

【連載：「NHラボセミナーより」第4回】

ハイレゾを考える

NHラボ(株) 茶谷 郁夫

良い音を求めて苦節 50 年、楽しみのはずの趣味が仕事となり、生きていくための手段となり、そしてまた楽しみに変わった。

ハイレゾは面倒な？オーディオを誰でも簡単に高性能で楽しめる素晴らしい世界を生み出してくれる。

ところで、NH ラボのセミナーは、学会ではなく、実証されていなくても、少し大胆な仮説で問題提起をし、より良いオーディオの為に議論が始まり、少しでもユーザーの為にすることが有れば良いのではないかとするスタンスで、開催している。

今回のテーマは 1 ユーザーとして、少し大胆な仮説も入れて話を進めた。過去の NH ラボセミナーでの説明をベースに補足を加えた内容で、デジタルの専門家から見ると相当無理があるかもしれない。スピーカを専門とする音響屋の私見として出来るだけ大目に見ていただき、また読者の皆様からの忌憚のないご意見を伺えることを期待している。

■ 良い音

Hi-Res は音が良いからこれからの主力になる。と私は確信している。音が「高解像度」でありそのために「高忠実度」をも実現しているのだから。この場合の技術的進化は、音愛好家からすれば音が良くなることだけである。

昔から良い音は人により違うから一律には決められない。との話を思い出した。私は良い音は皆同じとのスタンスである。この場合良い音から、“好みの音”を除外する。そして、自然な音、生の音、録音現場の音を基準にすると、一応答える。

生の音が録音できるのか？との問いに、出来るだけ生々しく、鮮度が良い、上質の刺身のような物と答える。私は仕事柄、現役後半はこのスタンスでスピーカの調整をしてきた。間違いは無かったと思っている。もちろん生の音は録音できていない。

皆が一番よく知っている、日ごろ聴きなれた自然の音の再現が目標である。もちろん、オーディオの世界が有ることも知っている。素晴らしいシステムを随所で聴かせていただいた。それは時として本物以上に生々しい、魅力的はオーディオの世界を再現する。

Hi-Res を楽しむうえで一番の問題はプログラムソースである。なかなか良いソースに巡り合わない。あ！このことですか、“Hi-Res=音が良い”ではないというのは？

もっとも CD でもそうであった、良い録音に巡り合うのは確立 10%程度でしょうか？

Hi-Res で最も問題なのは、デジタル機器や再生ソフトの中で何をやられているか分からないことに有る。そこでお願い、先ず音源ソースの素性を明確にしてほしい。できれば使用機材も明記してほしい。次に再生ソフトやデジタル機器も中で何をやっているか明確にしてほしい。

今自分が聴いているものがどのような素性の物が分からないと、音屋としては判断がしにくい。

録音時の機材スペック、PCMなのかDSDなのか。PCからの出力、DAC。デジタルアンプでのデジタル入力と、アナログ入力の取り扱い、等々。

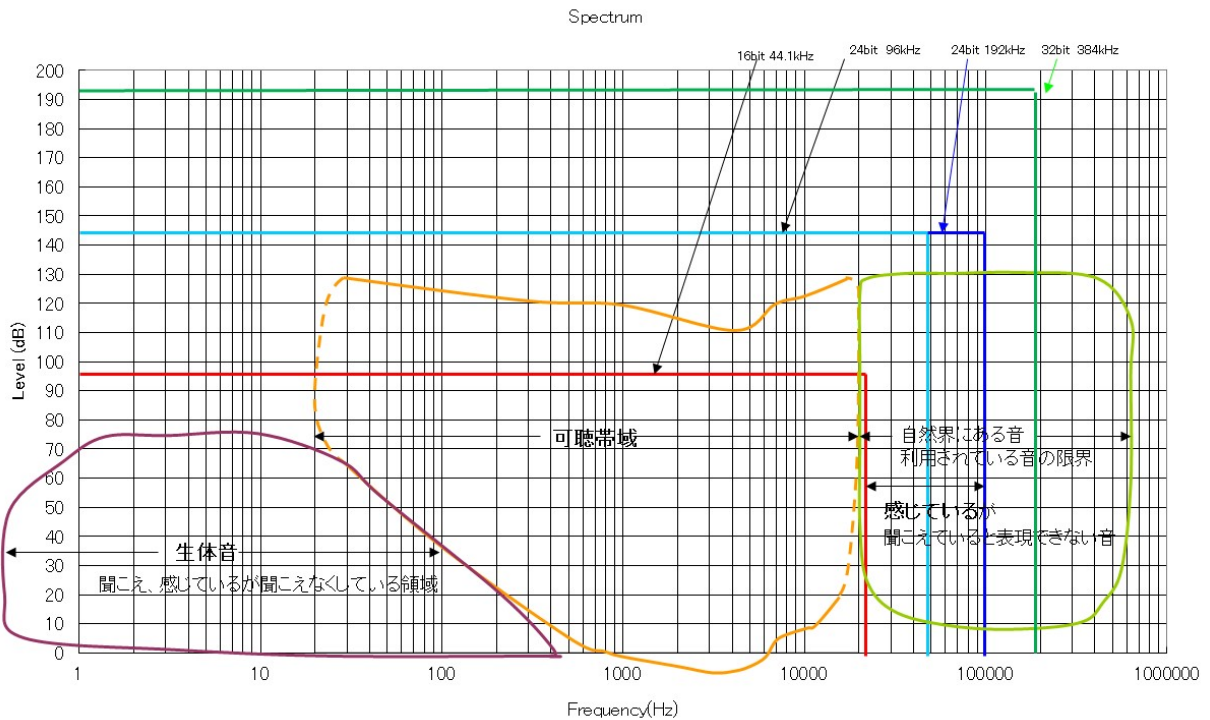


図1. 音の周波数帯域での拡がり

## ■ 周波数帯域

図1に音の周波数帯域での拡がりレベルを合わせて書いてみた。先ず真ん中に可聴帯域があり、低いほうに生体音やノイズが有る。超低域（可聴帯域外の低域）の音は聴こえているとさくて生活しにくいので普通は脳が聴こえなくしている。可聴帯域以上にも自然界に音はたくさん存在しているが、普通意識しても聴こえない。この帯域は感じているが（何らかの方法で？）表現できない。ピアノの鍵盤の最高音が4186Hzなのはこれ以上高い音の鍵盤を作っても高い音としてわからないからである。

私は可聴帯域の音と超高域が音の成分として同時に聴こえてきたときには、超高域成分の有り無しで違いを感じると仮説を立てている。意味のある超高域の場合には分かる（感じる）と思っている。

同じ図にPCMのサンプリング周波数とビットの違いでどうなるか書き込んでみた。ハイサンプリングとハイビットに従い周波数帯域が広がり、ダイナミックレンジが大きくなる。Hi-Resのすばらしさを表現するつもりのこの図は、実は間違えている。周波数帯域の広がりとしては良いのだが、基準値がビット毎に違う。これを同時に書くことはできない。

超高帯域の効果はハイパーソニック・エフェクトとして知られている。脳の活動に変化が現れるのには数十分の時間がかかる。また、ヘッドフォン受聴だけでは効果が出ないとの報告が有る。

Hi-Res ソースを聴くと細かいニュアンスが分かり、質の向上を感じる。瞬間的に音質の良さを感じる。またヘッドフォンでも同様に良さがわかる。

Hi-Res は可聴帯域内でも十分良さが分かる。骨伝導や体の表面での受聴を予測している超高帯域音の効果とは別の良さもある。

Hi-Res で良く再現される空間の広がり感はヘッドフォンではまだ上手く出ない。

### ■ ダイナミックレンジ

Hi-Res ではダイナミックレンジが相当広い。24 ビットであれば、概略  $24 \times 6 = 144\text{dB}$  である。従来のアナログ系では  $60 \sim 70\text{dB}$ 。CD の 16 ビットでは  $96\text{dB}$  で十分大きい。アナログ系ではリミッター処理をしてピークレベルに注意を払った。その結果、録音レベルを上げられるのでダイナミックな音にすることが可能であった。初期の CD ではクリップを恐れ、平均音圧レベルを下げて録音した。その為か？ユーザーが期待した、ダイナミックな音の再現に苦労した。最近の音の良い CD は総じて録音レベルが高い、ダイナミックな良さを感じる。

Hi-Res は、ダイナミックな音とは別に非常に小さい音での再現性の良い事に驚かされる。24 ビットで  $144\text{dB}$  と言うのは  $1\text{V}$  を基準に考えると  $1/2^{24} = 60\text{nV}$  の細かさで信号を拾えるということで、信号の密度と精度の向上が期待できる。 $60\text{nV}$  を基準として  $144\text{dB}$  のダイナミックレンジが有るといふことだ。単純に 16 ビットの  $96\text{dB}$  とは比較できない。Hi-Res だからダイナミックな音が出るわけではなく、むしろ細かい音の再現が出来るということだ。

### ■ 音の生々しさ

もう一つ Hi-Res の試聴において感じるのは、生々しさである。昔、録音スタジオを見学した際に、圧倒的な音圧とともに歪の少ない鮮度の高い音に驚いた。音楽スタジオ、TV 局、ラジオ局でモニターの目的による違いが有るにしても、生は生である。

その当時はテープの音であったが、素晴らしく、この音を家庭に持ち込めるのは何時のことになるかと思った。デジタルになれば録音現場の良さが伝送経路、媒体を経ることで劣化が少なくなり、生の音源を得る可能性が高くなる。デジタルデータの伝送では、デジタル信号の正確な伝送が保証されている（ビットパーフェクト）。

アナログ信号をデジタル化する際や、デジタル信号をアナログに戻す際に AD,DA が行われているが、この AD,DA の音の変化が大きそうである。

Hi-Res は周波数帯域だけでは無く、時間軸上の細かさ。ダイナミックレンジだけではなく、レベル（深さ）方向の細かさに効果がある。時間軸上の細かさとはレベルの細かさを合わせて見ると、その瞬間、瞬間の変化の精度が良くなるということであろう。立ち上がり特性の精度を考えると、可聴帯域以上の変化に追従できるか？となり、超高帯域も必要となる。

### ■ 音の立ち上がりを考える

オシロスコープの応答を概算するときの関係「周波数帯域×立ち上がり時間 (s) = 0.35」を無理やり当てはめると、 $50\text{kHz}$  では  $7\mu\text{sec}$ 、 $100\text{kHz}$  では  $3.5\mu\text{sec}$  となる。次ページの図 2 は  $10\mu\text{sec}$  で  $100\text{dB}$  の立ち上がりという極端な場合の例で有るが（本当にあるかは不明）、いず

れにしても 96kHz サンプリングで 10.4 $\mu$ sec、192kHz サンプリングで 5.2 $\mu$ sec 毎となるのでサンプリングが粗い。DSD の場合 PCM に比べ、細かく刻んでサンプリングしているのが分かる。

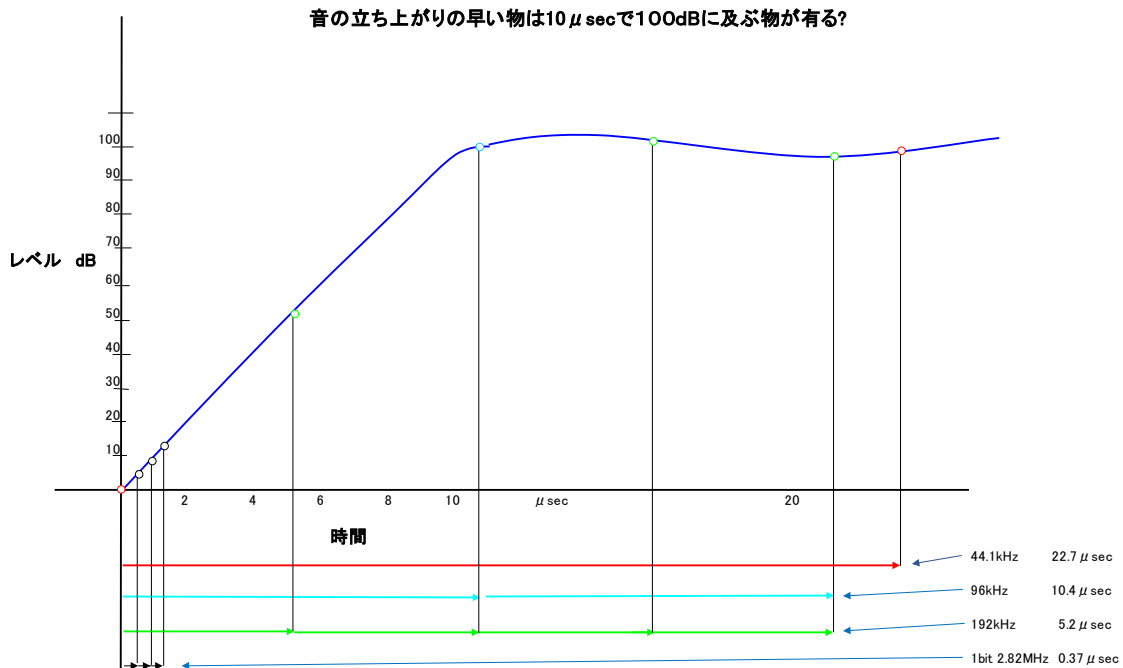


図2、立ち上がりの早い信号でのサンプリング

最近アナログレコードが挽回してきているらしい。少し乱暴な表だがアナログとデジタルの利点欠点をまとめてみた。未だにアナログの音の滑らかさは魅力だ。アナログで良い音を出すのはなかなか大変である。ただ、アナログで培われた良い音を出すためのノウハウはデジタルでも同じように生きる。問題点が多いにもかかわらずアナログの音は良いと言う。聴きやすく、暖かく、なめらかに聴こえる。オーディオ信号のデータ化にまだまだ問題が有るということだろう。デジタルは、硬く、ギスギスし、冷たいと言う。それでも利点が多い。

	<b>アナログ</b>	<b>デジタル</b>
問題点	ダイナミックレンジが狭い 周波数帯域が狭い ノイズが多い ワウフラッターのような変動が起こりやすい 伝送により信号が劣化する 左右のレベル差が生じやすい 左右の位相特性が変動しやすい 信号レベルが高いほど歪が多い Lpレコード再生では内周程歪が増える	オーディオ信号のデータ化で信号は劣化する 信号レベルが低いと歪が多い 雑音、歪の絶対量は振幅によらず一定 過負荷レベルを超えるとクリップ歪が急激に増える コピーしても劣化しない
利点	価格に比例した性能と言える 信号が連続しており、ギスギス感が少ない 過負荷での歪はソフトクリップである	機器の価格、省エネルギーの点で優れる 低価格でもかなりの性能を有する ダイナミックレンジが広い 周波数特性がフラットで広い ノイズもアナログより少ない 左右の信号変動が少ない 伝送により信号は劣化しない

表1、アナログとデジタルの問題点と利点

## ■ CD でも良いのではないか

CD も、制作側が慣れてすっかりその器を使いこなすようになった。初めのころに比べ、随分と聴きやすく、良いバランスになってきた。

そもそもデジタルは基本性能の高さと、伝送劣化のなさが売りで有った。音はアナログとは違う新鮮さが有った。だがユーザーが期待するのは滑らかな聴きやすい音だった。

一方、今の Hi-Res では基本特性は良いがまだバランスのおかしなソースがある。良いソースになかなか当たらない。早く制作側で新しい器 Hi-Res に慣れて使いこなしてほしい。

CD で十分楽しめる。そこに Hi-Res としなくても良いのではないか。ほとんどのマイクロフォンも音質の良いものの高域は 20kHz 以下であろう。しかし将来の可能性とその時代の文化レベルの低下を避けるために、Hi-Res でさらなる鮮度の高い音を期待する。目標は録音現場のモニターから出る生々しさである。身近な自然の音、仲間の声、これらも常日頃親しんでいるから基準になる。Hi-Res は基本性能が高いので、安価でもそこそこの音がする。ここが危険！本当の良さが伝わらないとわざわざ Hi-Res にしたのにデータ量ばかり増えて面白くないという話になる。

## ■ ハイビット・ハイサンプリングの精度

ハイビットでの疑問はこのような細かい電圧変化を正確に AD,DA 出来るのか？となる。

基準電圧 1V として、16bit:  $1/2^{16}=15\mu\text{V}$ 、24bit:  $1/2^{24}=60\text{nV}$ 、32bit:  $1/2^{32}=0.23\text{nV}$  の細かさ。

PC 内部やデジタル機器内で色々な電圧変動が発生し影響を受ける。ビット数が高いということは、その細かさで信号を切り取るということである。

ハイサンプリングの疑問はこのような細かい時間変化を正確に出来るか？となる。

44.1kHz:  $1/44.1\text{k}=22\mu\text{sec}$ 、192kHz:  $1/192\text{k}=5.2\mu\text{sec}$ 、384kHz:  $1/384\text{k}=2.6\mu\text{sec}$  の精度。

DSD の場合 2.8MHz:  $1/2.8\text{M}=0.36\mu\text{sec}$ 、11.2MHz:  $1/11.2\text{M}=0.09\mu\text{sec}$  の精度となる。

ソフト制作現場は、何をやっても高級機材であり関係者が良しとすれば OK としよう。また、高性能クロックも使用している。ユーザーの機器や試聴ソフトではその精度を維持できないと思う。

デジタル信号のままであればそれでも問題は少ないと思うが、DA,AD する場合、ノイズ、電源変動、クロックの乱れで影響を受ける。

Hi-Res を聴いていると細かい音の変化、柔らかさ、小さな音のエコー感などが違う。従って、低レベルの音での時間変化の精度が十分確保されないといけない。

先日もデジタル屋さんに怒られた、“デジタルケーブルで音が変わりますか？”

答えは小声で“はい”。

## ■ 耳の信号処理

耳の聴覚神経のパルス発生を見てみると、先ず周波数分析をし、その成分毎にパルスが発生する。パルスの発生は、位相固定の働きにより、位相を判ってパルスを出したり出さなかったりする。パルスの発生は、音の大きさによって強いパルスを出すのではなく、大きさによってパルスの発生個数が増える。これは、PCM 的ではなく、PDM 的である。DSD 的なのである。

当然、脳の他の伝達神経パルスも同様に PDM 的は信号伝達を行っている。

私は PCM の音を少し硬いと感じ DSD の音を柔らかいと感じる。PCM の音は明瞭でしっかり

しているとの意見もあるが、私は DSD が何か人間的で良いなと思う。

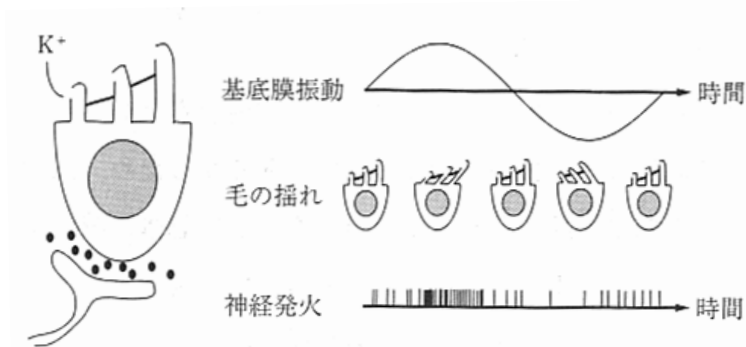


図3、内毛細胞の神経発火の模式 (音響学入門 日本音響学会編)

### ■ 混変調歪

初めに、穴沢氏が JAS ジャーナルに 2015 年 5 月に掲載した文を引用させていただく。

「40 年以上前に 100kHz まで平坦な特性を持つ測定用マイクを使って楽音を録音し、再生してみたことがある。そのまま再生すると混変調ひずみによって生音と全く違う音が出てくるが、15kHz 以上を切って再生すると元の音に近いまともな音を聴くことが出来た。これは、高域、特に可聴域外での非直線により混変調歪が発生し、可聴域内にひずみが発生するためであった。」

混変調に関しては小谷野氏の論文、穴沢氏の指摘、蘆原氏の出版物等で指摘されている。

そもそも人間の耳も相当の混変調を起こすので、なかなか難しいが、いずれにしても良い高音用スピーカ (TW) が必要となる。今までの経験ではフィルム TW、コンデンサー TW、ホーン TW 高剛性ドーム TW などが良かったが最近では小口径フィルムドーム TW でも良いものが有る。癖のない、伸びやかな、抜けの良い音のするものが良い。これを小入力で使用するのが良い。出来れば可聴帯域を入れないようにし、低振幅で使用するのが良い。

まあ、40kHz 以上伸びているものは大体伸びやかな音がする。

普通のスピーカに超高域音を追加すると艶やかな細かい表現に優れた再生音となり魅力的である。面白いのは CD でも効果があるということだ。艶やかな音のするものは過渡特性に優れ、可聴帯域の応答速度を向上させる。これが良い効果を導く。

余談だが、ヘッドフォンは振幅が小さいので混変調には強い。したがって Hi-Res の効果も良く分かる。

### ■ 音場感

Hi-Res ソースを聴くと音場感の再現に優れる。特にエコー感の変化が美しい。

通常、録音のマイクの位置で直接音と間接音の両方が到達し録音される。これを再生する場合、再生する部屋や環境で更に間接音 (付帯音) が追加される。この追加された付帯音が音場感を阻害する。間接音は何処から来ているか分からないような広い指向性で再生されるべきである。

本来は直接音も間接音も方向性を持ちその方向へ再現されるべきである。が方向の情報が無い。良い Hi-Res で音場感を感じられる場合、小さな音の変化、位相に忠実であるからに違いない。

マルチウェイスピーカでもタイムアライメントを考慮したものが良い音場感を再現する。

位相までフラットにしたものは少ないが、DSP を使用した特性補正でも、周波数特性だけでなく、位相までフラットにした FIR を使用した物は大変に印象が良い。EQ をするならこれ以外に無い。

フルレンジで可聴帯域を広くカバーするスピーカは少ないが、たまごスピーカはこれにあたる。何処から音が出ているか分からないほど、音場感の再現に優れる。音場感の再現には小さなレベルの音の変化を聴こえるようにする必要が有る。部屋のノイズを下げ、S/N を良くする。

たまごスピーカは余計な付帯音が少ないので邪魔をしない。音源の時間変化、音色の変化をしっかりと維持する、これが良さの原因である。

## ■ 耳の特徴

聴覚では高域での時間精度が良い。周波数分析するとき FFT では高い周波数での時間分解能が低い、聴覚では低い周波数では周波数分解能が高く、高い周波数では時間分解能が高い。これは、高い周波数の時間変化、位相情報が生きていくうえで大切な情報だからであろう。

左右の耳から来た神経パルスと比較し時間差を知る組織がある。5mの距離で1度の変化が分かるという実験が有り、その差はわずか 10 $\mu$ sec、100kHz の音の 1 波長分の長さである。

もう一つ面白いのは、神経パルスが音の立ちあがりにたくさん発生する。これも立ち上がりを大切にしている証拠になる。

聴覚は、ある周波数の変化に注目しているのではなく、スペクトルの形の変化、音色の変化に着目している。これは、プロフィール分析と呼ばれている。音色の変化を決めている要素は以下のようなものが有る。そのどれもが Hi-Res によって、より忠実に維持される。

周波数スペクトル 主としてパワー  
 振幅の時間エンベロープ  
 倍音構造の不規則な変動  
 雑音成分の混在  
 倍音周波数のズレ  
 ビブラート  
 ビート

特に立ち上がり部分の特性が音色知覚に大きな影響を及ぼす。

表 2、音色を規定する物理量

(音色の感性学 日本音響学会編)

## ■ Hi-Res で注意すべき点

以上考えてくると、デジタルデータは忠実に伝送されているから、問題は取扱である。

ユーザーとしては主に DA が問題となるが、AD も問題となろう。

音楽ソフト、デジタル機器の内部、再生ソフト内部で何をやっているかは明確にしてほしい。

聴覚の特殊な能力を考えると、可聴帯域内でもまだまだ大切にしてほしい点がありそうである。

穴沢氏の JAS ジャーナル 1982 年 12 月に有るように、AD 変換する場合の許される誤差から計算された時間精度は 20kHz、10V、16 ビットで 121psec というとんでもない値となる。また、蘆原氏の“超広帯域オーディオの計測”に有るように 22kHz、24 ビットでは 0.43psec、

192kHz、24ビットでは0.1psecとなる。再生機器には、より精度の高いクロックが求められる。

このように高い精度を維持するのは容易でないと思う。つついスペック競争に踊らされてしまう自分が情けないが、Hi-Resにはまだまだ改良ポイントが有ることが分かった。

## ■ まとめ

ソフト：ソフトの素性と処理過程を明示してほしい。そこから、Hi-Resの良い点を感じ取り、魅力的な違いを大切にするユーザーが育つと思う。

ハード：ハードや再生ソフトでどう処理しているか明示してほしい。どのスペックのHi-Resを聴いているのか分かれば、有難味も増加する。

ハイビット、ハイサンプリング：レベル方向と時間方向のデータ化する時の細かさを表していて、ダイナミックレンジや周波数帯域の広さだけではない。

可聴帯域以上の必要性：性能の良いツイーターを使用すれば、CDでもずっと良くなる。

音の鮮度：音の鮮度はオーディオ的再生技術とジッターの少なさで維持される。

音の質：CDグレードとHi-Resを比較すると瞬時的に、質、深み、ニュアンス、広がりが違う。聴覚はこのような小さな違いを瞬時に聴き分ける能力を持つ。

昨年の春の1ビットオーディオ研究会でPCM 384kHz/32bitと11.2MHz DSDの音を聴かせていただいた、どちらも質の点ではこれで十分と思わせる物であった。装置も申し分ないスピーカーも高級オーディオ用である。ジャズ演奏で有った。ところが、音場が違う。あそこまで質が向上すると、次にステージの再現を期待する。たぶん、スピーカーの指向性がいけないのではないかと考えた。各楽器の音はスピーカーから直接出てきて良い。定位、音像ともに申し分ない。

ところが、音場感を出す間接音成分は無指向で出ないといけない。そうすれば、ステージが有り、各楽器が演奏しており、それを包み込む空間が出来る。

最近話題のドルビーアトモスはついに天井にスピーカーを付けることを要求した。さすがにそのせいもあって、音の空間再現は見事である。もっとも、はなはだ人工的ではあるが。

## 参考文献

- ・ PC オーディオガイドブック 島幸太郎 インプレスジャパン
- ・ デジタルオーディオの基本と応用 河合 一 誠文堂新光社
- ・ デジタルオーディオにまつわる迷信とその克服 山崎芳男 JAS ジャーナル 1982年8月
- ・ 音響学入門 鈴木陽一他 共著 コロナ社
- ・ 高音用スピーカーの歪み発生要因の検討 小谷野進司 PIONEER R&D Vol.13 No.2
- ・ 超広帯域オーディオの計測 蘆原 郁 編著 コロナ社
- ・ アナログレコードの音質改善とデジタル録音の導入 穴澤健明 JAS ジャーナル 2015年5月
- ・ 聴覚の研究 京都大学
- ・ 空耳の科学 柏野牧夫 株式会社ヤマハミュージックメディア
- ・ 音色の感性学 岩宮眞一郎 コロナ社