

『連載ラウドネス講座』

第 1 回 ラウドネスの基礎とトゥルーピーク

技術委員会オーディオ基準小委員会 丸谷 正利

いま関心の高まっているラウドネスを中心に、オーディオ全般のトピックも混えて解説記事を連載します。第 1 回はラウドネスの基礎と True-peak（トゥルーピーク）です。

すでにご存知の方もいると思いますが、民放連では 2011 年 7 月の完全デジタル放送化に向けて音声レベルに関する新しい運用基準を策定中です。JPPA でもオーディオ基準小委員会を中心となって民放連と話し合いを行っています。新運用基準では、テレビ番組や CM の音声レベル管理にラウドネスや True-peak という考え方を導入するという話です。民放連の動向についてもこの連載の中で触れたいと思います。

1. なぜラウドネスなのか？

昔からテレビの音量問題は議論されてきました。CM になるとうるさい、チャンネルを変えると音量が変わる、番組によってはセリフがよく聴こえないなど、視聴者からはいろいろなクレームが上がって来ます。それでもこの問題はなかなか解決できないのが現状です。アナログ放送では音声信号の変調度の関係からある意味で音声レベルに制限がかかっています。しかし、デジタル放送になると変調度という制限が無くなり、物理的なレベルは 0dBFS まで使用することが可能となります。いまでもクレームとなっている音量問題がもっと顕著化する恐れさえあります。これを避けるためにはデジタル放送に対応した音声基準を作る必要があります。アナログ放送の経験から、現在の VU メータによる音声基準では適切な管理が困難であることがわかっています。そこで、定量的に音量レベルを測定する方法として注目されたのが「ラウドネス」です。

2. ラウドネスとは何か

ラウドネス(loudness)とは人の感じる音の大きさのことを言います。ラウドネスの詳細は次回に譲りここでは基本的な知識について述

べておきます。

「ラウドネス」は感覚量（心理量とも言います）を測定する方法です。つまり「人の感じる音の大きさ」を測定します。ラウドネスで音量管理を行うことは、音の大きさを音量管理を行うことになるので、視聴者に対しレベル差の少ない番組や CM の提供が可能となります。

これに対し我々が日常使用している VU メータによる音量管理は、電圧レベルの大小でメータを振らせているので物理量の測定となります。このため、同じメータの振れでも実際に人が感じ取る音の大きさが異なるということが多々あります。その理由は ISO 226 等ラウドネス曲線（図-1）を見てお分かりのとおりです。（実線が新規格(2003)、破線が旧規格）。

3. ラウドネスと音圧レベル

国際規格 ISO ではラウドネスの基準音は音圧レベル 40dB の 1kHz 純音となっており、この時のラウドネスが 1 sone（ソーン：ラウドネスの単位）となります。このラウドネスをデシベルと同じように対数表示にしたものが「ラウドネスレベル」で、単位に phon を用います。1 sone は 40 phon と定義されています。

したがって、1kHz 純音の場合は、

$$1 \text{ sone} = 40 \text{ phon} = 40 \text{ dB SPL}$$

という関係が成立します。等ラウドネス曲線の 40 phon のカーブは 1 sone のカーブでもあります。

1kHz 純音では、音圧レベルが 10dB 大きくなるとラウドネスは 2 倍 (2 sone) になります。しかし、図-1 から分かるように低域周波数になるにしたがって音圧レベルとラウドネスの比率が大きくなります。20Hz では音圧レベルが同じ 10dB の変化でもラウドネスは約 4 倍になります。

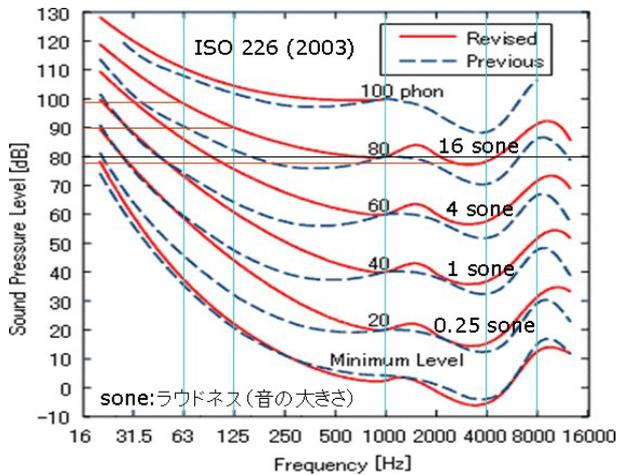


図-1 ISO 226(2003) 等ラウドネス曲線

4. 等ラウドネス曲線の読み方

図-1 左側の音圧レベル 80dB のラインをたどると、1kHz ではラウドネスレベルがちょうど 80phon になります。この曲線を左の方へたどり 125Hz のところで音圧レベルを読むと 90dB になります。つまり 1kHz と同じラウドネスレベルを得るには音圧レベルを 90dB にしなければならないことが分かります。この場合、VU メータは 1kHz に対して 10VU も大きく振れることとなります。63Hz の場合はさらにその差が大きくなります。一方、3kHz 近辺の周波数では同じ 80phon のラウドネスレベルを得るのに約 77dB の音圧レベルがあればよいことが分かります。つまり、耳の感度は低域ほど鈍く、3kHz 付近で最大となっています。

CM 制作の手法、「VU メータはそれほど振れないが音量のある音」は、この物理量と感覚量の差を利用したものといえます。

5. ラウドネスの測定

残念なことに我々が扱う音楽や番組音声のような非定常音 (時間変動をともなう音) のラウドネスを測定する方法はまだ確立されていません。ISO 532B で規定されているラウドネス測定方法は定常音 (ある一定サイクルで繰り返される音) が対象であり、非定常音のラウドネスを正確に測定することが出来ません。ITU-R のラウドネス算出方法も試行錯誤の中で決められたアルゴリズムのひとつで、この方法が正確ということではありません。例えば、純音に対しては正しいラウドネス値を求めることが出来ないという問題があります。

6. ITU-R の勧告文書

ITU-R ではラウドネスに関する勧告文書をいくつか出しています。ひとつは BS.1770 で、この文書はラウドネス測定アルゴリズムに関して記述しています。二つ目は BS.1771 で、この文書にはラウドネスメータに関する規定が記述されています。三つ目はまだ番号が取られていませんが BS.[LOUD]と呼ばれる、番組の国際交換基準で、番組制作におけるターゲットラウドネス (等価ラウドネスレベル) に関する文書です。この文書は 2009 年 11 月の ITU-R 会議で採択され、現在国際投票に掛けられています。ここで合意されると正式勧告文書となります。

7. ITU-R のラウドネス単位

ITU-R ではラウドネスレベルの単位として LKFS を用いています。これは ISO/IEC が定めているラウドネスレベルの単位 phon と異なります。phon は実際に聴取している音のラウドネスレベルを示す単位ですが、LKFS は録音

データのラウドネスレベルを表す単位です。60phon の音はどんな環境で聴取しても 60phon ですが、-24LKFS の音は再生時に音の大きさ（視聴ラウドネス）を変えることができます。この違いを認識しておく必要があります。また、ITU-R のアルゴリズムは補正カーブ K を用いた計算方法で、これは ISO の騒音レベル測定方法（A-weighted や C-weighted の sound pressure level）と原理は同じです。

3 月中旬に行われた EBU のラウドネスプロジェクトチーム P/LOUD の会議では、LKFS よりも LUFS (Loudness Unit Full-Scale) の方が ISO 80000-8 のネーミングルールに合っているとの見解を出しており、ITU-R に対し修正意見が出てきそうです。

8. 新勧告文書 BS.[LOUD]

先ほど述べた新勧告 BS.[LOUD]は、我々ポストプロダクション作業に関係する文書です。この勧告は番組制作時のターゲットラウドネスとその測定方法を決めているからです。BS.[LOUD]ではターゲットラウドネスが-24LKFS となっており、ラウドネス測定方法として音響信号全体を測定する方法（日本、ヨーロッパ方式）と、ダイアログレベルを測定する方法（米国方式）の二つが認められています。また、視聴者が快適に感じるラウドネスの範囲（Comfort Zone）についても記述されています。ただし、この新勧告に記述されている-24LKFS を-23LKFS に修正しようという動きがあります。4 月 19 日から開催される ITU-R 会議で議論されることになりそうです。

9. True-peak とは

True-peak とは“真のピーク”という意味ですが、要するにアナログ信号のピークレベルのことです。これと今使用しているピークレベルメータの“ピーク”と何が違うのでしょ

うか。

通常我々が使用しているピークレベルメータは、A/D 変換されたサンプルデータのピークレベル（これを **Sample-peak** とする）を基にレベル表示を行っています。この方法では周波数が高くなるとサンプルポイントが少なくなり、入力信号の“真のピーク”を捉えられない場合があります。特にサンプリング周波数と整数倍の関係にある周波数はサンプルポイントが固定されるため、**True-peak** との誤差が生じやすくなります（図-2 参照）。

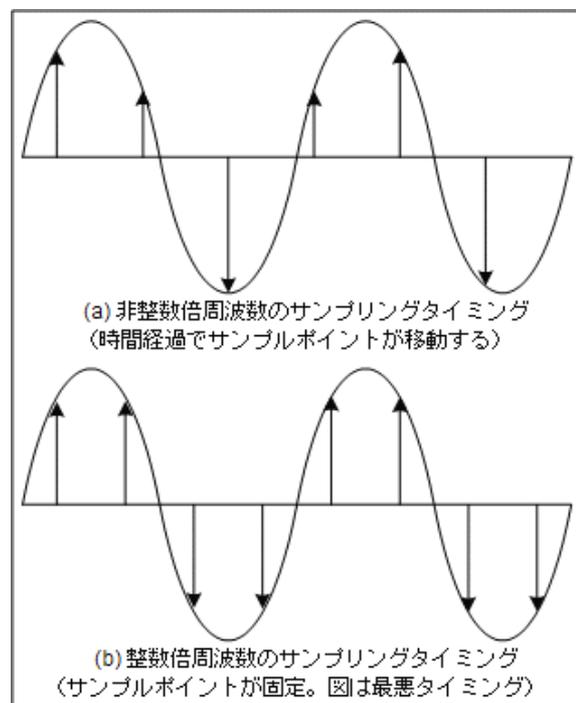


図-2 サンプリングタイミングによる誤差

10. Sample-peak とピークレベルオーバー

ピークレベルメータで 0dBFS 管理を行っても、True-peak との誤差を生じる（本当の信号はピークレベルオーバー）と、D/A 変換でオーバーロードが発生します。これはサンプリング周波数の 1/2 以下の周波数は正しく変換される（シャノンの定理）からです。A/D 変換の時は捉えることの出来なかったピークレベルであっても、D/A 変換ではそのピークが再現され、その結果オーバーロードになります。

ステレオやサラウンドの再生では、瞬間的なオーバーロードは気付かれることが少ないかもしれませんが、サラウンドからのステレオダウンミックスではオーバーロードがもっと顕著になる可能性があります。また、音声圧縮の過程でもピークレベルの増加が懸念されています[文献 1]。このような問題も含めて、ピークレベルの管理をしっかりと行おうというのが True-peak によるレベル管理です。

11. Sample-peak と True-peak の誤差

サンプリングでどの程度の誤差を生じるのか、48kHz/fs の Sample-peak と True-peak の最大誤差理論値を計算した結果が表-2 です。サンプルポイントが固定状態になる周波数の最大誤差は、その周波数のピークがサンプルポイントの間にある場合となります。計算もその条件で行っています[文献 1, 2]。

被測定周波数	True-peak との誤差	Sample-peak 0dBFS 時の True-peak 値
1kHz	-0.0186dB	+0.0186dBFS
2kHz	-0.0746dB	+0.0746dBFS
4kHz	-0.3011dB	+0.3011dBFS
8kHz	-1.2494dB	+1.2494dBFS
12kHz	-3.0103dB	+3.0103dBFS
16kHz	-6.0206dB	+6.0206dBFS

表-1 fs=48kHz の最大誤差理論値
(下 5 桁目を四捨五入)

True-peak との差 : True-peak を 0dB とした時の Sample-peak レベル

表-1 から分かるように、周波数が高くなるにしたがって True-peak と Sample-peak の差が大きくなります。16kHz/0dBFS の信号を入力した時、ピークレベルメータ上では最大 6dB の誤差を生じることになります。例えばピークレベルメータが 0dBFS を表示している時、True-peak は+6dBFS かもしれません。また、ピークレベルメータが-6dBFS を表示している時、True-peak は 0dBFS かもしれないということです。

現実には 8kHz や 16kHz 成分がフルスイ

グするような音響信号は考えられませんが、大振幅の中低域信号の上にこれらの高域信号が乗っている場合を考えると、誤差の発生は避けられません。

このような誤差を極力少なくして“真のピークレベル”を表示しようというのが True-peak の考え方です。BS.1770-1 では True-peak 処理のためにオプションを含め 6 つの条件が記載されています。そのひとつが PCM 信号の 4 倍オーバーサンプリング (48kHz x4 = 192kHz) 処理です。先ほどの 16kHz で考えると、4 倍オーバーサンプリングを行うことで Sample-peak と True-peak の差は 0.301dB まで小さくすることが可能となります。

12. SACD のピークレベル

同じデジタル信号でも SACD は DSD 方式で A/D 変換を行っています。これはデルタシグマ変調による 1bit A/D 変換ですが、サンプリング周波数は 44.1kHz の 64 倍 (2.8224 MHz) になっています。SACD の信号レベル規定では DSD 信号の 50%変調を 0dB (これを 0dB SACD と呼ぶ) と定めています。したがって 100%変調 (クリッピングレベル) の DSD 信号は+6dB となりますが、SACD に記録できる最大信号レベル (これを MaxPeak と呼ぶ) は+3.1dB までとなっています (50kHz 帯域)。100%変調の信号レベルが+6dB ですから D/A 変換では記録上の Max Peak に対して+2.9dB のマージンを持っていると考えることができます。

13. ピークレベルメータの運用マージンは?

それでは私たちが使用しているピークレベルメータの場合、ピークをどのレベルで管理するのが良いのか、つまり最大ピークレベルをマイナス何 dBFS で運用するのが適切なのかについて考えてみます。これは悩ましい問題で、高域信号のレベルをどの程度まで考慮するかということになります。例えば、

8kHz/0dBFS の入力信号に対して、そのピークレベルを保障するのであれば表-1 から、-2dBFS と考えることが出来ます。

同様に 12kHz/0dBFS の信号を保証したければ-3dBFS にする必要があります。

ITU-R のオーバーサンプリング後の誤差の考え方もこれと同じです（コラム参照）。48kHz サンプルングの時の理論上のマージンは 20kHz で約 12dBFS 必要となります。

一方、実際の音響信号を分析してみると 8kHz 以上の信号成分のレベルは-30dBFS 以下の場合がほとんどです。これを考慮すると実用的な最大許容レベルは-1~-2dBFS、安全を見て-3dBFS もあれば十分かも知れません（ただし、圧縮時のマージンは考慮していません）。まだ分析途中なので詳しい報告は別の回に行います。

14. 素材の録音レベル解析

最後に音声素材を使用した解析の一例を紹介します。波形とラウドネスメータ、VU メータ、QPPM（主にヨーロッパで使用されている IEC Type II レベルメータ）による録音レベル解析を行ってみました。使用した測定ソフトはラウドネス評価テスト用 LMCU v1.5（オーストラリア放送）です。少し古いバージョンなので現在 ITU-R で議論している新機能には対応していないことと、測定がモノラルのみ（ステレオ素材は Lch を測定）となっています。したがって、測定したラウドネス値は最新のラウドネスメータと比べ多少異なることをご了承ください。

15. VU メータとラウドネスレベルの検証

使用した素材は、ポップス・CM・スピーチ・バラエティの 4 種類で、それぞれ 10 秒間を解析しました。解析画像を図-3~図-6 に、それを表にまとめたものを表-2 に示します。表の「VU 平均」「LOUD 平均」は 10 秒間の解析値（平均値）です。「-24LOUD」と「平

均値 0VU」は解析値から計算で求めた値です。

「-24LOUD」はラウドネス値を-24LKFS にした時の VU メータの平均値、「平均値 0VU」は VU メータを 0VU 平均にした時のラウドネス値です。

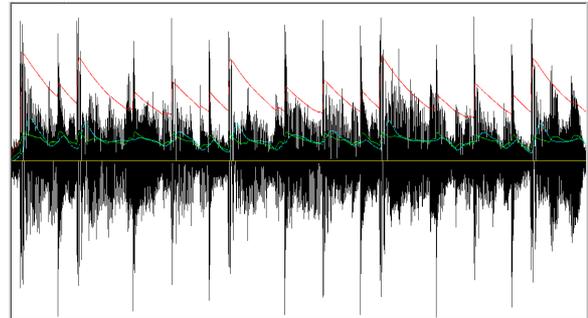


図-3 ポップス

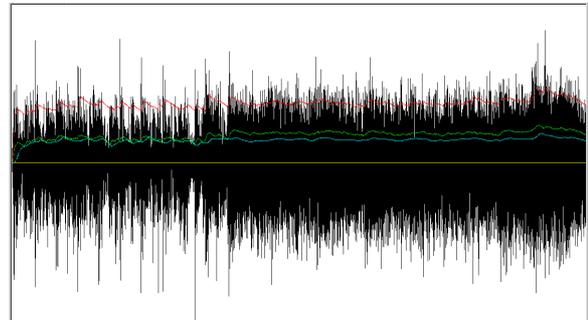


図-4 CM1

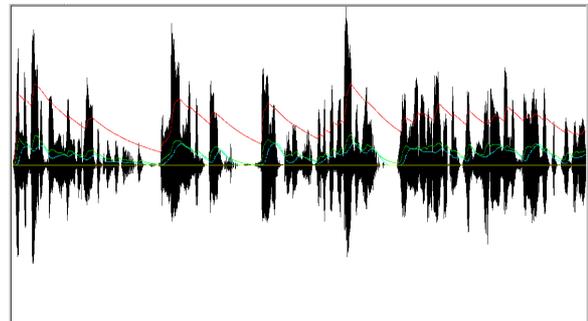


図-5 スピーチ 2

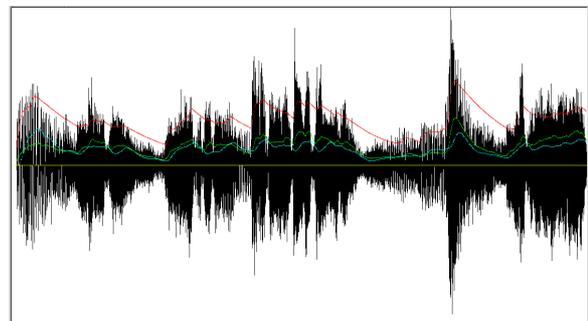


図-6 バラエティ

素材	VU 平均値	LOUD 平均値	-24LOUD	平均値 0VU
音楽	+5.3VU	-17.7LKFS	-1.0VU	-23.0LKFS
CM1	+2.2VU	-18.9LKFS	-2.9VU	-21.1LKFS
CM2	+1.5VU	-18.1LKFS	-4.4VU	-19.6LKFS
CM3	+4.3VU	-17.3LKFS	-2.4VU	-21.6LKFS
スピーチ 1	-4.9VU	-25.7LKFS	-3.2VU	-20.8LKFS
スピーチ 2	-5.5VU	-25.5LKFS	-4.0VU	-20.0LKFS
バラエティ	-3.5VU	-25.1LKFS	-2.4VU	-21.6LKFS

表-2 素材の解析結果

今回の素材解析例では、ラウドネスレベルを-24LKFS 前後にするための VU メータの振れは、①スピーチのように人の声を中心の場合は-3~-4VU 前後、②バラエティのようにスピーチと BGM が入っている素材は-2.5VU 前後、③CM のように COMP で固めた素材は+3~-4VU 前後、④ダイナミックレンジの広いポップス素材では-1VU 前後という結果になりました。

現在の制作手法で作業した場合、目標とするラウドネスレベルが-24LKFS ならば VU メー

タの平均レベルは-2~-3VU 以下にする必要があります。昔から「ナレーションの平均レベルは+3 から+4VU」と言われてきましたが、今議論しているラウドネスレベルにぴたりと当てはまりそうです。

少ない素材を使用した一例ですが、これからもいろいろな素材を使って検証してみたいと思います。

今回は、ラウドネスの測定要素と快適範囲 (Comport Zone) を中心に述べたいと思います。

[文献 1] ITU-R BS.1770-1, 2007

[文献 2] Audio Precision Application Note #5, Julian Dunn, 2001

【コラム】ITU-R の True-peak 計算

表は BS.1770-1 に記載されているオーバーサンプリング後の True-peak に対する Under-read の値、つまり、True-peak とオーバーサンプリング後の Sample-peak の理論的な誤差を示したものです。BS.1770 では Peak-sample という用語を使用していますが、ここでは本文と統一して Sample-peak を使います。表から分かるように 4 倍オーバーサンプリング (192kHz) を行っても $f_{norm}=0.50$ ($f=24kHz$) では最大 0.688dB の誤差を生じます。当然、48kHz での誤差はもっと大きくなります。

Over-sampling ratio (n)	Under-read(dB) Max $f_{norm} = 0.45$	Under-read(dB) Max $f_{norm} = 0.50$	16kHz 信号 最大誤差(dB) $f_{norm} = 0.3333$	20kHz 信号 最大誤差(dB) $f_{norm} = 0.4167$
4	0.554	0.688	0.3011	0.4736
8	0.136	0.169	0.0746	0.1168
10	0.087	0.108	0.0477	0.0746
12	0.060	0.075	0.0331	0.0518
14	0.044	0.055	0.0243	0.0380
16	0.034	0.042	0.0186	0.0291
32	0.008	0.010	0.0047	0.0073
1(fs=48k)	16.113	∞	6.0206	11.7400

表 オーバーサンプリング後の True-peak と Sample-peak の読取り誤差
(網掛けは 48kHz サンプリング時の最大誤差)

BS.1770-1 では下式を使って表の計算を行っています。

$$\text{maximum under-read (in dB)} = 20\log(\cos(\pi \cdot f_{norm} / n))$$

$$f_{norm} = \text{最大周波数} \div \text{サンプリング周波数} \quad (\text{例えば } 24\text{kHz} \div 48\text{kHz} = 0.50)$$

また、100% True-peak のレベルを 0dB TP とし、dBFS と区別しています。例えば -1dB TP とは 100% True-peak レベルから 1dB 下がったレベルのことで、-1dBFS とは意味が異なります。

表の網掛け部分は筆者が追加計算した 16kHz と 20kHz の最大誤差と、ratio 1 (=48kHz サンプリング) 時の f_{norm} 及び 16k、20kHz の最大誤差です。通常のピークレベルメータはサンプリング周波数のデータでピークレベル表示を行うので、ratio 1 の誤差を生じるおそれがあります。ただし、本文でも説明しているように 20kHz 信号はサンプリング周波数と整数倍の関係ではないので、サンプルポイントが固定されることはなく、ピークレベル表示上は大きな問題になりません。

BS.1770-1 にはオーバーサンプリングを行っていないピークレベルメータに対して、何 dB のマージンが適当なのか記述がありません。表から分かることは、ITU-R が推奨する 4 倍オーバーサンプリングを行った時、そのピークレベルから 1dB のマージン (-1dB TP) を取れば問題ないということです。ただし、これは通常のピークレベルメータに対し 1dB のマージンがあればよいという意味ではないので、誤解のないようにしてください。

米国 ATSC が発行した RP A/85 の中でも True-peak に対するマージンの記述があります。そこでは後処理も考えて -2dB TP が適当としていますが、やはり通常のピークレベルメータに対するマージンは記載されていません。ピークレベルメータで監視する場合のマージンに対する筆者の考えの一部については本文で述べているとおりです。