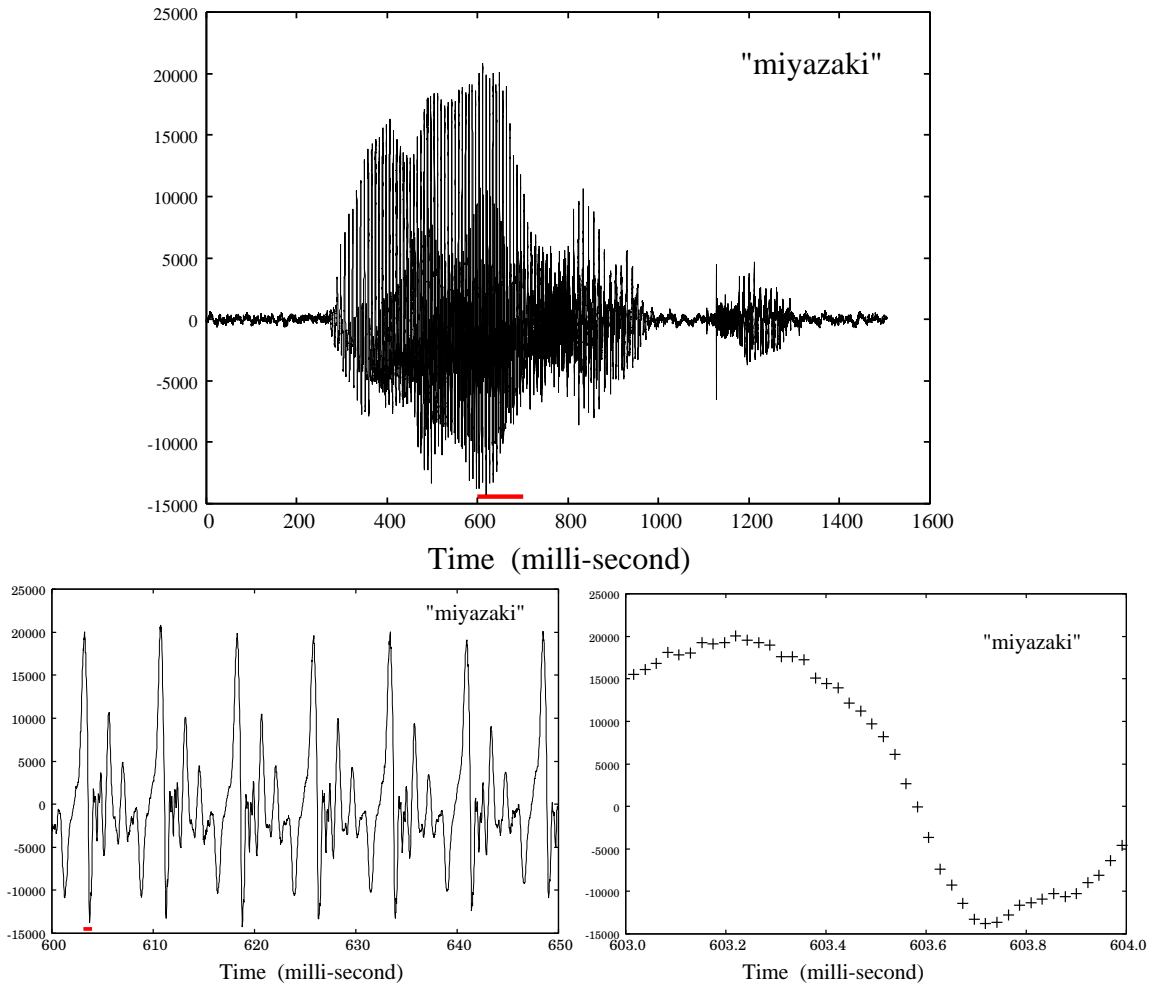


## パターン情報処理：音，画像

- アナログ信号 デジタル信号 (AD 変換): 標本化 (サンプリング) と量子化
- 音，音声  
波形の例 (「みやざき」, 約 1.5 秒間のデータ)



サンプリング周波数：44.1 kHz (1 秒間に 44100 個のサンプル点 → 横軸)  
量子化ビット数 (深度) 16 bit (2byte). (縦軸の取り得る値： $2^{16} = 65536$  通り.)  
1 チャンネル (モノラル)  
1 秒間のデータ量  $\Rightarrow 44100 \times 16\text{bit}(2\text{byte}) \times 1\text{channel} = 705,600\text{bit} = 88.2\text{kbyte}$

% ls

```
-rw-r--r-- 1 date date 132682 2009-07-12 18:40 miyazaki001.wav
```

- それぞれの人の音波形「みやざき」の共通点と違いは何か【自由研究課題】  
音を聞いて (音の波形を見て) → 発せられた「単語」が判別できる。  
音を聞いて → 「誰」の声が判別できる。  
 $\Rightarrow$  波形「みやざき」の本質は何か。

- CD (CD-DA: Compact Disc Digital Audio, 1981 年)

サンプリング周波数 : 44.1 kHz .

量子化ビット数 (深度) 16 bit. ( $2^{16} = 65536$  通り)

2 チャンネル (ステレオ)

1 秒間のデータ  $\Rightarrow 44100 \times 16\text{bit} \times 2\text{channel} = 1,411,200\text{bit} \approx 1.4 \text{ Mbps}$  (bit per second).

1 分間のデータ  $\Rightarrow 60\text{sec} \times 44100 \times 2\text{byte} \times 2\text{channel} = 10,584,000\text{byte}$ .

74 分 42 秒 :  $10.584 \times 74.7 = 790.6248 \text{ Mb}$

- MP3 (MPEG Audio Layer-3)

標準とされる品質 : 128 kbps (これに対し CD は 1.4 Mbps) .

約 11 倍 , 圧縮されている  $\Rightarrow$  どのように情報を圧縮しているか

(考え方)

各サンプル点の振幅の値を直接記録しない .

(録音の仕方)

もとの波形を , ある短い時間ごとの波形  $f^0, f^1, \dots$  (例えば  $m = 1024$  サンプル点ごと) に分ける . 分けられた個々の波形の断片  $f^\alpha$  ( $m$  次元ベクトル) を , 様々な周波数の正弦波  $e_1, e_2, \dots, e_n$  ( $m$  次元ベクトル ,  $n$  は  $m$  より小さい) の重み付けの和で表現することを試みる【線形代数 . フーリエ級数】 .

$$f^\alpha \approx \sum_{i=1}^n x_i e_i$$

ともかく係数  $x_1, x_2, \dots, x_n$  を計算し , この  $n$  次元ベクトル  $x^\alpha = (x_1, x_2, \dots, x_n)$  を記録する . この処理をすべての断片について実行し ,  $x^0, x^1, \dots$  を記録する .

$n < m$  でこれを実現すれば , データを圧縮して記録できる (情報は劣化する) .

(再生時)

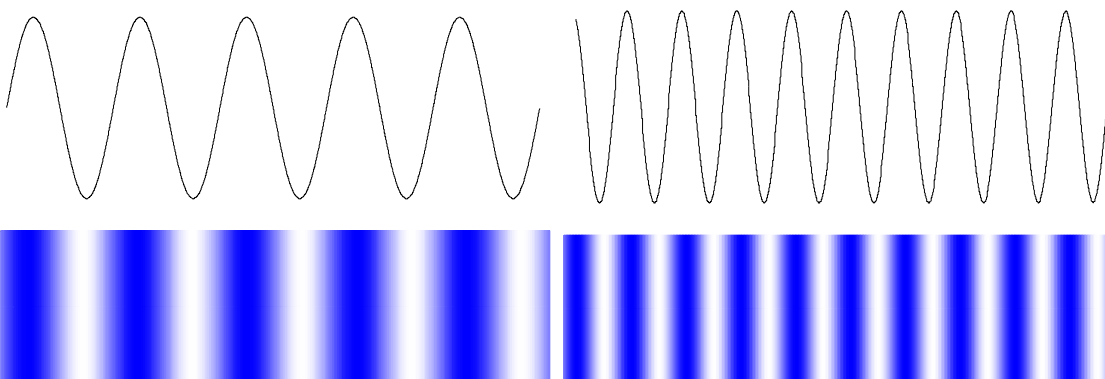
各音の断片について係数  $x_0, x_1, \dots, x_n$  を読み ,  $\sum_{i=1}^n x_i e_i$  を計算し ,

波形 (短い時間)  $f^\alpha$  の復元を試みる . これの繰り返し .

- 音と画像

音 : 1 次元の信号 . 画像 : 2 次元の広がりをもつ信号 .

$\Rightarrow$  「信号処理」の立場から見ると共通点が多い .



- 画像のデジタル表現

- ・ピクセル (画素) . 画像の最小単位 .
- ・RGB カラー画像 . 赤 (R) , 緑 (G) , 青 (B) の三色の光の点 .  
各 RGB の明るさを指定  $\implies$  どんな色でも表現
- $\iff$  グレースケール画像 . 色に関する情報を含まない画像 .

- 動画 : 連続した静止画像を高速に映し出している . (パラパラ漫画)

映画 24 枚/1 秒 . テレビ 30 枚/1 秒 .

- アナログ信号 デジタル信号 (AD 変換) : 標本化 (サンプリング) と量子化

例 : DVD (Digital Versatile Disk , 1995 年)

画素数  $720 \times 480$  pixel . 階調 : 24 bit (RGB 各 8bit) . フレーム数 : 30 枚/秒 .

- ・1 秒間のデータ :

$\implies 720 \times 480 \times 24 \times 30 = 248,832,000 \text{ bit} = 31,104,000 \text{ byte} \approx \underline{31.104 \text{ Mb}}$  .

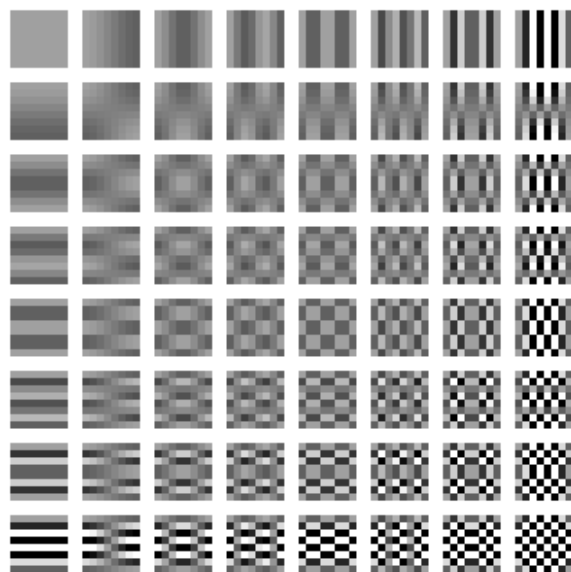
- ・DVD : 約 4.7Gb (giga byte) のデータを保存 .

$\implies 4700\text{Mb} / 31.104\text{Mb} \approx 151 \text{ 秒/DVD1 枚}$  (動画像を情報圧縮しないで保存した場合)

- データをいかに圧縮するか (切実な問題) .

$\implies$  表現したい画像を小さいブロックに分割 ( $f^\alpha$  , 例えば  $8 \times 8$  ピクセル) .

少数の土台画像 ( $e_i$  , 基底) の重ね合わせで各ブロック画像を再構成し表現 .  $f^\alpha = \sum_{i=1}^n x_i e_i$



- 情報圧縮の先端的研究 (スパース符号化) : 考え方

一つの問題 (画像の断片) に対して , 少数 (の基底) が対応する (全員は動員されない) .