

音韻認識における連続分布型HMMの教師なし話者適応

中藤 良久* ・ 松本 弘**

Unsupervised Speaker Adaptation of Continuous Density HMM in Phoneme Recognition

Yoshihisa NAKATO* and Hiroshi MATSUMOTO**

This paper discusses an application of the unsupervised spectral codebook adaptation method based on a minimum fuzzy objective function to the continuous density Hidden Markov Model (HMM). The following adaptation techniques are examined: a) normalization of input spectra by fuzzy mapping with the adapted codebook, b) mean vector adaptation of samples derived by fuzzy mapping, and c) a direct adaptation of mean vectors in the process of codebook adaptation. Through 16-phoneme recognition tests using a single Gaussian distribution HMM with uncorrelated covariance trained by a male speaker, it is found that the methods a) and c) improve the average recognition accuracy for 9 male speakers from 61% for no adaptation to 72% for unsupervised speaker adaptation by ten word training. Furthermore, a modified adaptation method with frequency warping attains a higher average correct score of 70% for ten female speakers than a supervised method.

1. はじめに

不特定話者を対象とした大語彙連続音声認識システムの実現にとって、話者適応化法の確立は重要な課題である。話者適応の方法は、既知の音韻情報を利用する教師ありの方法と、任意の音声を用いる教師なしの方法に大別できるが、不特定話者の音声認識システムを構成するためには、任意の発話内容の短時間の音声を用いて、教師なしに学習可能なことが望まれる。これまで教師なし話者適応化法として、ベクトル量子化符号帳の写像を基本とする方法が提案されており¹⁻³⁾、主にSPLIT法による単語認識を通してその有効性が明らかにされている。特に、ファジイ目

*松下電器産業株式会社 ** 電気電子工学科 教授

的関数を適応化の基準とすることで短時間学習が可能となり、ある程度動的な話者適応も可能であることが確かめられている³⁾。

ここでは、更に、この適応化法を連続音声認識における有力手法である隠れマルコフモデル(HMM)による認識手法に適用する方法とその有効性について検討する。ベクトル量子化符号帳を用いた話者適応は、標準話者の符号帳を入力話者に適応させるため、離散型HMMに適した方法であり、実際にその有効性も確認されている^{4,5)}。しかし、離散型HMMでは、ベクトル量子化誤差により認識精度が劣化するため⁶⁾、連続分布型HMMに話者適応を組み込むことが好ましい。その方法として平田らは、HMMモデルの平均ベクトルを代表点(コードベクトル)と見なした区分量子化誤差ベクトル平均化法により²⁾認識精度を改善できることを示した⁷⁾。

ここでは、連続分布型HMMの高精度な話者適応を実現するため、先に提案したファジイ目的関数最小化基準に基づく教師なし話者適応化法に、更にスペクトル周波数軸の伸縮を取り入れた改良を加え、入力音声スペクトルを標準話者に合わせて正規化する方法と、HMMモデルの平均ベクトルのみを入力音声スペクトルに適応化させる方法を提案し、韻認識実験により比較検討する。

2. 話者適応化法

2.1 教師なし話者適応化法

少数の学習サンプルにより、教師なしにスペクトルの話者適応を行うためには、入力話者と標準話者とのスペクトル差を少数パラメータにより正確にモデル化する必要がある。先に提案した部分空間毎の個人差ベクトルの内挿による従来のスペクトル変換は、男女間のようにホルマントの個人差が極端に大きい場合に十分な適応が不可能であった。そこで、ここでは従来のモデルに加え、周波数軸伸縮を取り入れた改良型教師なし話者適応化法⁸⁾について述べる。ケプストラムで表される標準話者のスペクトル空間内に設定されたM個の代表点 v_k における入力話者との未知のスペクトル差 $\{\Delta_k\}$ に加えて周波数軸の伸縮を導入し、標準話者のコードブック y_i を、次式で入力話者のコードブック x_i に変換する。

$$x_i = \Psi \left\{ y_i + \sum_{k=1}^M w_{ik} \Delta_k; a_i \right\}, \quad (1)$$

$$a_i = \sum_{k=1}^M w_{ik} \alpha_k, \quad (2)$$

$$w_{ik} = \frac{d(y_i, v_k)^p}{\sum_{r=1}^M d(y_i, v_r)^p}, \quad (3)$$

$$d(y_i, v_k) = \|y_i - v_k\|. \quad (4)$$

但し、 p は内挿の平滑度を制御するパラメータであり、 $\Psi\{y_i; a_i\}$ は係数 a_i (以下伸縮係数と呼ぶ)の1次オールパスフィルターの特性を利用した非線形周波数変換

を表す. \mathbf{y}_i としてケプストラム係数を用いる場合, 第 n 次の成分の周波数変換前後のケプストラム係数 c_n と \tilde{c}_n とは次式で関係づけられる.

$$\sum_{n=-\infty}^{\infty} c_n z^{-n} = \sum_{n=-\infty}^{\infty} \tilde{c}_n \tilde{z}^{-n}, \quad (5)$$

$$\tilde{z} = \frac{z^{-1} - a_i}{1 - a_i z^{-1}}. \quad (6)$$

ここで式(1),(2)の個人差ベクトル Δ_k と伸縮係数 α_k は, 入力話者の学習サンプル $\{\mathbf{x}_i\}$ ($i = 1, \dots, N$) と適応化後のコードブック $\{\hat{\mathbf{x}}_i\}$ に関する, ファジイ目的関数

$$J_F(\{u_{ik}\}, \{\Delta_k\}, \{\alpha_k\}) = \sum_{i=1}^N \sum_{j=1}^L u_{ij}^F d(\mathbf{x}_i, \hat{\mathbf{x}}_j)^2 \quad (7)$$

を, 拘束条件

$$\sum_{j=1}^L u_{ij} = 1, \quad (8)$$

$$\sum_{k=1}^M \|\Delta_k\|^2 \leq \eta(E_x^2 - E_y^2) \quad (9)$$

のもとで最小化するように推定する. ここで, η は Δ_k の大きさを制限するパラメータである. E_y は $\{\mathbf{y}_j\}$ 作成時の学習サンプル自身の平均量子化誤差であり, E_x は \mathbf{x}_i ($i = 1, \dots, N$) の符号帳 $\{\mathbf{y}_j\}$ による平均量子化誤差である. 以上に基づく適応化アルゴリズムでは, 代表点数を順次増加させる段階的最小化手法を用いる^{9,10}. 代表点数を固定した場合, 次の反復手順により J_F は最小化される.

1. $\{\Delta_k\}$ に関する最小化

- (1) $\Delta_k = 0, \alpha_k = 0, (k = 1, \dots, M)$ として u_{ij} について最小化する.
- (2) (1) の u_{ij} を用いて Δ_k について最小化する.
- (3) 式1により \mathbf{y}_i を更新し, J_F の減少率が δ 以下であれば (2) に, その他は (1) に戻る.

2. $\{\alpha_k\}$ に関する最小化

- (1) $\Delta_k = 0, \alpha_k = 0, (k = 1, \dots, M)$ として u_{ij} について最小化する.
- (2) (1) の u_{ij} を用いて α_k について最小化する.
- (3) オッペンハイムの漸化式により¹¹, \mathbf{y}_i の各成分を変換し, J_F の減少率が δ 以下であれば終了し, その他は (2) に戻る.

なお, ステップ 2-(2) で, J_F は α_k に関して非線形であるため, ノルム制限付きの線形化法によって最小化した¹². 但し, ケプストラム係数の α に関する微係数は, $\delta\alpha = 0.01$ の差分により計算した.

2.2 教師あり適応化法

上記の方法の上限を知る意味で次の教師あり適応化法と比較する¹³⁾。この方法では、入力話者の任意のスペクトル \mathbf{x}_i を入力話者のスペクトル空間内にあらかじめ設定された代表点 \mathbf{v}_k ($k = 1, \dots, M$) (教師なしと同数) における標準話者との推定個人差ベクトル Δ_k ($k = 1, \dots, M$) を用いて式 (1) と同様 (但し、周波数伸縮なし) の次式で変換 (正規化) し、標準話者の HMM モデルで認識する。

$$\mathbf{y}_i = \mathbf{x}_i + \sum_{k=1}^M w_{ik} \Delta_k, \quad (10)$$

$$w_{ik} = \frac{d(\mathbf{x}_i, \mathbf{v}_k)^p}{\sum_{r=1}^M d(\mathbf{x}_i, \mathbf{v}_r)^p}. \quad (11)$$

ここで、個人差ベクトル $\{\Delta_k\}$ ($k = 1, \dots, M$) は、入力話者学習サンプルの変換スペクトル $\mathbf{y}_{i(r)}$ と対応付られた標準話者のスペクトル $\mathbf{y}_{j(r)}$ との差の総自乗誤差

$$J(\{C_r\}, \{\Delta_k\}) = \sum_{r=1}^N \|\mathbf{y}_{j(r)} - \mathbf{y}_{i(r)}\|^2 \quad (12)$$

を最小化することによって推定される。ここで、目的関数 J_F の最小化は、DP マッチングによる $\{C(r)\}$ に関する最小化と、次の正規方程式による $\{\Delta_k\}$ に関する最小化を反復することにより与えられる。

$$\sum_{k=1}^M W_{mk} \Delta_k = \sum_{r=1}^N w_{i(r)} e_{i(r)j(r)}, \quad (13)$$

$$(m = 1, 2, \dots, M).$$

ただし

$$W_{mk} = \sum_{r=1}^N w_{i(r)m} w_{j(r)k}, \quad (14)$$

$$e_{i(r)j(r)} = \mathbf{y}_{j(r)} - \mathbf{y}_{i(r)}. \quad (15)$$

なお、本方法は連続写像であるため、ベクトル量子化誤差による劣化は生じない。

3. HMM音韻認識における話者適応

3.1 HMMモデルの構造

連続分布型 HMM のモデル構造については、これまでもいくつかのタイプについて種々検討され、音韻認識では 4 状態 3 ループモデル (出力確率分布は同じ遷移元については共通とする) が良いことが確認されている⁶⁾。本研究でも同じモデル構造を用いる。また、出力率密度分布としては、混合多次元正規分布が望ましいが、パラメータの推定精度の向上には多くの学習データが必要であることを考慮し、ここでは単一無相関正規分布を使用する。

3.2 入力スペクトルのファジイ写像による正規化法

教師なしにHMMモデルを入力話者に適応させる替りに入力音声スペクトルをファジイ写像¹⁴⁾により正規化する方法(正規化法)を示す. この方法では, 学習によりあらかじめ求めた標準話者から入力話者への適応化コードブック $\{\hat{\mathbf{x}}_j\}$ を用いて入力話者のスペクトル \mathbf{x}_i をファジイベクトル量子化し, 級関数 u_{ij} を求め, 標準話者コードブック \mathbf{y}_j を用いて次式で復号化し, 標準話者のHMMモデルで認識する¹⁵⁾.

$$\hat{\mathbf{y}}_j = \frac{\sum_{u_{ij} > u_T} u_{ij}^F \cdot \mathbf{y}_i}{\sum_{u_{ij} > u_T} u_{ij}^F}. \quad (16)$$

但し, 復号に用いる級関数の閾値 u_T は, $1/L$ (u_{ij} の平均値)とした.

3.3 学習ベクトルのファジイ写像による適応

この方法は, 標準話者の符号帳 $\{\mathbf{y}_k\}$ と, 適応化符号帳 $\{\hat{\mathbf{x}}_k\}$ を用いて, 3.2と同様のファジイ写像(但し, 写像の方向は逆)により, 入力話者にファジイマッピングを行った標準話者の学習サンプルを用いて, HMMモデルの平均ベクトル $\{\mu_{ij}\}$ のみを再学習することに等価な方法である(以下写像学習法と呼ぶ). すなわち, 標準話者のHMMモデルを学習する際に, 第 n 単語の学習サンプル \mathbf{y}_i^n を標準話者コードブック \mathbf{y}_k でファジイベクトル量子化(F)した時の級関数 u_{ntk} の平均 h_{ijk} を予め学習しておき, 標準話者から入力話者への変換コードブック \mathbf{x}_k を用いて次式で適応化後の平均ベクトルを求める.

$$\mu_{ij} = \sum_{k=1}^L h_{ijk} \mathbf{x}_k, \quad (17)$$

$$h_{ijk} = \frac{\sum_n \sum_t \gamma_n(i, j, t) \tilde{u}_{ntk}}{\sum_n \sum_t \gamma_n(i, j, t)}, \quad (18)$$

$$\tilde{u}_{ntk} = \frac{u_{ntk}^F}{\sum_{u_{ntk} < u_T} u_{ntk}}. \quad (19)$$

但し, $\gamma_n(i, j, t)$ は, 第 n 単語の第 t フレームが状態 $i \rightarrow j$ に遷移する確率を示す.

3.4 平均ベクトルの直接的適応化法(直接法)

本方法(以下直接法と呼ぶ)では, ファジイ目的関数最小化に基づく標準話者コードブックの適応化の過程で, 標準話者のHMMモデルの平均ベクトル μ_{ij} を, コードベクトルの適応化と全く同様に, 式(1),(2),(3)において, \mathbf{y}_i の代わりに μ_{ij} を代入して逐次的に適応化し, 変換後の平均ベクトル μ_{ij} に変換する. 但し, μ_{ij} はスペクトル符号帳の一部とはなっていないため, J_F は \mathbf{y}_i のみで評価される.

4. 音声データと実験条件

4.1 音声データと分析条件

音声データは、男女各 10 名 (M 1~M 10, F1~F10) の発声した 212 単語である。また、音声の分析条件を Table 1, 適応化条件を教師なしについては Table 2, 教師ありについては Table 3 に示す。但し、適応化学習サンプルは認識用音声の一部 (最初の 10~50 単語) を用いた。評価する音韻は、有声破裂音 /b/, /d/, /g/, 鼻音および撥音 /m/, /n/, /N/, 摩擦音 /s/, /h/, 流音 /r/, 半母音 /w/, /y/, 母音 /a/, /i/, /u/, /e/, /o/ の 5 カテゴリー 16 音韻である。使用した音声データは、視察により音韻境界が求められており、認識時の音素境界は既知とした。また、子音については後続母音とのわたりにも認識に有効な情報を含んでいることが確認されているので⁶⁾、わたり部 10ms (1 フレーム) を付加して用いた。Table 4 に男女 20 名に対する各音韻の平均サンプル数とフレーム数を示す。HMM モデルの作成に用いる標準話者としては、男性話者のうち本人以外の男性 9 名の 16 音韻を、適応化なしに認識したときの平均認識率の高い 3 名 (M 4, M 7, M 8) を選び、各話者についてモデルを学習した。

4.2 音韻 HMM の学習条件

状態遷移確率の初期値としては、同じ状態から出ている全てのアークに等しい値を設定する。出力確率分布は、予め全データから平均ベクトル及び分散を求めておき、各出力確率分布に与える。HMM のパラメータ推定には、最尤推定を基本とした Baum-Welch アルゴリズムを用い、学習回数は予備実験から 10 回とした。認識には Viterbi アルゴリズムを使用する。また特徴パラメータは 15 次メルケプストラムである。

Table 1 Analysis conditions.

Sampling Frequency	10kHz
Hamming Window	20ms
Frame Period	10ms
Preemphasis	$1 - 0.90z^{-1}$
LPC Analysis Order	12 次
Mel-Cepstral Order	15 次

Table 2 Conditions in unsupervised speaker adaptation.

Codebook Size	L	256
No. of Typical Points	M	16
Interpolation Parameter	p	2.0
Fuzziness	F	1.5
Norm Constraint	η	1.0

Table 3 Conditions in supervised speaker adaptation.

No. of Typical Points M	16
Slope Constraint in DP	3~1/3
End Point Constraint	Free
Interpolation Parameter p	2.0

Table 4 Average numbers of phoneme samples over 20 speakers.

Phone	No. of Samples	No. of Frames		
		Mean	Min	Max
/b/	23	6	4	9
/d/	18	6	4	9
/g/	8	6	4	9
/m/	41	7	4	12
/n/	28	7	4	12
/N/	54	14	7	24
/s/	72	13	6	27
/h/	29	8	4	13
/r/	20	5	4	7
/w/	7	6	4	8
/y/	55	7	4	15
/a/	134	12	5	29
/i/	131	11	4	26
/u/	103	12	4	31
/e/	80	13	5	29
/o/	140	16	4	36

5. 認識結果

5.1 適応化法の比較

以下の実験では、先ず周波数伸縮を行わない従来の方法で各種適応化法を比較するため、各標準話者について標準話者以外の男性話者（9名）に関して有声破裂音または母音のみを認識対象とした認識実験を行った。教師あり・なし共に10単語（約700フレーム）の適応学習サンプルを用いたときの各適応化法における適応前後での有声破裂音と母音の話者平均認識率をTable 5に示す。また参考に、特定話者close（HMMの学習に用いたデータを認識させる）での認識率も示す。有声破裂音の認識では、直接法でわずかに認識率が改善されるが、他の手法については大きく認識率が低下し、適応化が悪影響を与えている。一方母音については、各手法とも6.3%~9.0%認識率が改善され、特に正規化法が優れており、教師あり正規化法と比較し同等の認識率が得られていることがわかる。なお、標準話者の違いによる適応化後の認識率のばらつきは、母音に比べ有声破裂音で大きい。

Table 5 Average recognition rates over the reference and test speakers in the various adaptation methods.

Phoneme Category	Without Adap.	Normal. Method	Mapped Training	Direct Method	Super. Normal.	Speaker Depend.
Voiced Plosive	68.3	59.2	60.6	68.8	64.3	92.9
Vowel	82.7	91.7	89.5	89.0	90.5	96.5

5.2 適応化における周波数軸伸縮の効果

ここでは、改良型教師なし適応化法における周波数伸縮の効果を見るため、前節と同じ実験条件で周波数伸縮有り/無しの教師なし正規化法を用い、標準話者以外の男性話者（9名）と女性話者（10名）に関して、有声破裂音及び母音を対象とした認識を行った。

適応前後における有声破裂音と母音に対する平均認識率をTable 6に示す。これより、男性話者に関しては周波数軸伸縮の効果は、有声破裂音及び母音とも僅かであるが、女性話者に関しては伸縮を導入することにより従来法に比べ、認識率は有声破裂音で平均8.3%、母音で平均9.5%改善され、男性テスト話者と同程度の認識率が達成されている。個人差が大きい場合に本方法が有効であることが確認された。

Table 6 Effects of frequency warping in speaker normalization.

Method	Voiced Plosive		Vowel	
	Male	Female	Male	Female
No Adap.	68.3	43.7	82.7	59.1
No Warping	59.2	54.7	91.7	80.5
Warping	60.0	63.0	92.1	90.0

5.3 16音韻の認識による評価

ここでは、前節までの結果をもとに、男性話者に関しては周波数伸縮なしの正規化法と直接法について、また女性話者に関しては周波数伸縮ありの正規化法について、16音韻を対象とした認識実験を行った。前節と同じ実験条件による認識結果を

男性及び女性について、それぞれ Table 7, Table 8 に示す。表中には全 16 音韻の平均認識率とともに音韻カテゴリー毎の平均認識率も示している。まず、男性について見ると標準話者によってばらつきがあるものの、前節同様有声破裂音を除くすべてのカテゴリーで適応の効果があり、摩擦音や流音・半母音では直接法が良く、鼻音や母音では正規化法が若干良いが、16 音韻の平均認識率では正規化法と直接法とはほとんど同じ認識率(約 73.2%)を与える。結局、適応なしに比べ約 12.0% 認識率が改善され、また教師あり適応化法の結果を若干上回る認識率が得られた。一方、女性話者では、適応なしと比較して各カテゴリーの平均認識率は約 20%~50%改善され、16 音韻の平均認識率で教師あり適応化法に比べ 3.8%高い 70.0% の認識率が得られた。結局、男女間のような大きな個人差がある場合にも改良型適応化法により男性間の適応に近い効果が得られた。

Table 7 Average recognition rates over male test speakers in the various adaptation methods.

Phoneme Category	Without Adap.	Normal. Method	Direct Method	Super. Normal.	Speaker Depend.
Voiced Plosive	53.5	48.8	45.0	34.5	83.6
Nasal	44.8	54.0	52.0	57.6	87.7
Fricative	86.8	84.9	88.2	83.1	97.0
Liquid, Semivowel	43.3	42.6	51.6	48.6	78.1
Vowel	63.6	81.8	80.9	80.8	92.1
16 Phonemes	61.2	73.2	73.3	72.6	90.3

Table 8 Average recognition rates over male test speakers in the various adaptation methods.

Phoneme Category	Without Adaptation	Normalization Method	Supervised Normalization
Voiced Plosive	31.4	55.7	28.1
Nasal	18.4	37.8	38.1
Fricative	60.7	81.7	75.5
Liquid, Semivowel	16.6	44.1	31.8
Vowel	29.2	79.2	78.0
16 Phonemes	30.2	70.0	66.2

5.4 学習サンプル長の効果

最後に、男性テスト話者のみを認識対象とし、教師なし正規化法と教師あり正規化法について学習サンプル長の効果を調べた。有声破裂音、母音及び 16 音韻について、5.1 と同じ標準話者 3 名の平均認識率と学習サンプル長との関係を Fig.1 に示す。但し、代表点数は教師なしについては 16、教師ありについては 10 単語学習で 16、30 および 50 単語学習で 32 とした。これより、教師なし正規化法は母音では学習サンプル数にかかわらず教師あり正規化法と同程度の認識率を与える。しかし、

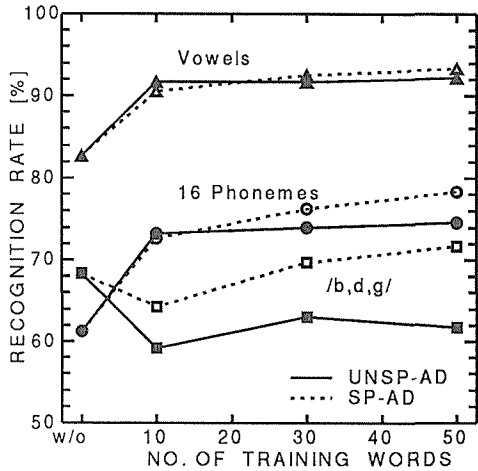


Fig.1 Effects of number of training words in the normalization method for the male test speakers.

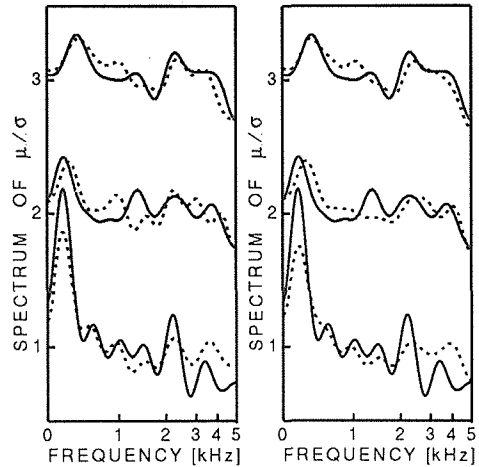


Fig.2 Comparison of the target and adapted models (μ/σ) for /b/.

有声破裂音では、学習単語が増すに従い教師あり正規化法による認識精度は改善されるが、教師なしではほとんど改善が見られない。以上より、50単語学習では16音韻の認識で教師あり正規化法が教師なしに比べ約4%高い認識率を与えるが、学習サンプルが少ない場合には教師なしも教師ありの方法に近い認識率を与えることがわかる。

5.5 考察

以上の適応化において、有声破裂音では各手法とも適応後に認識率が低下し、適応化が悪影響を与えている。しかし、教師あり正規化法においても認識率が低下していることから、これは学習時のサンプル数が少ない (Table 4) ため、モデルの推定精度 (特に分散) が低いことに起因しているものと思われる。特に本実験では、メルケプストラムを用いたため分散の小さい高次のケプストラム成分の僅かな違いが悪影響を与えている可能性がある¹⁶⁾。これをスペクトル上で考察するため有声破裂音/b/における話者M4→M1の適応について、入力話者M1(実線)及び適応したM4のモデル(点線)の各状態の μ/σ (平均ベクトル/標準偏差)に対応するスペクトルを比較しFig.2に示す。Fig.2 (a)は適応化の効果のあった直接法の場合、Fig.2 (b)は逆効果であった写像学習法による場合である。両手法とも全体的には大きな差は見られないが、低域周波数のピーク付近で写像学習法の方が入力話者のモデルとの差が大きく、これが認識率に大きな影響を及ぼしていたものと思われる。

6. まとめ

先に提案したファジイ目的関数最小化基準による教師なし話者適応法を連続分布型HMMに適用する方法を提案した。その結果ファジイ写像による正規化法と直接

法により、10 単語程度の学習音声は教師有りの方法と同程度の認識率が得られることを示した。また、適応化モデルに周波数軸伸縮を加えることにより、男性話者で女性の音声を認識させる場合にも、男性話者に近い認識率を達成できることを示した。しかし、適応化で得られる認識精度はまだ十分ではない。特に、有声破裂音では適応化により逆に認識率が低下するため、更に改良する必要がある。

今後は、さらに学習サンプル数を増やすとともに、混合分布によるHMMモデルの高精度化、複数話者による学習マルチテンプレート法や話者重畳型等による動的個人差への対応等について検討を進める予定である。

参考文献

- 1) 古井：音響学会音声研究会資料，**SP87-67**(1987).
- 2) 山下，松本：電子情報通信学会論文誌，**J72-A**，869(1989).
- 3) H.Matsumoto and Y.Yamashita: J. Acoust. Soc. Jpn. **E14**, 353(1993).
- 4) 中村，花沢，鹿野：日本音響学会誌，**45**，942(1989).
- 5) 山本，牧野，城戸：音響学会講演論文集，2-P-26(1990-3).
- 6) 中川，平田，橋本：日本音響学会誌，**46**，486(1990).
- 7) 平田，中川：音響学会音声研究会資料，**SP90-16**(1990).
- 8) 中藤，松本：音響学会講演論文集，2-P-15(1989-10)
- 9) N.R.Draper and H.Smith: *Applied Regression Analysis*, (Jhon Wiley & Sons,Inc., New York, 1966), Chap. 6.
- 10) 白木，誉田：音響学会音声研究会資料，**SP87-67**(1987).
- 11) A.V.Oppenheim and D.Johnson: Proceedings of IEEE, **59**，299(1971).
- 12) D.W.Marquardt: J. Soc. Indust. Math., **11**，431(1963).
- 13) 松本，丸山，井上：日本音響学会誌，**50**，549(1994).
- 14) 中村，鹿野：日本音響学会誌，**45**，107(1989).
- 15) 中藤，近藤，松本：音響学会講演論文集，2-P-13(1989-10).
- 16) Y.Tohkura: "A Weighted Cepstral Distance Measure for Speech Recognition," *Proceedings of IEEE ICASSP*, pp.761~764(Tokyo,1986).