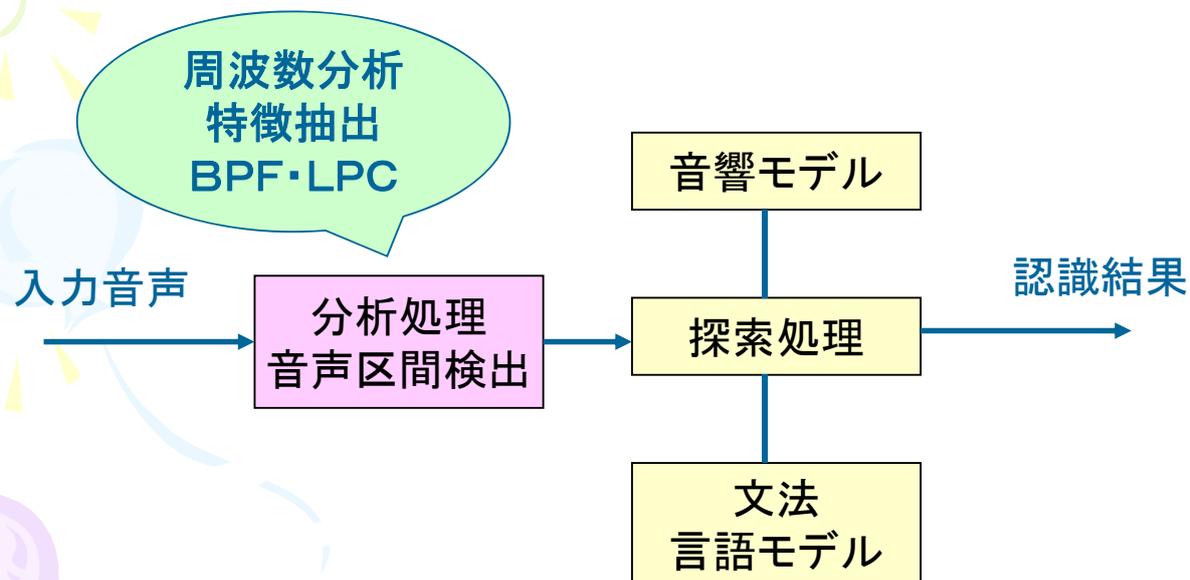


音声認識システムの構成 (2) 再掲



音声認識

Copyright © by Takeshi Kawabata

技術各論 > 特徴抽出

- 声道の形状

⇒

⇒

- 音声(時間波形)

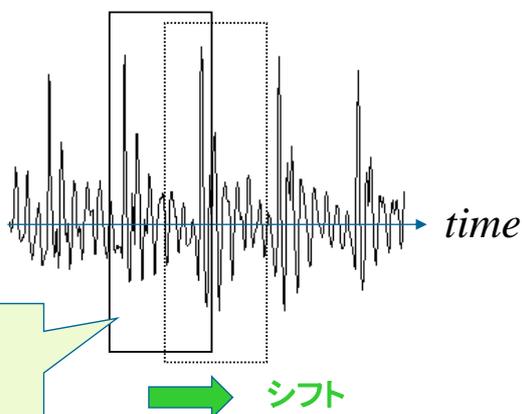
⇒

⇒

音声認識

Copyright © by Takeshi Kawabata

技術各論 > 特徴抽出 (2)



音声認識

Copyright © by Takeshi Kawabata

技術各論 > 特徴抽出 (3)

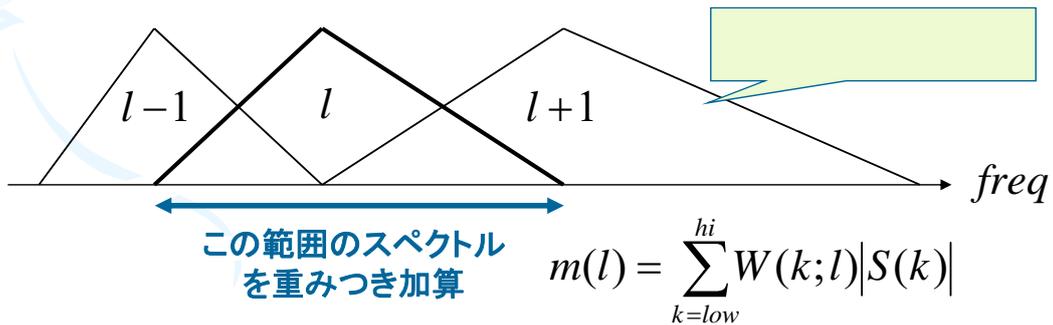
- 音声認識に用いられる特徴量

音声認識

Copyright © by Takeshi Kawabata

技術各論 > 特徴抽出 (4)

- MFCC (Mel-Frequency Cepstrum Coefficients)



音声認識

Copyright © by Takeshi Kawabata

技術各論 > 特徴抽出 (5)

- MFCC (Mel-Frequency Cepstrum Coefficients)

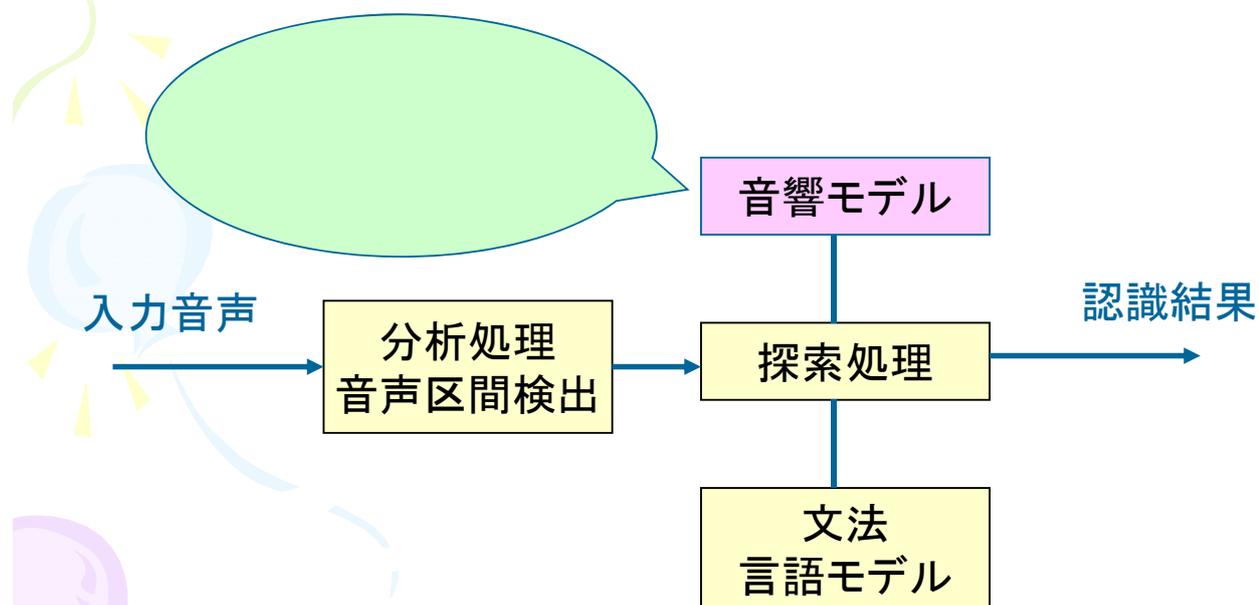
$$mfcc(i) = \sqrt{\frac{2}{N}} \sum_{l=1}^L \log m(l) \cos \left\{ \left(l - \frac{1}{2} \right) \frac{i\pi}{L} \right\}$$

離散コサイン変換: 離散フーリエ変換の一種、情報圧縮率が高い

音声認識

Copyright © by Takeshi Kawabata

音声認識システムの構成 (3)



音声認識

Copyright © by Takeshi Kawabata

技術各論 > 音響モデル

音声認識

Copyright © by Takeshi Kawabata



技術各論 > 音響モデル (2)

• 統計的音響モデル



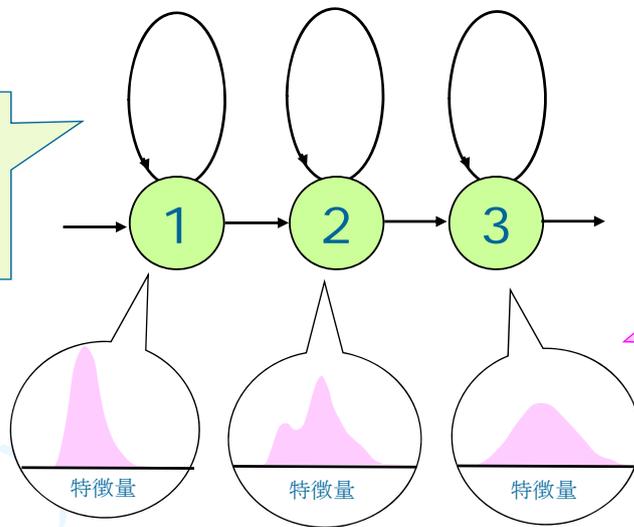
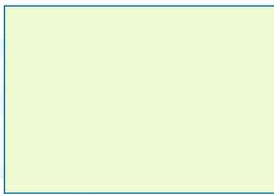
音声認識

Copyright © by Takeshi Kawabata



技術各論 > 音響モデル (3)

• HMM



音声認識

Copyright © by Takeshi Kawabata

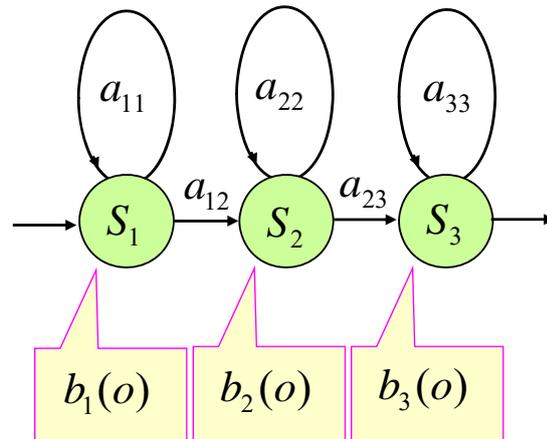
技術各論 > 音響モデル (4)

- HMMの構成要素

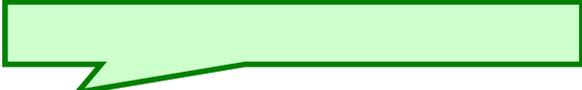
$$\Sigma = \{S_i\}$$

$$A = \{a_{ij}\}$$

$$B = \{b_i(o)\}$$



技術各論 > 音響モデル (5)


$$b_i(o) = \ln \left\{ \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left(-\frac{(x - \mu_i)^2}{2\sigma_i^2} \right) \right\}$$
$$= -\ln \sqrt{2\pi}\sigma_i - \frac{(x - \mu_i)^2}{2\sigma_i^2}$$



技術各論 > 音響モデル (6)

- HMMの学習 (A, B の決定)



「

」



音声認識

Copyright © by Takeshi Kawabata



技術各論 > 音響モデル (7)

- HMMによる音声認識



「

アルゴリズム」による高速近似計算



音声認識

Copyright © by Takeshi Kawabata

ここまでのポイント

1. 技術各論＞特徴抽出

- MFCC

2. 技術各論＞音響モデル

- 隠れマルコフモデル

(HMM, Hidden Markov Model)