

連続量を変形要因とする重回帰モデルを内包するHMM

藤永 勝久[†] 中井 満[†] 下平 博[†] 嵯峨山 茂樹[‡]

[†]北陸先端科学技術大学院大学 情報科学研究科
〒923-1292 石川県能美郡辰口町旭台 1-1

[‡]東京大学大学院 工学系研究科
〒113-8656 東京都文京区本郷 7-3-1

あらまし 基本周波数 (F_0) や発話スタイルなどの補助特徴量を用いて認識時に動的にモデルの適応を行う重回帰HMMを提案する。音声スペクトルは F_0 や発話スタイルなど多くの変形要因により変化する。重回帰HMMは連続量を持つ変形要因に対する音声スペクトルの変化を重回帰モデルで表す。重回帰モデルをHMMの各隠れ状態に組み込むことにより、音素環境や音素内での時間的位置毎に、別々の重回帰モデルを持つことが可能となる。本報告では重回帰HMMの概念や基本的な定式化を説明し、EMアルゴリズムを用いて尤度最大基準でパラメータを推定するアルゴリズムを示す。変形要因に F_0 を用いて切り出し音素認識と孤立単語認識を行った結果、従来のHMMと比較してそれぞれ15.3%、22.0%の誤り削減率が得られた。

キーワード 音声認識, 隠れマルコフモデル, 重回帰モデル, F_0 , 適応

Multiple-Regression HMM for Speech Variability

Katsuhisa Fujinaga[†], Mitsuru Nakai[†], Hiroshi Shimodaira[†], Shigeki Sagayama[‡]

[†]School of Information Science, Japan Advanced Institute of Science and Technology
1-1 Asahidai, Tatsunokuchi, Ishikawa, 923-1292 JAPAN

[‡]Graduate School of Engineering, University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656 Japan

Abstract This paper proposes a new class of hidden Markov model (HMM) called multiple-regression HMM that utilizes auxiliary features such as fundamental frequency (F_0) and speaking styles that affect spectral parameters to better model the acoustic features of phonemes. Though such auxiliary features are considered to be the factors that degrade the performance of speech recognizers, the proposed MR-HMM adapts its model parameters, i.e. mean vectors of output probability distributions, depending on these auxiliary information to improve the recognition accuracy. Formulation for parameter reestimation of MR-HMM based on the EM algorithm is given in the paper. Speaker-dependent speech recognition experiments demonstrated that error rates reduced by more than 15.3% and 22.0% in phoneme and isolated word recognition, respectively, compared with the conventional HMMs.

Key words speech recognition, hidden Markov model, multiple-regression model, F_0 , adaptation

1 はじめに

音声スペクトルは性別や話者の違い、文脈依存のみならず、発話スタイルや基本周波数 (F_0)、背景雑音レベルの影響 (Lombard 効果) などの多くの変形要因により変形することが知られている。これらの変形要因は音声認識時に得ることができるため、これらを有効に用いれば音声認識の精度を向上することができる。従来の研究では、話者性の違いに関しては話者適応 (MAP[1], VFS[2], MLLR[3]) の手法が考案され、文脈依存に関しては、環境依存モデル [4] の構築が考案されてきた。しかし、発話スタイルや F_0 の影響に関してはあまり行われていない。我々は発話スタイルや F_0 の影響などを考慮した音声認識の手法を研究する。

音声による対話システムを利用する時、同じ話者でも様々なスタイルの発話が入力される。例えば、システムが誤認識をして再度入力を促した場合、利用者は再び誤認識されないように、単語をより明瞭にゆっくりと強調して発声する傾向がある。このような普段とは異なる発話は、しばしば誤認識の原因となる [5]。この問題を解決する手法には、まず、発話スタイル毎に別々のモデルを利用するという手法が考えられる [6]。次に、発話スタイルや F_0 の変化に対して、モデルの適応や入力特徴量の正規化を行くことが考えられる。この場合、フレーム毎に適応や正規化を行えるなら、更に有効だと考えられる。

Singer と嵯峨山は、音声スペクトルと F_0 の相関を利用した入力特徴量の正規化手法を提案した [7]。この手法は、音素毎にケプストラムと F_0 の関係を線形モデルで持ち、入力された音声を音素毎の線形モデルにしたがって正規化する。これにより音声認識の精度向上が確認されている。この手法は各音素内では線形モデルが一定だと仮定している。しかし、同じ音素でも音素環境や、前後の音素との接続部や音素中心などの位置により、音声パラメータと F_0 の関係は変化すると考えた方がより自然と思われる。その為、 F_0 による音声パラメータの正規化やモデルの適応と音素の認識を同時に行い、同じ音素内でも時間毎に複数のモデルを持つ事が有効だと考える。これは HMM の定式に適応処理を含め、 F_0 などの変形要因により HMM のパラメータが時間同期で適応されるモデルを構築することで実現できると考える。我々はこのようなモデルの1つとして、HMM に重回帰モデルを組み込み、時間同期で HMM のパラメータを適応させる重回帰 HMM (MR-HMM: Multiple-Regression HMM) を提案する。本報告では各隠れ状態の出力確率に重回帰式を組み込み、 F_0 などの変形要因を説明変数とすることにより平均値を変化させる事を試みる。重回帰 HMM は Singer らの研究に対し

- 音素毎ではなく状態毎に重回帰モデルを持つ
- 変形要因の変化に対して、入力特徴量の正規化を行うのではなく、HMM のパラメータの適応を行う
- 学習データに対し尤度最大基準で重回帰係数を求

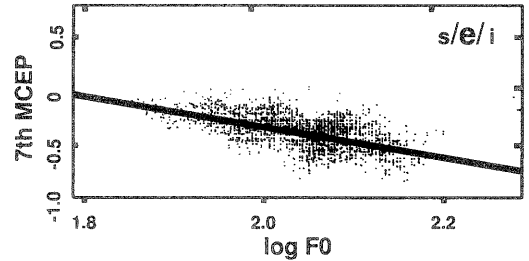


図 1: F_0 と 7 次 の MCEP の間の相関

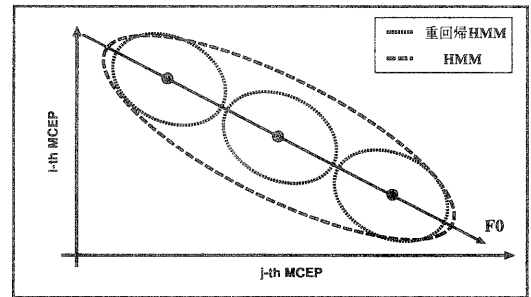


図 2: F_0 の変化に対してモデルの適応を行うことにより、分布の分散を小さくする

める

という違いがある。

なお、Juang 等が提案している自己重回帰 HMM (AR-HMM: Autoregressive HMM) [8] は重回帰 HMM と全く異なるモデルである。AR-HMM は観測ベクトルの自己相関を利用したモデルで、各状態には自己相関値を持つ。この自己相関値は過去の入力を説明変数とした重回帰式に等価に変換され、出力確率の平均値を変化させる。これに対し、重回帰 HMM は観測に直接用いない要因を説明変数とするという違いがあり、両者全く異なる。

第 2 章では重回帰 HMM と、EM アルゴリズムを用いた重回帰 HMM のパラメータの推定法の説明を行う。第 3 章では変形要因として F_0 を用いた話者依存の孤立単語認識実験を行い、重回帰 HMM の性能評価を行う。第 4 章は結びである。

2 重回帰 HMM

2.1 重回帰 HMM の概要

音素 s/e/i (/s/と/i/に挟まれた/e/) の $\log F_0$ と 7 次 のメルケプストラム係数 (MCEP) との相関を図 1 に示す。この図より、7 次 の MCEP と $\log F_0$ にはある程度の相関が読み取れる。 F_0 の音声スペクトルに対する影響は従来より研究されており、母音のホルマント周波数と F_0 の関係は生理的な観点などから説明されている [9]。音声スペクトルと F_0 の相関は文脈や音素環境により変化するために、音声パラメータの F_0 を用いた正規化は、音

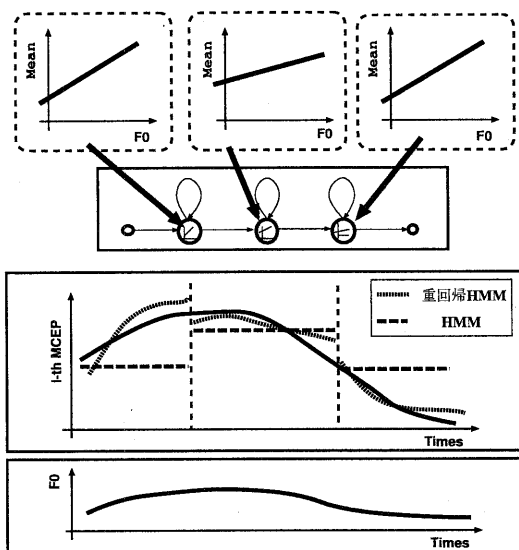


図 3: F_0 の変化に対して出力確率の平均値を変化させる

声認識に同期して行われる方が望ましい。重回帰 HMM は F_0 などの変形要因によりモデルパラメータを変化させることにより、この問題を解決する。

重回帰 HMM と HMM の出力確率の分布の概念的な違いを図 2 に示す。従来の HMM では大きな分散を持つ。これに対し、重回帰 HMM では変形要因の変化に対し HMM のパラメータを変化させる事により小さな分散を持つため、曖昧さが減少する。

重回帰 HMM が F_0 により平均値を変化させる様子を図 3 に示す。図の上段は重回帰 HMM の各状態の出力確率の平均ベクトル (この例では 1 次元のスカラー) に対する F_0 の回帰直線を示し、下段は MCEP と F_0 、HMM と重回帰 HMM の各隠れ状態の出力確率の平均値を示している。従来の HMM では各状態の平均値は一定値をとり変化しない。それに対し重回帰 HMM では平均値が回帰式と F_0 により時間毎に変化する。この場合の時刻 t での平均値 μ_t は以下のようにモデル化する。

$$\mu_t = r_0 + r_1 \log F_0(t), \quad (1)$$

ここで r_0, r_1 は回帰係数、 $F_0(t)$ は時刻 t での F_0 である。 M 個の変形要因 ($\xi_{1t}, \dots, \xi_{Mt}$) の場合、上式は以下のように書き直せる

$$\mu_t = r_0 + r_1 \xi_{1t} + \dots + r_M \xi_{Mt}. \quad (2)$$

本来、重回帰 HMM は連続値を持った変形要因を扱うように定式化している。しかし、離散値を持つ変形要因 (性別など) も、数量化法 (MDS) を利用することにより、変形要因が離散的なものでも扱うことができる。

重回帰 HMM が回帰式により平均値を変化させる点は MLLR と似ている。MLLR は認識環境への事前適応の

手法であり、予め認識環境の適応データから観測される特徴量 (例えば MCEP) を用いて HMM のパラメータを更新する。これに対し、重回帰 HMM では特徴量以外の観測量 (例えば F_0) を変形要因として、時間同期に HMM のパラメータを変化させるという違いがある。

2.2 重回帰 HMM の確率計算

従来の HMM と重回帰 HMM の確率計算の違いは、重回帰 HMM は変形要因による条件付き確率を求めるという点にある。しかし、定式化の多くは従来の HMM と同じであり、同様の手順で確率を計算できる。特に、遷移確率 a_{ij} と、初期確率 π_i は従来の HMM と全く同様に定義する。

重回帰 HMM の各状態の出力確率の分布は正規分布と仮定し、平均ベクトルを μ 、共分散行列を U とする。 M 次元の変形要因ベクトル $\xi_t = [\xi_{1t}, \xi_{2t}, \dots, \xi_{Mt}]'$ が与えられた場合、 N 次元の観測ベクトル $x_t = [x_{1t}, x_{2t}, \dots, x_{Nt}]'$ を状態 i が出力する確率を以下のように定義する。

$$b_i(x_t | \xi_t) = \frac{1}{(2\pi)^{\frac{N}{2}} |U|^{\frac{1}{2}}} e^{-\frac{1}{2}(x_t - \mu(\xi_t))' U_i^{-1} (x_t - \mu(\xi_t))}, \quad (3)$$

ここで $\mu(\xi_t)$ は以下のように求める。

$$\mu(\xi_t) = R_i \hat{\xi}_t, \quad (4)$$

$$\hat{\xi}_t = (1, \xi_t) = (1, \xi_{1t}, \xi_{2t}, \dots, \xi_{Mt})'$$

ここで R_i は $N \times (M + 1)$ 次元の重回帰行列である。

変形要因ベクトル系列 $\xi = (\xi_1 \dots \xi_T)$ が与えられた場合の、モデルが観測ベクトル系列 $X = (x_1 \dots x_T)$ を出力する確率は以下のように定義する。

$$\begin{aligned} f(X | \lambda, \xi) &= \sum_{\theta \in \Theta} f(X, \theta | \lambda, \xi) \\ &= \sum_{\theta \in \Theta} \pi_{\theta_1} b_{\theta_1}(x_1 | \xi_1) \prod_{t=2}^T a_{\theta_{t-1}, \theta_t} b_{\theta_t}(x_t | \xi_t), \end{aligned}$$

ここで、 λ は重回帰 HMM のパラメータ、 θ は状態系列 ($\theta_1, \dots, \theta_T$)、 Θ は全ての可能な状態系列を表す。 $f(X, \theta | \lambda, \xi)$ は、 ξ と θ が与えられた場合に、モデルが X を出力する確率である。

forward/backward アルゴリズムや Viterbi アルゴリズムなど、従来の HMM で用いられるアルゴリズムに関しても、従来の HMM の出力確率を式 (3) に置き換え、変形要因ベクトル系列を与えることにより利用できる。

なお、従来の全共分散を持つ HMM においては、変形要因を特徴量に組み込む事で、変形要因と他の特徴量の相関を利用することができる。全共分散を持つ HMM と重回帰 HMM の違いは、全共分散の HMM は変形要因により尤度の最大値が変化するが、重回帰 HMM の場合は平均値が変化するだけで尤度の最大値は変わらないという点にある。重回帰 HMM も全共分散を持つことも可能なので、特徴量に入れるべきものと、変形要因に組み込むべきものを見極めていく必要がある。

2.3 重回帰 HMM のパラメータ推定

重回帰行列 R_i を含む重回帰 HMM の各パラメータは全て尤度最大化基準により求めることができる。パラメータの推定には EM アルゴリズムや Viterbi 学習などを用いることができ、本報告では EM アルゴリズムを用いた重回帰 HMM のパラメータ推定について説明する。

重回帰 HMM のパラメータ推定は次のような補助関数 $Q(\lambda, \bar{\lambda})$ を繰り返し最大化する事により行う。

$$Q(\lambda, \bar{\lambda}) = \sum_{\theta \in \Theta} f(X, \theta | \lambda) \log(f(X, \theta | \bar{\lambda})),$$

ここで λ は現在の重回帰 HMM のパラメータ、 $\bar{\lambda}$ は更新されたパラメータを表す。

重回帰 HMM の再推定式は以下のように導出できる。

$$\bar{R}_i = \left(\sum_{t=1}^T \gamma_t(i) x_t \hat{\xi}_t' \right) \left(\sum_{t=1}^T \gamma_t(i) \hat{\xi}_t \hat{\xi}_t' \right)^{-1},$$

$$\bar{U}_i = \frac{\sum_{t=1}^T \gamma_t(i) (x_t - R_i \hat{\xi}_t) (x_t - R_i \hat{\xi}_t)'}{\sum_{t=1}^T \gamma_t(i)},$$

$$\bar{\pi}_i = \gamma_1(i), \quad \bar{a}_{ij} = \frac{\sum_{t=2}^T \gamma_{t-1}(i, j)}{\sum_{t=2}^T \gamma_{t-1}(i)},$$

ここで、 γ_t は時刻 t に状態 i に存在する確率、 $\gamma_{t-1}(i, j)$ は時刻 $t-1$ に状態 i に存在し、時刻 t に状態 j に存在する確率を表す。 γ_t と $\gamma_{t-1}(i, j)$ は従来の HMM と同様に forward/backward アルゴリズムを用いて求める。

変形要因を考慮しない場合、上の再推定式は従来の HMM の再推定式と等価である。

実際の計算では \bar{R}_i は以下の式の $N \times (M+1)$ 個の連立方程式を直接解くことにより求める。

$$\bar{R}_i \left(\sum_{t=1}^T \gamma_t(i) \hat{\xi}_t \hat{\xi}_t' \right) = \sum_{t=1}^T \gamma_t(i) x_t \hat{\xi}_t'$$

ここで左辺の括弧内の行列のコンディションが悪いときには \bar{R}_i が安定して求められず、括弧内の行列式が 0 の時には解が求まらない。これらの問題に対しては、特異値分解や、重回帰分析で行われる回帰係数の棄却などを行うことで解決することが考えられる。本報告では、括弧内のコンディションが悪いときには、従来の HMM と同様に状態の平均値を固定値として求めた。この場合、括弧内の行列の 1 行 1 列の値と、右辺の 1 列目の値を用いることで平均値を求めることができる。求められた平均値を \bar{R}_i の 1 列目に代入し、残りの列を全て 0 にすることにより、HMM と重回帰 HMM を混在させたモデルを構築することができる。

パラメータの推定の手順は以下の通り:

Step 1: 重回帰 HMM の初期パラメータを設定する

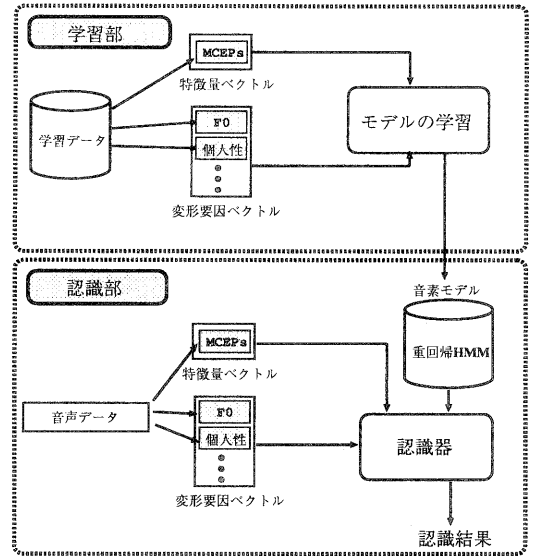


図 4: 重回帰 HMM を用いた音声認識の概要

- Step 2:** 終了条件を満たすまで、Step 2.1, 2.2 を繰り返す
- Step 2.1:** 現在のパラメータを用いて γ_t と $\gamma_{t-1}(i, j)$ を計算する
- Step 2.2:** γ_t と $\gamma_{t-1}(i, j)$ を用い再推定式からパラメータを求め、更新する

3章で行う実験では終了条件は 10 回の繰り返しとした。

2.4 重回帰 HMM を用いた認識システム

重回帰 HMM を用いた認識システムの例を図 4 に示す。

学習部では、学習データより特徴量の抽出を行うと同時に、 F_0 や、個人性や発話スタイルのような変形要因を抽出する。本報告の実験では F_0 のみを用いている。個人性や発話スタイルや感情などは、それらを認識するシステムの出力(怒り、興奮など)を用いることも、認識される前の何らかの特徴量(声道長、平均 F_0 など)を用いることもできる。そして、抽出された特徴量と変形要因を元に、重回帰 HMM を学習する。

認識部では、入力された音声データから特徴量と変形要因を学習部と同様に抽出し、それを元に各モデルの尤度を計算し、最も尤度の高いモデルを認識結果として出力する。

3 実験

3.1 実験条件

重回帰 HMM の性能評価の為に、特定話者の切り出し音素実験と孤立単語認識実験を行った。実験データには、ATR 研究用日本語音声データベース・セット A 中、男性 2 話者 (MAU, MHT) 女性 2 話者 (FFS, FMS) を用いた。

表 1: 重回帰 HMM と HMM の音素認識結果

分散	モデル	特徴量	要因	% 誤り率				% 誤り削減率			
				母音	有声子音	無声子音	全体	母音	有声子音	無声子音	全体
対角	HMM	C	-	17.6	28.4	38.5	18.7	-	-	-	-
		C+P	-	19.3	34.0	41.4	20.3	-9.5	-17.5	-9.1	-8.8
	重回帰	C	P	15.5	20.2	27.1	15.8	11.1	29.8	27.0	15.3
全共	HMM	C	-	11.3	9.3	5.7	8.3	-	-	-	-
		C+P	-	12.2	10.3	8.3	9.3	-9.0	-9.0	-52.3	-11.3
	重回帰	C	P	10.6	8.0	6.0	8.0	6.9	14.5	-3.6	4.3

(C: MCEP(13)+ Δ MCEP(13), P: $\log F_0$, 母音:/a,i,u,e,o/, 有声子音:/b,d,g/, 無声子音:/p,t,k/, 全体:26 音素)

表 2: 重回帰 HMM と HMM の孤立単語認識結果

# 状態数	モデル	特徴量	変形要因	% 誤り率	% 誤り削減率
84 (monophone)	HMM	C	-	14.6	-
		C+P	-	17.1	-18.0
	重回帰 HMM	C	P	11.1	22.0
206 (triphone)	HMM	C	-	6.2	-
		C+P	-	6.7	-11.2
	重回帰 HMM	C	P	4.5	26.5
406 (triphone)	HMM	C	-	4.7	-
		C+P	-	5.1	-9.5
	重回帰 HMM	C	P	3.5	22.0

(C: MCEP(13)+ Δ MCEP(13), P: $\log F_0$)

16 kHz サンプリングした波形データをブラックマン窓 25ms、フレーム周期 5ms で分析し、13 次の MCEP と 13 次の Δ MCEP と $\log F_0$ を抽出した。ここで、MCEP、 Δ MCEP は 0 次をパワーとして持つ。 $\log F_0$ の抽出にはケプストラム法を用い、無声部分は $\log F_0 = 0$ とした。

実験は重回帰 HMM (特徴量 MCEP+ Δ MCEP、変形要因 $\log F_0$) と、HMM (特徴量 MCEP+ Δ MCEP と特徴量 MCEP+ Δ MCEP+ $\log F_0$ の 2 種類) で行った。

3.2 切り出し音素認識実験

音素ラベル / n, a, b, t, d, e, f, g, h, i, s, k, m, n, o, p, q, r, s, j, t, ts, u, w, j, z / の計 26 音素を用いて、環境独立、話者依存の切り出し音素認識実験を行った。モデルは HMM、重回帰 HMM 共に、対角あるいは全角共分散の分布を持つ 1 混合 3 状態 left-to-right 型とした。学習データは、セット A 中の重要語 5240 単語の奇数番目と音素バランス単語 216 単語を使用し、評価データには重要語 5240 単語の偶数番目の全音素を利用した。

誤り率と誤り削減率の各話者の平均を表 1 に示す。重回帰 HMM は全共分散でも対角共分散でも性能の向上が見られる。それに対し、特徴量に $\log F_0$ を加えた従来の

HMM は性能の低下がみられる。

母音/a,i,u,e,o/、有声子音/b,d,g/、無声子音/p,t,k/の認識結果から、重回帰 HMM では有声子音の性能の向上が高いことがわかる。また、 F_0 を変形要因としているために、無声子音では性能の向上が見られないと予想したが、対角共分散の場合は性能の向上が見られた。これは前後の母音の影響などが原因だと考える。

3.3 孤立単語認識実験

対角共分散の分布を持つ 1 混合、left-to-right 型の重回帰 HMM/HMM を用いて特定話者孤立単語認識実験を行った。また、音素環境毎に F_0 の変化に対する音声ベクトルの変化が違ふと考えられるため、環境依存モデルを構築することにより回帰行列の推定性能が向上すると思われる。その為、状態数 84(monophone)、206(triphone)、406(triphone) の 3 種類を用意し、状態数毎の結果も確認した。monophone は全て 3 状態のモデルとした。triphone の状態共有構造は時間方向に最大 3 状態とし ML-SSS[10] を用いて構築した。この実験では話者毎に従来の HMM (特徴量 MCEP+ Δ MCEP) によって状態共有構造を構築し、他のモデルも同じ共有構造を利用した。学

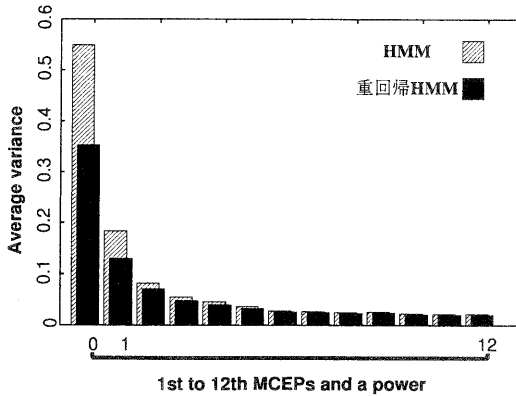


図 5: 重回帰 HMM と従来の HMM の分散の平均

習データは、セット A 中の重要語 5240 単語の奇数番号と音素バランス単語 216 単語を使用し、評価データには重要語 5240 単語の偶数番号で学習データに無い環境を含む 225 単語を除いた 2395 単語を利用した。辞書は評価データ全体を使い、認識は評価データの 4 の倍数の番号である 1193 単語で行った。

実験結果を表 2 に示す。全ての組み合わせにおいて、重回帰 HMM は従来の HMM(MCEP+ Δ MCEP) より最低 22.0% の性能の向上が見られる。それに対し、特微量に $\log F_0$ を加えた従来の HMM は、切り出し音素認識実験と同様に性能の低下がみられる。これらの実験より、重回帰 HMM は従来の HMM より $\log F_0$ の情報を有効に利用できると考える。

図 5 は重回帰 HMM と従来の HMM(特微量 MCEP+ Δ MCEP) の出力確率の分散の平均を MCEP の 0 次から 12 次までプロットした図である。この図から、 F_0 により平均値を変化させることで分散が減少していることが確認できる。

4 むすび

本報告では、変形要因による音声スペクトルへの影響を考慮したモデルとして重回帰 HMM を提案し、EM アルゴリズムを用いて尤度最大基準でパラメータを推定するアルゴリズムを導いた。重回帰 HMM を評価するために、変形要因を F_0 とする音声認識実験を行った。切り出し音素認識と孤立単語認識実験を行い、それぞれ従来の HMM に対して 15.3%、22.0% 以上の性能向上を見せた。従来の HMM において特微量に F_0 を加えた場合は性能が低下しており、重回帰 HMM は従来の HMM よりも、 F_0 を有効に活用したと考えられる。また本報告では F_0 のみを扱ったが、特微量と相関のあるものならば変形要因として扱うことができる。

今後の研究としては

- F_0 以外の変形要因を用いた認識実験
- 回帰係数の棄却

- 非線型モデルへの拡張
- 分散や遷移確率の拡張

などを行う予定である。更に、HMM から音声パラメータを生成するアルゴリズム [11] を重回帰 HMM に適用し、発話スタイルや F_0 を用いて音声を変化させる音声合成の実験を行うことも考えている。

謝辞

重回帰 HMM/HMM のプログラム製作に関してご教示下さった北陸先端大博士後期課程の松田繁樹氏に感謝します。なお、本研究の一部は文部省科学研究費補助金(基盤研究 C) 課題番号 12680375 の助成を受けて行われた。

参考文献

- [1] C.H. Lee, C.-H. Lin, B.-H. Juang. "A Study on Speaker Adaptation of the Parameters of Continuous Density Hidden Markov Models". *IEEE Trans. SP*, 39(4):806-814, April 1991.
- [2] 服部 浩明, 嵯峨山 茂樹. "移動ベクトル平滑化話者適応の原理とアルゴリズム". 信学技報, SP92(15):15-22, 1992.
- [3] C. J. Leggetter, P. C. Woodland. "Speaker Adaptation of HMM's Using Linear Regression". Technical Report TR.182, Cambridge University, 1994.
- [4] Jun Takami and Shigeki Sagayama. "A Successive State Splitting Algorithm for Efficient Allophone Modeling". In *Proc. ICASSP*, volume I, pages 573-576, 1992.
- [5] H. Soltau, A. Waibel. "On the influence of hyperarticulated speech on the recognition performance". In *Proc. ICSLP98*, page 736, 1998.
- [6] H. Soltau, A. Waibel. "Specialized acoustic models for Hyperarticulated Speech". In *Proc. ICASSP-2000*, pages 1779-1782, 2000.
- [7] H. Singer, S. Sagayama. "Pitch Dependent Phone Modeling for HMM Based Speech Recognition". In *Proc. ICASSP-1992*, volume I, pages 273-276, 1992.
- [8] B.-H. Juang. "Mixture autoregressive hidden Markov models for speech signals". *IEEE ASSP*, 33(6):1404-1413, 1985.
- [9] K. Honda. "Relationship Between Pitch Control and Vowel Articulation". In *Vocal Fold Physiology*, pages 286-289. College-Hill Press, 1983.
- [10] M. Ostendorf, H. Singer. "HMM Topology Design Using Maximum Likelihood Successive State Splitting". *Computer Speech and Language*, 11(1):17-41, 1997.
- [11] 徳田 恵一, 益子 貴史, 小林 隆夫, 今井 聖. "動的特微量を用いた HMM からの音声パラメータ生成アルゴリズム". 音響学会誌, 53 巻, 3 号, pages 192-200, 1997.