

サウンドファイルの紹介

wave, aiff, MP3

サウンドファイルとは

- AD変換: analog→digital
- MIDI系とPCM系: AD変換されたサウンドファイル
 - MIDI系.midはミディ関係
 - PCM系はミディ以外でwave, aiff, MP3など
- PCM pulse code modulation
 - 音をある時間間隔で,ある深さで数値化: 採取する毎秒回数Hz,と割り当てメモリー b
 - さらにチャンネル数がある。

サンプリングレート

- アナログデータの毎秒取得回数Hz
 - eg. 10回/sなら,10Hz
 - CDなら44.1kHz,DVDなら46kHz, 96kHz
 - レートが低いと籠もった音になる。
 - 電話:11kHz,AMラジオ:22kHz,FMラジオ:
33kHz,DAT(CDの原盤):48kHz

何故44.1kHzか

- 人間が認識できる周波数は200Hz~20kHz。それゆえ人間の可聴域20kHzを超えるサンプリングレートなら原音のアナログデータと区別がつかない。
- ナイキストの定理Nyquist'Theoremによると音波の再現性は音源の1/2の周波数
- → $20\text{kHz} \times 2 = 40\text{kHz}$ の音源が必要となる。さらに約10%の余裕。
 - なお, $1\text{Hz} = 1\text{cycle/s}$ を現す。単位を含んだ計算にはHzの代わりにcpsを使った方が便利。

量子化ビット(サンプリングビット)

- eg. 1 bitなら大小の2段階 3 bitsなら8段階。
- CDなどのサウンドファイルでは16 bitsつまり65,536段階。
- 16 bits は人間に粗さを感じさせない均衡点。waveファイルを直接編集する際にはエディタは24 bitsで内部処理。

チャンネル数←音の立体性

- 1チャンネルはモノーラル。2チャンネルはステレオ(LR)。
- 4チャンネルはサラウンド(LR+LsRs)。
- 5チャンネルは+C(センターチャンネル)
- 5.1チャンネルは+LFE(重低音用チャンネル Low Frequency Effect=woofer)
- 現在はさらに高度に(20以上)

サウンドファイルのサイズ

- サンプルングレート(1秒間の回数)
- × 量子化ビット bit (割り当てメモリ)
- × チャンネル数 (いわば本数)
- = $44.1 \text{ k(cps)} \times 16 \text{ (b/c)} \times 2 \text{ (channel)}$
- = $1.4 \text{ M (bps)} = 1400 / 8 \text{ (kB/s)} = 175 \text{ (kB/s)}$ ただし, $1 \text{ byte (B)} = 8 \text{ bits}$
- 例えば, この 1.4 Mbps を, この操作の **ビットレート** と呼ぶ。
- さて, 1分間あたり, $175 \text{ kB/s} \times 60 \text{ s} = 10 \text{ MB}$ のメモリを使用。

MP3ファイル

- MPEG1 Audio Layer-IIIの略称。
- 最も一般的なサウンドファイル。
- 欠番のMPEG3とは関係ない。
- 高い音質を維持しながらデータサイズが小さい。
- CD相当の音質とされる128kbpsで圧縮した場合のファイルサイズは元(1400)の1/10

Bit rateとMP3の音質

音源	ビットレート kbps	圧縮率
CD	128	1/10
FMラジオ	64	1/20
AMラジオ	32	1/40
電話	16	1/80

MP3の限界

- 音が籠もっている。
- 理論上は聞こえない筈の音であっても音質の悪さを感じることもある。
- 不要なと思われるデータが間引かれている。
- エンコード(ファイル変換)によってデータ自体を変質させるので非可逆圧縮。デコード(復原)できない。

注記: codec

- Coder Decoderの略。エンコードエンジンと呼ばれる。
- これで一定の規則に従って、信号やデータを解析し、原音の間引きと再生をする。MP3は音楽圧縮コーデックの代表例。

参考：ヘルツとサイクル

- ヘルツHz: 単位利用の取り決めでは1997年以来, 次のcpsに代わった。Hzには「毎秒」も組み込まれているので異なる単位の計算をする際に不便。
- サイクルcps: 周波数の単位。電気や音などの振動現象で, 1秒間に同じ位相がめぐってくる回数をサイクル毎秒(記号c/s, cps)または単にサイクル(記号c)で表わす。後者ではHzの問題点が改善されないから, cpsを使った方がいいだろう。1 Hz = 1 cpsとする。cpsにはcharacter per second(4 bytesを1単位)とする使い方もあり, 注意が必要。

参考：音量（正確には音強）

- 音波の振幅の大きさ。周囲との圧力のずれ。
- 周辺との音圧差(音圧レベル L_p)であり,通常デシベルdBで表す。下の式で p (Pa)は音圧のRMS(二乗平均平方根)、 p_0 は基準となる音圧 ($20\mu\text{Pa}$)。音声ファイルのサイズとは関連がない。

$$L_p = 10 \log_{10} \left(\frac{p^2}{p_0^2} \right) = 20 \log_{10} \left(\frac{p}{p_0} \right) \text{ dB}$$