

5. 音声からの特徴抽出

5.1 特徴抽出の手順

5.2 音声信号のディジタル化

5.3 人の聴覚をまねて – スペクトル分析

5.4 もうひと工夫 – ケプストラム分析

5.5 雑音の除去

5.1 特徴抽出の手順

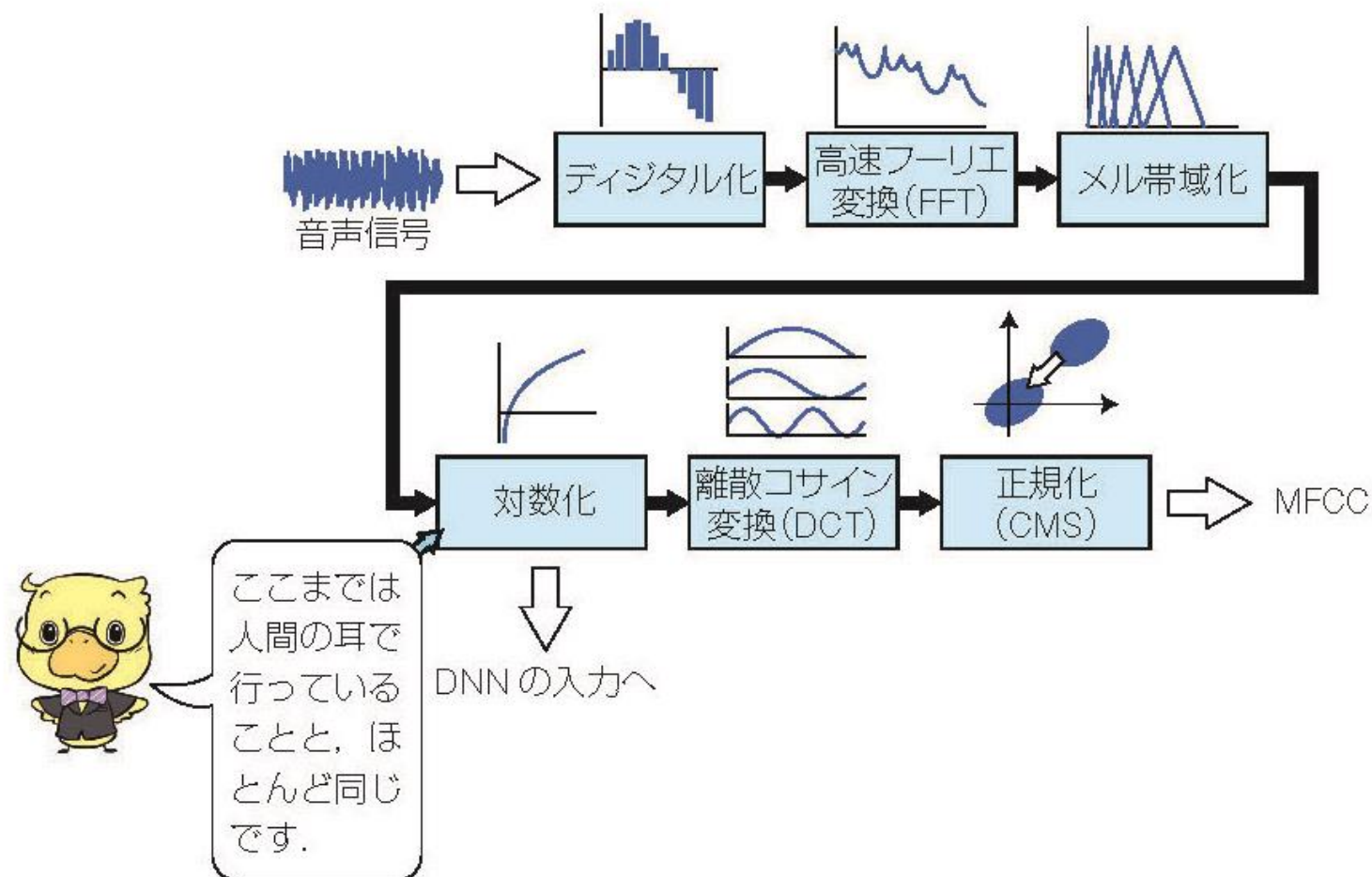


図 5.3 音声の特徴量を求める手順

5.2 音声信号のディジタル化

- 波のディジタル化
 - 標本化：時間軸方向の分割
→ 標本化定理
 - 量子化：強度方向の分割

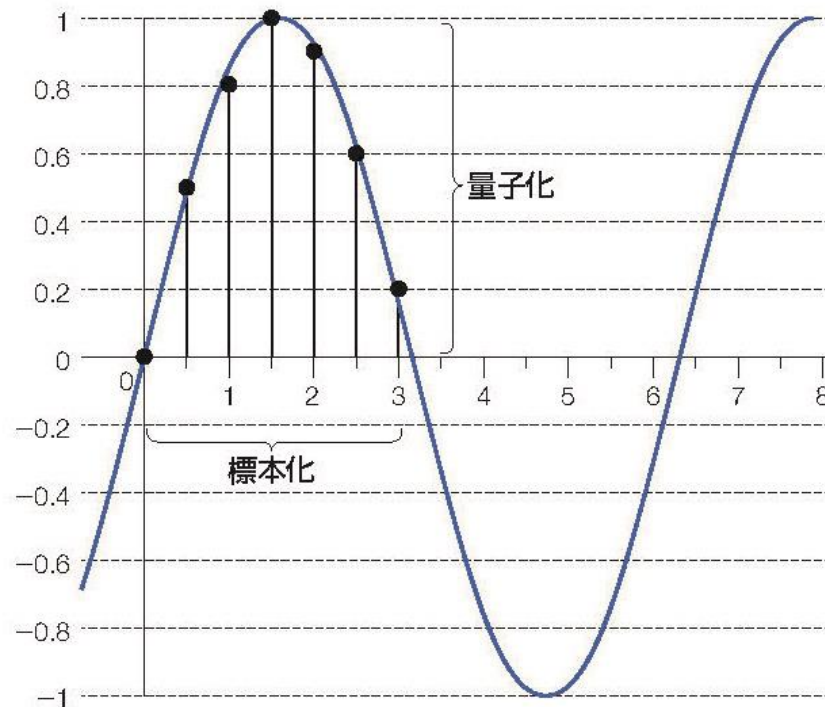


図 5.4 標本化と量子化

5.3 人の聴覚をまねて – スペクトル分析

- 連続信号の分割：フレーム化

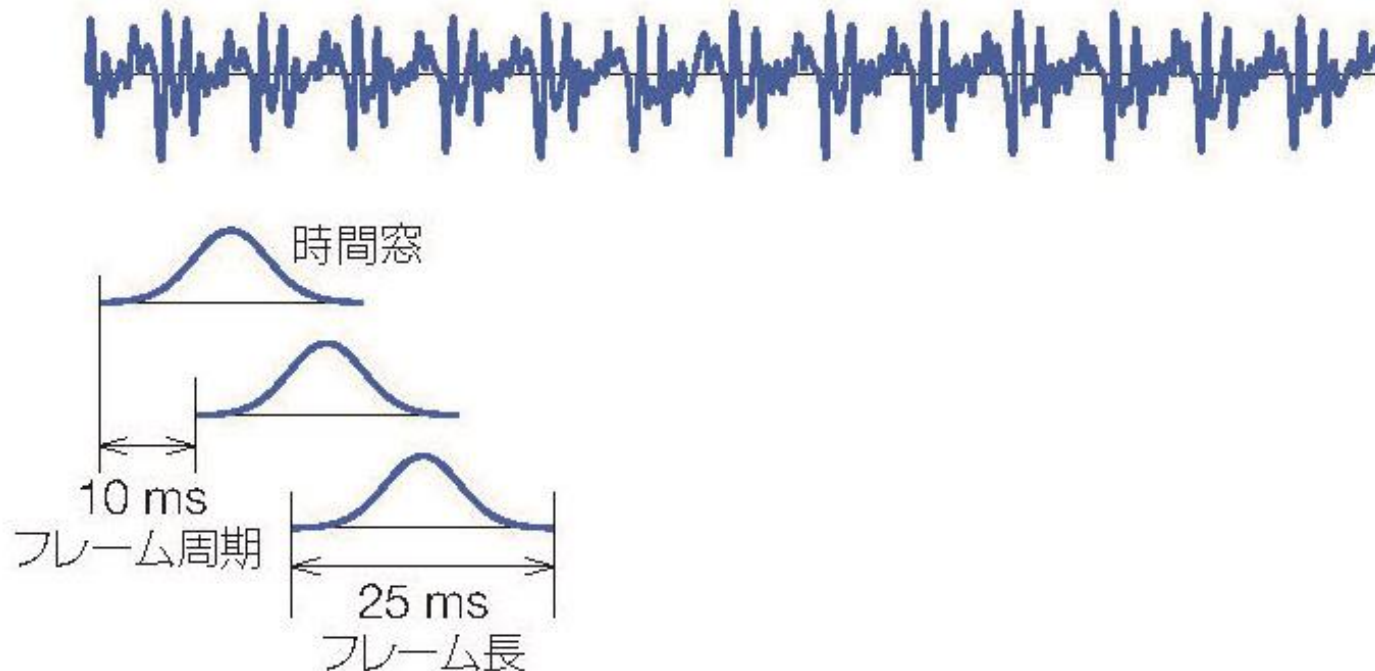


図 5.5 音声信号のフレーム化

5.3 人の聴覚をまねて – スペクトル分析

• スペクトル分析

- フレームとして切り出した音声信号をフーリエ変換し、パワースペクトルを計算
- 低周波数ほど周波数の違いに敏感という人間の知覚を反映したメルフィルタバンク処理を行う

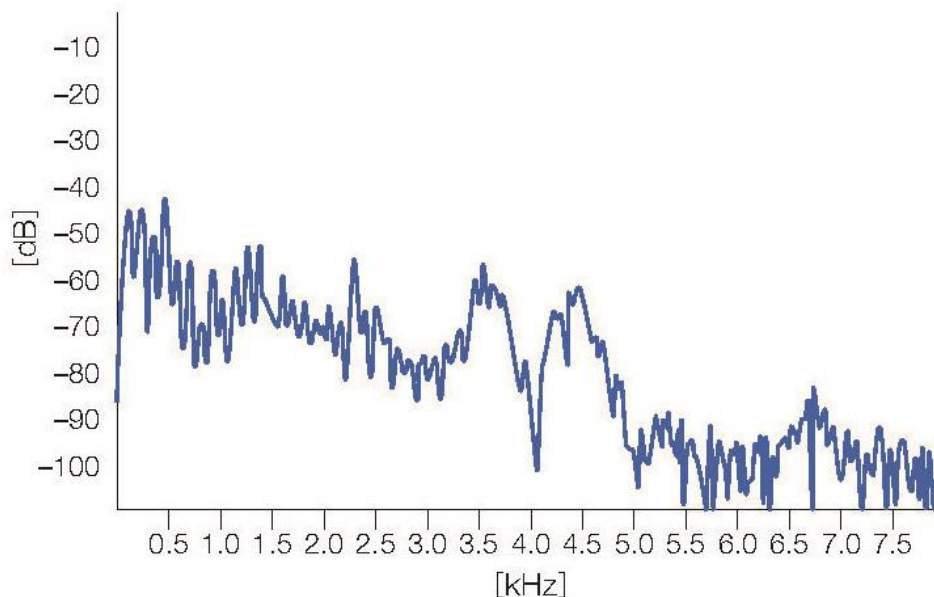


図 5.6 パワースペクトル

5.4 もうひと工夫 – ケプストラム分析

- メルスペクトルの概形を抽出
 - 離散コサイン変換でケプストラムを計算
 - ケプストラムの低次情報がスペクトル概形に相当
→ MFCC (mel-frequency cepstral coefficient)

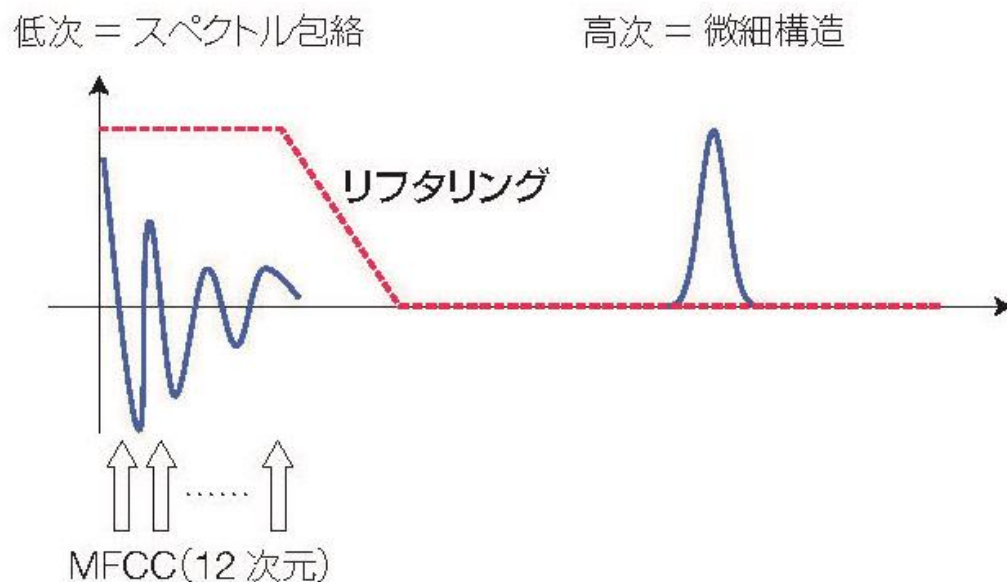


図 5.9 MFCC の抽出

5.4 もうひと工夫 – ケプストラム分析

- 変化量の抽出
 - Δ MFCC: 前後2フレームのMFCCから傾きを抽出
 - $\Delta\Delta$ MFCC: Δ MFCCの変化量を抽出
- 特徴量としての音声のパワー
 - 単純なパワー（声の大きさ）は特徴としては不適
 - Δ パワー、 $\Delta\Delta$ パワーは有効な特徴

5.5 雑音の除去

- 雑音の種類
 - 加法性と乗法性

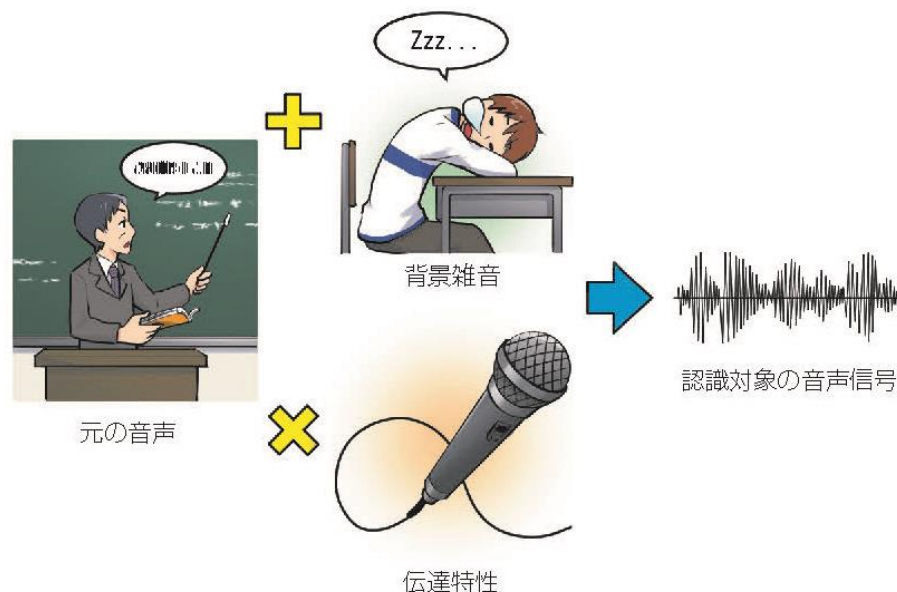


図 5.11 雑音の種類

- 雑音除去の方法
 - CMS (cepstrum mean subtraction)
 - 発話全体のケプストラム平均を求め、各フレームのケプストラムから引く