

[\[前へ\]](#) [\[次へ\]](#) [\[目次へ\]](#)

webを見ていると、たまにスペクトラムアナライザーを使用したコーデックやフォーマットの評価を見かけます。しかし残念ながら、これは非可逆圧縮オーディオコーデックの評価としては、全く意味がありません。これは可逆でないフォーマット全てに共通する事柄です。

非可逆圧縮の最大の目標はデータ量あたりの主観の質を高めることです。波形の形を整えることではありません。人の耳はフラットではないので、歪みが大きくても知覚しにくい場合があったり、その反面非常に微妙な差を聞き分けられることもあります。近代の非可逆音声圧縮エンコーダはそういった知覚特性を利用して主観の質を高めている訳ですが、それには人の聴覚・心理特性をベースに符号化の精度を部分部分で変えなければなりません。

実際に、波形レベルで明らかに問題が生じていても、知覚上明らかな問題として認識できない場合があったりしますし、逆に20m/secくらいに拡大した視覚的には同じような波形であっても、聴覚的にははっきりとまずく聞こえることもあります。このことは波形の観察が音質の評価の決定打にはなりえないことの証明でしょう。

スペクトラムアナライザーでの分析に関しても、単純に高域が出た・出ないというのは意味が無いです。人は目で音を感じることはできません。エンコーダやフォーマットの開発者もまた、綺麗な周波数分布に見えることを目標としません。例えば私的な話ですが、私がエンコーダを弄る時、自分にとって嫌な音を無くす、若しくは出難くすることを目標とします。もちろんそこにはビットレートと質との間のバランスを始めとする様々な取引がある訳ですが、そこで最大の頼りとなるものはあくまで自分の耳なのです。目で確認できるのはせいぜい完全に間引かれた周波数域くらいであり、音質を判断するのは不可能です。

閑話休題、一般的に言われる高域 (16kHz~)は実際にはそれほど重要ではありません。一般的な放送などでは15~16kHzまでのレンジしかありませんし、だからといって、それが原因で音を楽しめないという話はあまり聞きません。高域を真っ先に削るのは実際にそれが人には聞こえ難い領域であるからです。もちろん聞こえ方には個人差がありますが、高域の感度は鈍さは、ほぼ全ての人にとって共通です。もし低いビットレートで無理して高い周波数にビットを割けば、より感度の高い下方の帯域で嫌な音が聞こえてくる可能性は増大します(トータルのビットレートが同じだとして)。しかし、高域の感度が鈍いからと削りすぎれば、それはそれで聴覚的に嫌と感じる人が多く出てくるかも知れません。その辺の取引はエンコーダの重要な仕事の一つですが、それは聞いた時により良い結果(ビットレート対主観の質比)が得られるように調整されるべきです。決して波形の形やスペクトラムの見た目を保つために、ではないです。

カテゴリ: [\[音声圧縮\]](#) - [TrackBack\(0\)](#) - 2004年11月21日

[\[前へ\]](#) [\[次へ\]](#) [\[目次へ\]](#)