

### 1.はじめに

現在、話者認識に用いられる手法の1つとして隠れマルコフモデル(Hidden Markov Model:HMM)法がある。HMMを用いた話者認識を行う際に認識率に影響を及ぼす要因の1つとして、学習時にどのような音素単位を用いてモデルを作成するかということが挙げられる。

従来、音声認識において半音節を単位とするHMMが用いられている。半音節単位を用いることで、従来のCVCあるいはVCV単位よりも少ないデータ量で学習が行えるという利点も挙げられ、小規模学習データに適した認識方法である。また、音素単位を用いるよりも調音結合を考慮したモデルの作成が可能である。本研究では、半音節単位を各話者ごとのモデル作成に用いて半音節HMMを作成することにより話者認識を行う方法を提案した[1]。

### 2.半音節HMM

半音節HMMは、原音声を半音節単位でセグメンテーションした半音節単位から作成された初期モデルから連結学習により作成した。

半音節HMMは、Left-to-Right型のHMMモデルで与える。状態数1とし、出力確率は、無相関正規分布(各次元互いに独立なガウス分布)で与える。

### 3.連結学習による話者認識

半音節HMMを用いたテキスト指定形話者認識は、各話者の学習データを用いて各話者の半音節モデル(半音節HMM)を作成する。

図1に半音節HMMを用いた連結学習による話者認識法を示す。認識時には、指定されたテキストに基づき、各話者の半音節HMMを連結し、モデルを作成する。入力音声を与えられたとき尤度を計算し、その尤度により話者およびテキストの判定を行い、テキスト指定形話者認識の認識結果とする[2]。

Speaker Recognition based on Concatenated HMM Using Demi-syllable Units

Toshie Namiki, Tadaaki Simizu, Naoki Isu, Kazuhiro Sugata  
Dept. of Information and Knowledge Engineering Tottori Univ., 4-101 Koyama-minami, Tottori 680, JAPAN

### 4.認識実験

男性話者6名が同時期に5回発声した単語(都道府県名)6種類を認識用とし、学習用には、認識用とは別の11種類の単語を用いて半音節HMMを作成し実験を行った。音声データをサンプリング周波数11.025kHz、量子化16bitで収集した後、1次の適応プリエンファシスをおこなった。特徴パラメータとして、フレーム長23.2ms、フレーム更新周期8.0ms、ハミング窓を用い、16次のLPCケプストラム係数を求めた。

話者認識実験の結果、高い認識率が得られたことにより、本研究で提案した半音節HMMの話者認識への適応が有効であることを確認した。

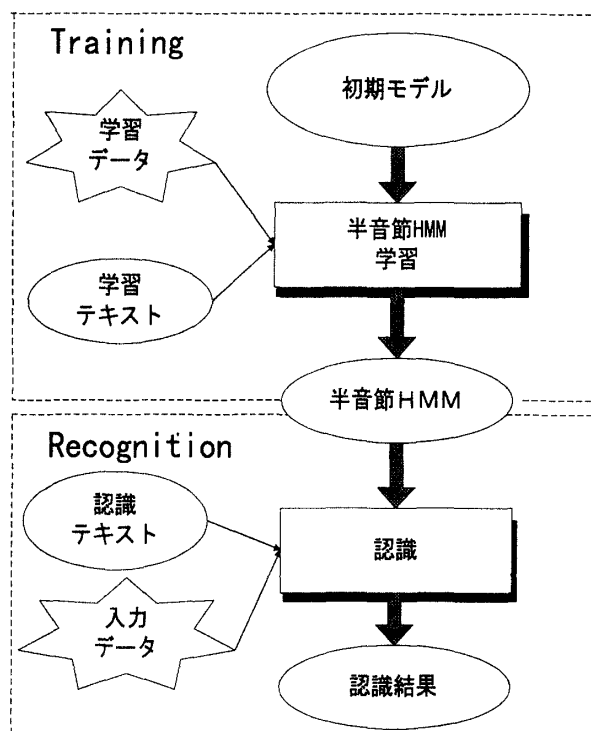


図1 半音節HMMを用いた話者認識方法

### 参考文献

- [1]渡辺隆夫, 吉田和永, 古賀真二: 半音節を単位としたHMMを用いた大語い音声認識, 信学論, D-II, Vol.J72-D-II, No. 8, pp. 1264-1269(1989)  
[2]松井知子, 古井貞熙: 連結音韻モデルによる話者認識, 信学技法, SP92-128(1993-01)