

論文 / 著書情報
Article / Book Information

論題(和文)	パワーの分散を考慮した拡張HMM合成法
Title(English)	
著者(和文)	南泰浩, 古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会 1996年秋季講演論文集, Vol. , No. 2-Q-6, pp. 141-142
Citation(English)	, Vol. , No. 2-Q-6, pp. 141-142
発行日 / Pub. date	1996, 9

○南 泰浩 古井貞熙

(NTTヒューマンインタフェース研究所)

1. はじめに

本論文では、雑音とひずみにHMMを適応させる方法として我々が提案した拡張HMM合成法[1][2]の改良について述べる。従来のHMM合成法[3][4]では、パワー（ケプストラムの0次）の分散をHMM合成に使うと認識率が低下するという問題があった。これは、ケプストラム0次の分布を一つの正規分布で表しているため、合成したあとの分布の近似精度が落ちるためであると考えられる。この劣化を改善するため、従来のHMM合成法では、ケプストラムの0次の分散を0に固定していた。本論文では、HMM合成法での近似の精度をより高める手法を提案する。拡張HMM合成法に、この手法を導入して数字認識実験を行い本手法の有効性を調べる。

2. 拡張HMM合成法による適応化

雑音が重畳し、かつ、ひずんだ音声の生成過程を図1のようにモデル化する。音声信号 S は音声HMMによって生成される。また、雑音信号 N は雑音HMMによって生成される。 H 、 G は乗算性の歪みを表す。 k は音声と雑音のパワー比に対応する係数である。このモデルでの信号は $X = (-HGS + kHN)$

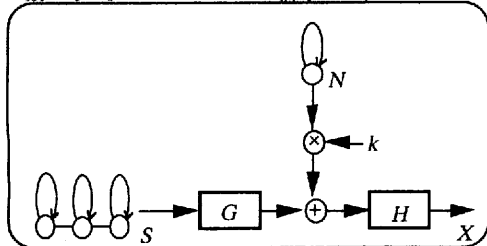
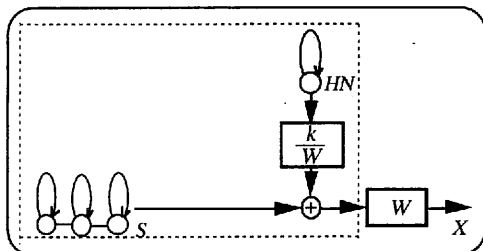


図1 雑音とひずみを含む音声のモデル化

図2 雑音とひずみを含む音声モデルの変形 ($W=HG$)

*Extended HMM composition incorporating power variance.

By Yasuhiro Minami and Sadaoki Furui
(NTT Human Interface Laboratories)

となる。ここで $W = HG$ と置くと、 $X = WS + kHN$ となる。この式を用いて図1のモデルを図2のように書き直す。

S を生成するHMMとしては不特定話者のHMMを用いる。また、雑音 N は直接求めることはできないので、マイクロフォン特性などを含んだ HN を生成するHMMを環境雑音から推定する。図中の+の演算には、HMM合成法を使う。 k, W は固定値であり分散がないとする。雑音と歪みの推定はこの k, W を同時に推定することにより実現する。

合成によって生成される雑音の重畳したHMMのモデルパラメータセットを $M(k, W)$ とする。尤度 $P(O|M(k, W))$ を最大にすることによって k, W を求める。 $O = \{x_1, x_2, \dots, x_T\}$ は入力音声の時系列である。

このアルゴリズムを以下に示す。

- (1) W を初期設定する。
- (2) k を変化させてHMMを合成し、 $P(O|M(k, W))$ が最大になる k を選択する。
- (3) 次に点線の枠内の部分のHMMを、 $1/W$ を固定し、(2)の k を用いて合成する。
- (4) Sankarのケプストラムバイアス推定法を用いて点線枠外の W を推定する[5]。
- (5) $1/W$ を新しい W を用いて更新する。
- (6) (2)から(5)を収束するまで繰り返す。

3. パワー分散の導入

HMM合成法では、以下のような式を使ってケプストラム領域の分布から対数スペクトル領域の分布を求め、さらに線形スペクトル領域に変換してHMM合成を行っていた。

$$\mu_{S_{\log}}^S = \Gamma \mu_{S_{\log}}^S \quad (1)$$

$$\Sigma_{S_{\log}}^S = \Gamma \Sigma_{S_{\log}}^S \Gamma^T \quad (2)$$

ここで、 Γ はコサイン変換、 t は転置を表す。

$\mu_{S_{\log}}^S$ 、 $\Sigma_{S_{\log}}^S$ はケプストラム領域での音声の正規分布の平均値と分散である。 $\mu_{S_{\log}}^S$ 、 $\Sigma_{S_{\log}}^S$ は対数スペクトル領域での正規分布の平均値と分散である。従来までの手法ではケプストラムの分散の0次の項を0とおいてHMM合成を行っていた。本論文では、ケプストラム0次の分布を図3のように量子化する。この量子化によって元のケプストラムの多次元分布は、量子化数分の混合分布となる。この各々の分布に対してHMM合成法による変換を行う。各分布の重みとして、ここでは、量子化した点の確率密度の値を用いた。ただし、総和が1.0になるように正規

化した。

4. 認識実験

4.1 実験条件

本論文では、4桁数字の認識実験を行った。表1に実験に用いたデータを示す。数字HMMの状態数は数字によって異なるが、状態毎の分布数は4である。雑音HMMとして1状態1混合分布のHMMを用いた。乗算性の歪みとして、バウンダリマイクロフォンを用いた。加算性雑音としては電子協データベース中の2つの雑音を用いた(表1)。雑音は音声とは別にバウンダリマイクロフォンを使って収録し、計算機上で所望のS/Nになるように音声データに加えた。この過程では、個々の音声データに対してS/Nを合わせたのではなく、音声データ全体に対してS/Nを求め、そのS/Nに応じた係数を雑音に掛け、この雑音を音声データに加えた。適応は、数字列毎に、教師なしで行った。この際、適応用のHMMとして1状態16混合分布の全音声HMMを使った。分布の量子化は、音声のHMMの分布に対してだけ行った。

4.2 実験結果

表2に実験結果を示す。一番上の段にHMM合成

表1 実験に用いたデータ

不特定話者数字HMM 学習用データ	電子協 4桁数字9702発声(70人)
評価用データ	4桁数字1785発声(51人)
雑音データ	電子協雑音データベース中の 計算機雑音と走行自動車内雑音

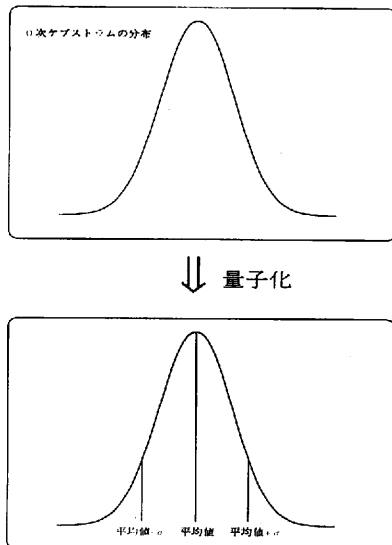


図3 0次ケプストラムの量子化

法で分散をそのまま使った場合の認識率を示す。また、その下段に0次のケプストラムの分散を0に固定した場合の認識率を示す。この結果から、0次のケプストラムの分散を使うと認識率の低下をまねいていることが分かる。特に計算機雑音では、この低下が大きい。

その下段に提案手法の結果を示した。量子化は3点(平均値、 $\pm\sigma$)と5点(平均値、 $\pm\sigma$ 、 $\pm2\times\sigma$)の2種類で行った。表から量子化数3の場合、分散を固定した場合に比べ認識率が向上していることが分かる。特に計算機雑音に対してはこの効果が大きい。

また、本実験では、量子化数を5に増やしても認識率の向上が見られなかった。

5. まとめ

本報告では、雑音とひずみにHMMを適応させる方法として我々が提案した拡張HMM合成法で、ケプストラムの0次の分布を量子化してHMM合成法を行う手法を提案した。この方法を用いて数字音声認識実験を行いその有効性を確認した。

謝辞

日頃、有意義な議論をしてくださる古井特別研究室の方々に感謝致します。本実験では、電子協のデータベースを利用した。

文献

- [1] 南 泰浩, 古井貞熙, "最尤推定法を用いたHMM適応化", 音講論集, 1-2-1, (1995-9).
- [2] 南 泰浩, 古井貞熙, "雑音と歪みを含んだ音声へのHMM適応化手法の評価", 音講論集, 2-5-14, (1996-3).
- [3] F. Martin, K. Shikano, and Y. Minami, "Recognition of noisy speech by composition of hidden Markov models", EUROSPEECH 93, pp. 1031-1034, (1993 9).
- [4] M. J. F. Gales and S. J. Young, "An improved approach to the hidden Markov model decomposition of speech and noise", ICASSP 92, pp. 233-236 (1992 3).
- [5] A. Sankar, C.-H. Lee, "Robust speech recognition based on stochastic matching", ICASSP 95, pp. 121-124, (1995 5).

表2 計算機雑音, 自動車内雑音中の4桁数字認識実験結果

雑音		計算機雑音		車内雑音	
S/N比		12dB	18dB	12dB	18dB
従来法	0次ケプストラム 分散有り	44.3%	74.7%	89.9%	95.8%
	0次ケプストラム 分散=0	60.3%	84.6%	90.1%	96.0%
提案法	量子化数 =3	61.4%	86.2%	91.1%	96.1%
	量子化数 =5	61.8%	85.9%	90.6%	95.7%