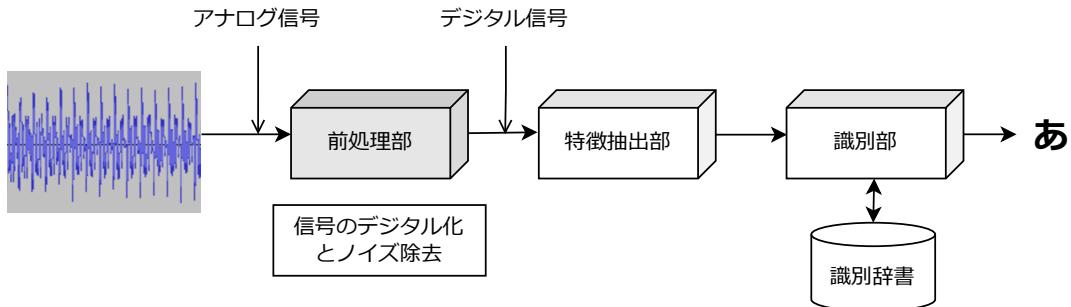


2. データをきちんと取り込もう



- 2.1 アナログ信号のデジタル化
- 2.2 人の知覚に近づける
- 2.3 特徴抽出をしやすくする処理

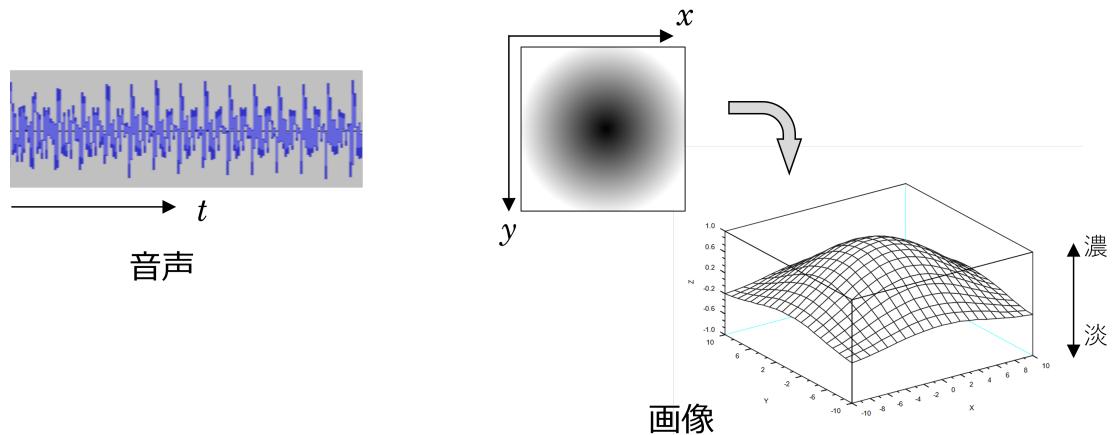


- 荒木雅弘:『フリーソフトでつくる
音声認識システム(第2版)』(森北
出版, 2017年)
- スライドとJupyter notebook
- サポートページ

2.1 アナログ信号のデジタル化

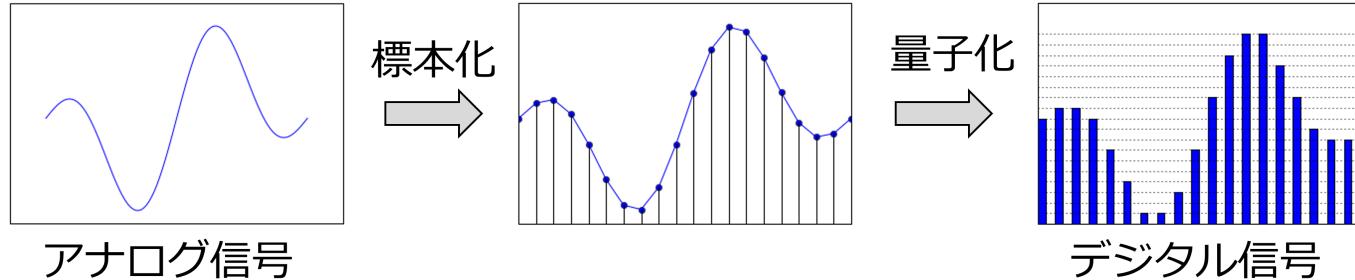
2.1.1 アナログ信号は波である

- 波としてのパターンの表現



2.1.2 標本化と量子化 (1/3)

- 波をデジタル化する手順
 - 標本化: 一定間隔で波をサンプリング
 - 量子化: 離散値に丸める
- 目標
 - なるべく情報を落とさずに、かつ、なるべくコンパクトに



2.1.2 標本化と量子化 (2/3)

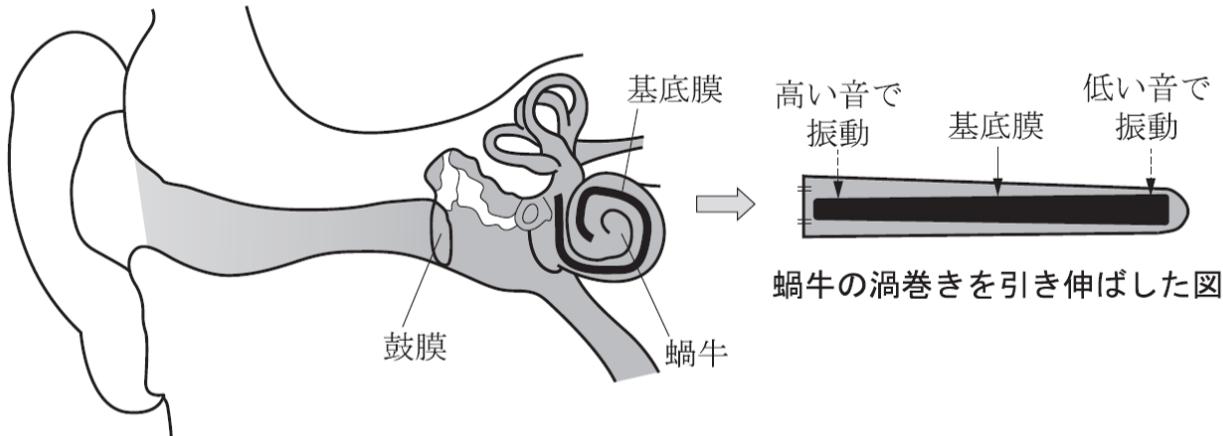
- 標本化
 - 時間または空間に連続して分布している信号を、離散的な観測点で代表させる
- 標本化定理
 - 元の波に含まれる周波数の最も高いものを f としたときに、 $2f$ より高い周波数で標本化すれば、元の波を完全に再現できる
 - 例) 人間の可聴範囲は 20 Hz ~ 20,000 Hz 程度
⇒ コンパクトディスク (CD) は 44,100 Hz で音を標本化

2.1.2 標本化と量子化 (3/3)

- 量子化
 - 連続値を取る信号強度を、有限の離散値で近似
 - 人間の識別能力を基準にする
 - 例) 聴覚のダイナミックレンジはほぼ100dB
= 聞き取れるもっとも小さな音の約100万倍の大きさまで聞こえる
 - よく使われる量子化ビット数: $16\text{bit} = 65,536\text{段階} \approx 96\text{dB}$
 - デシベルの定義: 教科書 演習問題2.1参照

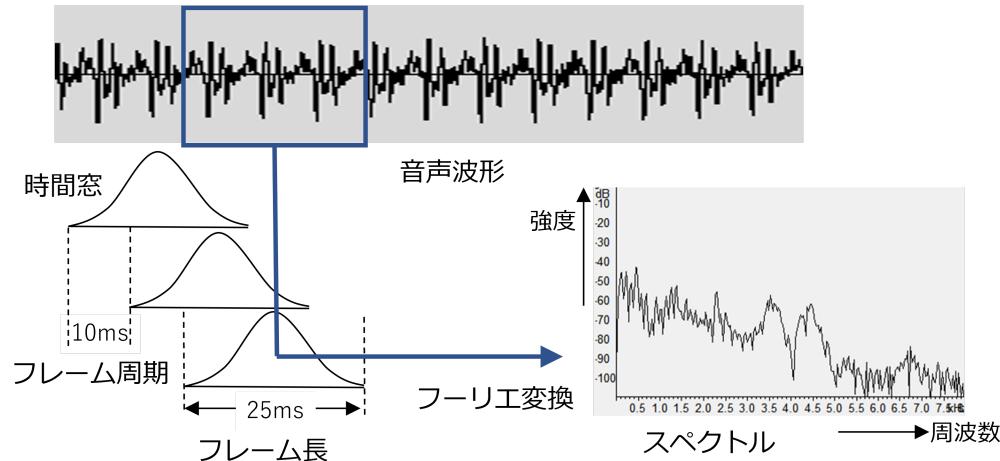
2.2 人の知覚に近づける (1/5)

- 音声の知覚
 - 音は空気の粗密波
 - 鼓膜を振動させる → 蝸牛内の基底膜で共振周波数をピックアップ



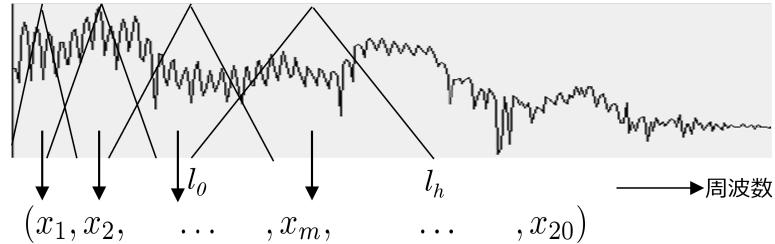
2.2 人の知覚に近づける (2/5)

- 音声の知覚原理を活用
 - 音声信号をフーリエ変換し、ピークとなる周波数を見つける



2.2 人の知覚に近づける (3/5)

- 音声の知覚原理を活用
 - 低い音ほど周波数分解能が高い → メルフィルタバンクの適用

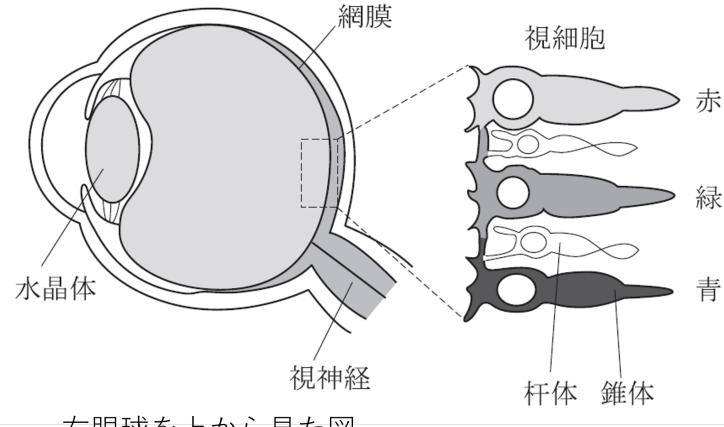


- 各帯域の振幅スペクトル値
 - m 番目の三角窓関数 W_m にパワースペクトル S をかけたものの帯域内周波数についての和

$$x_m = \sum_{k=l_0}^{l_h} W_m(k) |S(k)|$$

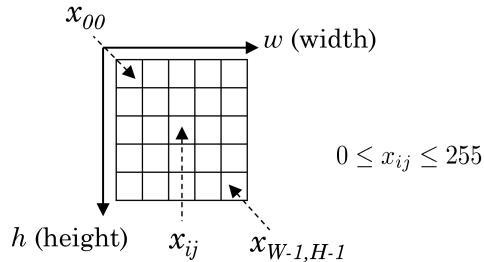
2.2 人の知覚に近づける (4/5)

- 画像の知覚
 - 光の波長によって感度が異なる赤錐体・緑錐体・青錐体が脳に信号を伝えている



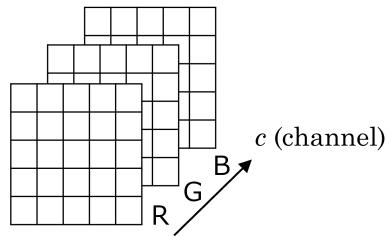
2.2 人の知覚に近づける (5/5)

- 画像の知覚原理を活用
 - 光の強さを感じる網膜の細胞を2次元配列で表現



0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9

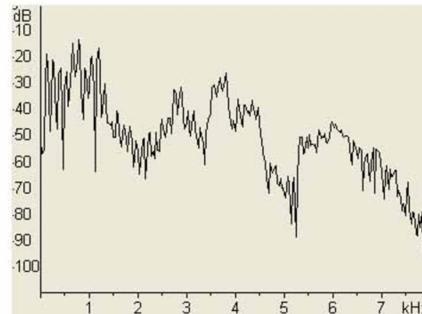
濃淡画像



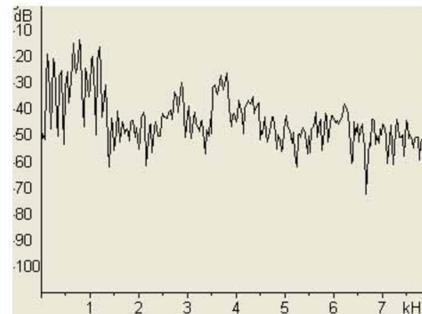
カラー画像

2.3 特徴抽出をしやすくする処理 (1/5)

- 音ノイズの除去
 - 背景雑音(加法性):周波数空間で引き算
 - マイクの特性(乗法性):周波数の対数空間で引き算



(a) 「あ」のスペクトル

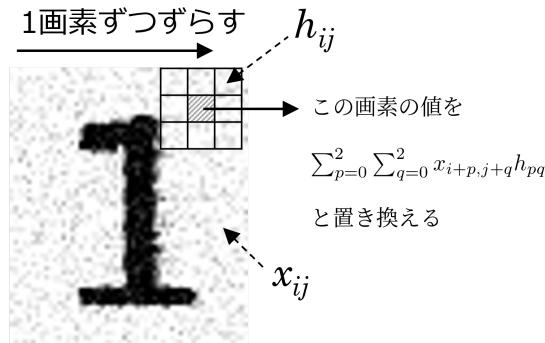


(b) 白色雑音を加えたスペクトル

加法性雑音の例

2.3 特徴抽出をしやすくする処理 (2/5)

- 画像ノイズの除去
 - フィルタの適用
 - 特定の画像入力に反応する脳の視覚野領域の処理に対応



1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

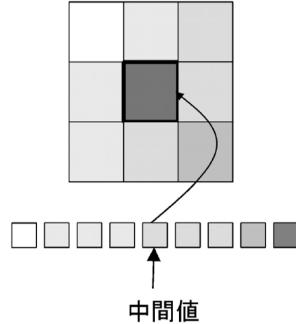
平均値フィルタ

-1	0	1
-1	0	1
-1	0	1

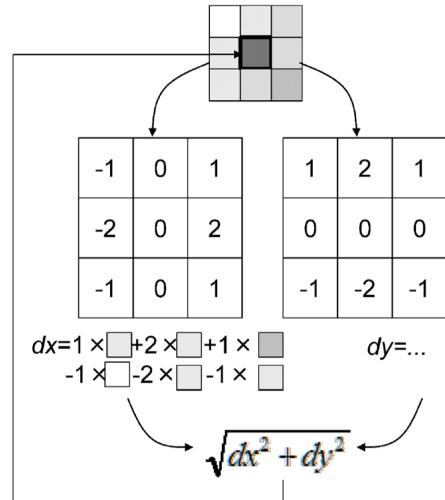
(縦) エッジフィルタ

2.3 特徴抽出をしやすくする処理 (3/5)

- さまざまな画像フィルタ



(a) メディアンフィルタ



(b) Sobelフィルタ

2.3 特徴抽出をしやすくする処理 (4/5)

- 画像フィルタ適用上の注意点
 - 周辺部の画素の処理
 - 0-padding: 元画像の端の画素にもフィルタが適用できるように画像を広げて、フィルタの適用前後で画素数を変えないようにする
 - 上記の処理を行わない場合、フィルタの適用により画素数は小さくなる
 - ストライドの設定
 - フィルタをずらす画素数(ストライド)を2以上に設定すると、画像を圧縮することになる

2.3 特徴抽出をしやすくする処理 (5/5)

- ・ メディアンフィルタ適用の結果



適用前



適用後

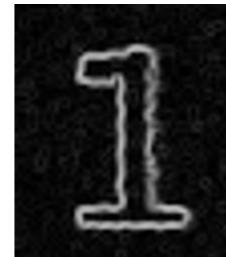
- ・ Sobelフィルタ適用の結果



元画像



ノイズ付加



Sobelフィルタ適用後

まとめ

- 前処理部の役割
 - アナログ信号のデジタル化
 - 標本化:一定間隔で波をサンプリング
 - 量子化:離散値に丸める
 - 後の特徴抽出のために人の知覚に近づけておく
 - 特徴抽出を容易にする処理
 - ノイズ除去
 - 音声はノイズを引き算できる形に変形する
 - 画像はフィルタを用いる
- Jupyter notebook