

オーディオデータを近似的に圧縮 しよう

SuperCon2017 本選問題解説

大阪大学サイバーメディアセンター

背景

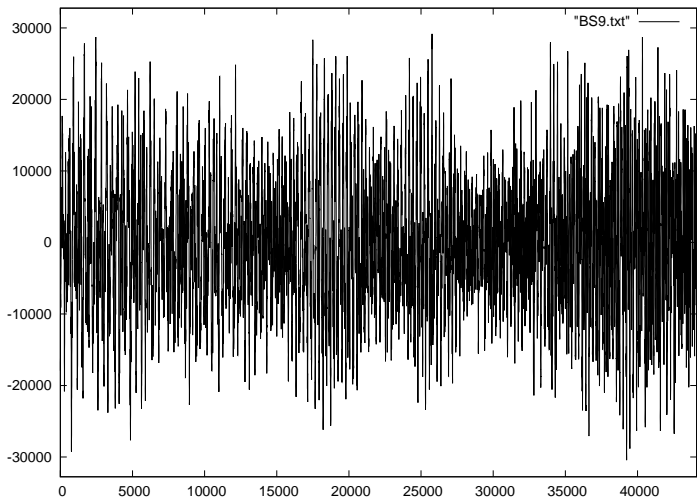
CDなどのデジタル・オーディオは元となるアナログ信号の音量を短い時間間隔で「サンプリング」して、整数のデータにしている (Pulse Code Modulation)。

- 音楽のデータは整数が並んだ数列として表現される
 - 元のアナログ信号は時間も音量も連続だが、デジタル信号は時間も音量も飛び飛び

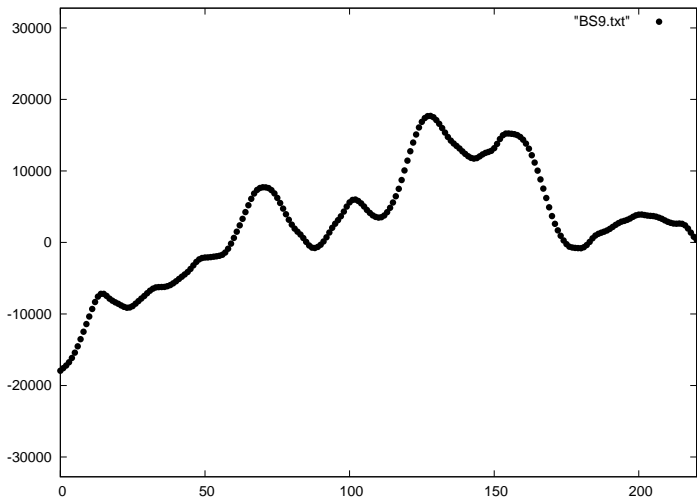
デジタル・オーディオの基礎概念

- サンプル周波数: 1秒間の信号を何個のデータにするか。CDでは1秒を44100に分割(サンプル周波数44100Hz)
- ビット深度: 音量を何ビットの整数で表すか。CDでは符号ありの16ビット整数(-32768から+32767。無音は0)

どちらも大きいほどアナログ信号を忠実に再現できるが、CDでは人間の可聴域や分解能を考慮してこの数値にしてある。



1秒間のオーディオ信号の例 (フルトヴェングラー
指揮、ベートーヴェン交響曲「第九」より)



0.005 秒間 (1/44100 秒ごとに 16 ビットの整数がひとつあることがわかる)

目標

音楽の PCM データを簡単な方法でなるべく精度よく圧縮したい。

- MP3 や AAC などの方式が知られているが、敢えて別の方法を考えてみよう

1ビット・エンコーディング

音量の**変化**を前の時刻に比べて増えるか減るかの1ビットデータで表す

元の PCM データをなるべく正確に表現する1ビット・データを作ろう!

設定

エンコード (符号化)

- 元の PCM データ列

$$P_i, i = 0, \dots, N - 1$$

- エンコードされた 1 ビットデータ列

$$B_i, i = 0, \dots, N - 2$$

B_i は 0 または 1 のいずれか。1 は次の時刻で音量が増えること、0 は減ることを表す

デコード (復号)

- B からデコードされるデータ

$$S_i = P_0 + A \sum_{k=0}^{i-1} (2B_k - 1)$$

A は任意に設定した「増減幅」を表す整数

- **クリッピング**: S_i が -32768 以下になる場合は $S_i = -32768$ 、 $+32767$ 以上になる場合は $S_i = +32767$ となる (それ以前の履歴は忘れる)。

デコード (続き)

- **平滑化**: 注目している時刻とその前後5点ずつの計11点のデータを合計する (これを11で割れば平均を取ったことになる)

$$\bar{S}_i = \sum_{k=i-5}^{i+5} S_k$$

この操作はおおむね 4000Hz 以上の高周波をカットすることに相当する (高周波のノイズを除去する)

近似のよさの定義

- 元のPCMデータも平滑化

$$\bar{P}_i = \sum_{k=i-5}^{i+5} P_k$$

- 平滑化されたデータの差の絶対値の合計

$$\Delta = \sum_{k=5}^{N-6} |\bar{P}_k - \bar{S}_k|$$

を「近似のよさ」とする

問題

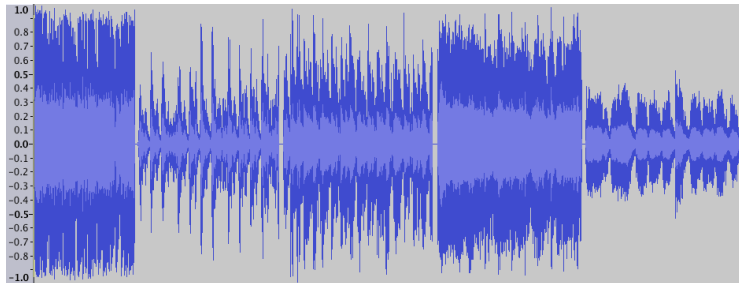
- 問題として PCM データが与えられる
 - CD と同様、44100Hz, 16 ビットの PCM データ
 - データ点の数 $N \leq 5292000$ (2 分以内)
- これに対して、 Δ ができるべく小さくなるような 1 ビットデータ列 B と増減幅 A を求めるためのプログラムを作る
 - 出力は A と B
 - A は 32767 以下の正の整数

勝利条件

- 問題を 5 問解く
 - 計算時間 (経過時間) の上限は合計 10 分
- 問題ごとに Δ が小さい順に 1 から 5 位を決める
- 問題ごとに 1 から 5 位に 5 から 1 点を与え、5 問の点数の合計で総合順位を決定

本選使用データ

- ① ryo 「ワールドイズマイン」 (30 秒)
 - 初音ミク
- ② 伊福部昭 「ゴジラ」 (34 秒)
 - オリジナルサウンドトラック
- ③ ザ・ビートルズ 「サージェント・ペパーズ・ロンリー・ハーツ・クラブ・バンド」 (1 分)
- ④ 相馬民謡 「相馬二遍返し」 (2 分)
 - ZABADAK
- ⑤ ムソルグスキー 「展覧会の絵」 (ラヴェル編曲) より 「キエフの大門」 (1 分 58 秒)
 - フリッツ・ライナー指揮シカゴ交響楽団



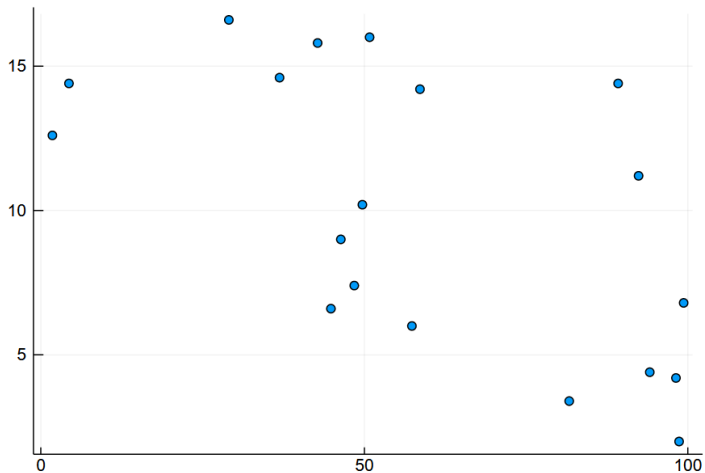
曲による波形の違い: 曲によってデータの性質は
ずいぶん違う



今回は阪大サイバーメディアセンターのSX-ACE
を使う

- **マルチコア・ベクトル並列型計算機**
- 1 CPU/node, 4 core/CPU
 - 各コアにベクトル演算器とスカラー演算器を持つ
 - ノード性能 256GFlops, メモリー 64GB
- **20 チームで 21 ノードを占有使用**
 - 1 ノード (4 並列) で実行できるプログラムを作る

上手なベクトル化が鍵!



平均ベクトル化率 vs. 平均順位