と実用化等で2012年日本オーディオ協会大賞(団体)、2017年前島密賞受賞。

電子情報通信学会情報・システムソサイエティ会員、IEEE Computer Society 会員