

TR-H-276

## 識別的特徴抽出に基づく最小誤り 話者認識のための特徴抽出

宮島千代美 (ATR-HIP/名古屋工業大)、  
渡辺秀行、徳田恵一 (名古屋工業大)、  
北村 正 (名古屋工業大)、片桐 滋

1999.9.8

## ATR人間情報通信研究所

〒619-0288 京都府相楽郡精華町光台2-2 TEL: 0774-95-1011

ATR Human Information Processing Research Laboratories  
2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan

Telephone: +81-774-95-1011  
Fax : +81-774-95-1008

識別的特徴抽出法に基づく最小誤り話者認識のための特徴抽出

**Feature Representation for Minimum Error Speaker  
Recognition Based on Discriminative Feature Extraction**

宮島 千代美<sup>†‡</sup>  
Chiyomi Miyajima

渡辺 秀行<sup>†</sup>  
Hideyuki Watanabe

徳田 恵一<sup>‡</sup>  
Keiichi Tokuda

北村 正<sup>‡</sup>  
Tadashi Kitamura

片桐 滋<sup>†</sup>  
Shigeru Katagiri

話者認識システムは特徴抽出器および分類器から構成されている。多くの場合、話者の特徴はリニアスケールもしくはメルスケールに基づいて抽出される。また、高周波数領域に話者の特徴が多く存在するという説もある。しかし、分類器の設計規準と異なる規準で選ばれたこれらの特徴は、必ずしも最適な認識性能に結び付くとは限らない。そこで本研究では、最小分類誤り/一般化確率的降下法 (MCE/GPD) の拡張である識別的特徴抽出法 (DFE) により、話者認識性能にとって最適な特徴抽出器の設計を行う。ここでは、GMM に基づくテキスト独立型話者識別において、ある特定の周波数域を強調することが可能な 2 次オールパス関数による周波数ワーピングを利用したメルケプストラム推定法を特徴抽出に適用し、最小誤り規準で周波数ワーピングと分類器の同時最適化を行う。その結果、従来とは異なる周波数ワーピングが得られ、同時に、従来の特徴量を用いて分類器のみを最適化した場合に比べて高い認識性能が得られることを示す。

<sup>†</sup> ATR 人間情報通信研究所

ATR Human Information Processing Research Laboratories

<sup>‡</sup> 名古屋工業大学 知能情報システム学科

Department of Computer Science, Nagoya Institute of Technology

©ATR 人間情報通信研究所

©ATR Human Information Processing Research Laboratories

# 目次

1	まえがき	1
2	話者認識システム	2
2.1	話者識別	2
2.2	テキスト独立型話者認識	2
2.3	混合ガウスモデル (GMM)	2
2.4	GMM に基づく話者モデル	3
2.5	EM アルゴリズムによる GMM の学習	4
3	話者の特徴抽出 —メルケプストラム推定法—	7
3.1	1次オールパス関数による周波数変換	7
3.2	2次オールパス関数による周波数変換	9
3.3	メルケプストラム係数の推定	10
3.4	最小化問題の解法	11
4	最小分類誤りのための識別の特徴抽出法	15
4.1	識別の特徴抽出法 (DFE)	15
4.2	話者識別のための DFE 学習則	15
4.3	損失関数の微分	18
4.4	GMM 出力確率のワーピングパラメータに関する微分	19
4.5	GMM 出力確率の GMM パラメータに関する微分	24
5	実験	27
5.1	識別タスクの設定	27
5.2	ML 学習による GMM を用いた識別実験	29
5.3	GPD 学習による GMM パラメータの最適化	30
(5.3.1)	学習のチューニングパラメータの設定	30
(5.3.2)	損失関数の検討	30
(5.3.3)	損失と識別率の変化	31
5.4	DFE による認識システムの最適化	33
(5.4.1)	ワーピングパラメータの最適化に関する検討	33
(5.4.2)	最適ワーピングパラメータの全探索	36
(5.4.3)	他の分析法との比較	38
6	あとがき	39
7	謝辞	40
	参考文献	41

---

付録 A	2次オールパス関数による周波数ワーピング	43
付録 B	逆フーリエ変換値の変換行列	45
B.1	実偶関数の場合	46
B.2	純虚奇関数の場合	46
付録 C	ケプストラム係数の周波数変換行列	47
C.3	ケプストラムからメルケプストラムへの変換	47
C.4	メルケプストラムからケプストラムへの変換	48
付録 D	ML 学習によるシステムの識別誤り率	50
付録 E	GPD 学習によるシステムの識別誤り率	62

## 1 まえがき

入力音声から、その言語的内容と独立に、発話者を機械により自動認識するテキスト独立型話者認識は、個人情報へのアクセス制限、秘密機関への入出チェックなど、セキュリティシステムにおける認証手段として重要であり、長年に亘り広く研究されている。また、セキュリティのみならず、言葉を理解する音声自動認識に対しても、テキスト独立型話者認識の有用性は明らかに高い。例えば、音声認識システムが日常的騒音環境の中で特定話者の発話のみを強調してその言語内容を理解するためには、誰が発話したのかを自動認識することが極めて重要なステップとなる [1]。

一般に話者認識システムは、音声認識システムと同様、front-end の特徴抽出器と back-end の分類器 (話者モデル) から構成されている。分類器の設計に対しては、ベイズ決定理論に基づく数理統計学的な設計手法が確立され、関連する数学の十分な整備に支えられて大いに発展している。一方、分類器の設計のみならず、話者認識にとって有効な話者特徴を取り出すことの出来る特徴抽出器を設計することも、実環境に耐えうる精度の高い話者認識のためには極めて重要である。

話者の特徴抽出法としては、音声認識の場合と同様、フィルタバンク分析や線形予測分析など、入力信号のスペクトル分析 (周波数分析) に基づく手法が一般的である。実環境下においては、騒音などの話者以外の音の特徴がスペクトル上に現れ、話者認識性能の劣化をもたらす。よって、認識器の性能向上に対して本質的な周波数帯域を選択あるいはワーピングすることが、ロバストな認識に対して極めて重要となってくる。話者認識ではなく、音声認識 (すなわち発話内容の認識) に対しては、メルスケールに基づく非線形周波数ワーピング等、音声スペクトルの低周波数帯域の解像度を大きくする特徴抽出法が有効な手法として定着している。これは、音韻を区別するための特徴が、ホルマント構造に代表される比較的低い周波数帯域に集中しているという、長い年月を経て培われてきた音声科学の知見に基づいている。しかし、話者特徴として有効な周波数帯域は、未だ十分に明らかにされていない。したがって多くの話者認識システムでは、周波数ワーピングを行わないリニアスケールの特徴抽出法や、せいぜいでも音声認識システムと同様の、低周波数帯域の解像度を大きくする特徴抽出手法が用いられている。明らかにこれらの特徴抽出法は、話者認識に対する最適性は保証されない。また一方で、比較的高い周波数帯域に話者固有の特徴が集中しているという報告もある [2], [3]。しかし、一般にこの帯域は信号対雑音比 (SNR) が低いため、雑音の影響を受けやすく、機械が実環境で話者を同定するには必ずしも最適とは言い難い。

このような従来の特徴抽出法の問題点は、特徴抽出器の設計 (有効な周波数帯域の選択) が後段の話者分類器 (話者モデル) の設計と異なる規準で行われており、抽出される特徴が認識器全体の性能にとって必ずしも最適ではないことに起因している。そこで本研究では、機械による話者自動認識にとって最適な話者特徴を見つけ出すことを目指し、話者認識システムの設計に識別的特徴抽出法 (DFE) [4] を応用する。DFE は、統計的パターン認識における新しい設計概念: 一般化確率的降下法 (GPD) [5]–[8] の拡張として開発された概念であり、front-end の特徴抽出器を back-end の分類器とともに最小認識誤りの目的で統一的に最適化するものである。文献 [9] では、話者認識において、ケプストラム係数ベクトルに対する変換行列を DFE に基づき学習している。しかし、有効な周波数帯域の発見を目指したものではない。ここでは、混合ガウスモデル (GMM) に基づく話者分類器を用いたテキスト独立型話者識別において、ある特定の周波数領域を強調することが可能な 2 次オールパス関数による周波数ワーピングを利用したメルケプストラム推定法 [10] を特徴抽出に適用し、DFE により、周波数ワーピングのパラメータおよび GMM パラメータの同時最適化を行う。このことにより、話者識別性能の向上のみならず、識別のための特徴として有効な周波数帯域を見つけ出すことができる。

女性 13 名、男性 22 名の話者識別実験により、本設計法は、従来の特徴量を用いて GMM のみの最尤推定や GMM のみの最小分類誤り学習を行った場合と比較し、より低い認識誤り率が得られ、パラメータ数の少ない簡略化されたモデルを用いた場合においても高い認識性能が得られ、同時に従来のリニアスケールやメルスケールと明らかに異なる周波数ワーピングが得られることを示す。このことは、システム全体の性能向上の目的と一貫した特徴抽出器の設計が極めて重要であり、それにより、従来の経験的な特徴とは異なる新たな特徴が得られ、同時に認識性能も向上することを示すものである。

## 2 話者認識システム

本研究では、混合ガウスモデル (GMM: Gaussian mixture model) を話者モデルとしたテキスト独立型の話者識別を行う。本節ではまず、一般的な話者認識手法について述べ、次に提案システムの基礎となる GMM に基づく話者モデルについて説明する。

### 2.1 話者識別

話者認識 (speaker recognition) は、話者識別 (speaker identification) と話者照合 (speaker verification) に分けられる。話者識別とは、入力音声があらかじめ登録してある話者のうち誰の声であるかを判定するものである。話者識別の応用例としては、会議録の自動ラベリングや犯罪捜査における容疑者の絞り込み、不特定話者音声認識のための話者固有の音韻モデルの選択的使用などが挙げられる。一方話者照合とは、入力音声の名乗った本人の音声であるか否かを、本人の音声の標準パターンと比較し、その類似度によって判定するものである。話者照合の応用例としては、バンキングサービス、個人情報へのアクセス制限など、セキュリティシステムにおける認証や、特定の話者のみに聞き耳を立てる音声理解システムへの応用などが考えられる。本研究ではこれらのうち話者識別を扱うが、後に説明する GPD 学習の誤分類測度などを変更すれば話者照合への応用も可能である。

### 2.2 テキスト独立型話者認識

話者認識の方法には、認識に用いるキーワードがあらかじめ限定されているテキスト依存型話者認識 (text-dependent speaker recognition) と、どのような発声内容でも認識ができるテキスト独立型話者認識 (text-independent speaker recognition) がある。また、これらの方法の中間的な手法として、認識の都度ユーザにキーワードを指定するテキスト指定型話者認識 (text-prompted speaker recognition) [11] がある。

テキスト依存型の場合は、キーワードが既知であるため、そのキーワードの時間パターンに関する情報も認識に利用することができ、その分高い認識率を得ることが容易となる。話者認識の多くは、このようにキーワードをそれぞれの人について固定することが可能であるが、必ずしも同じ言葉同士を比較することができないケースもある。このような目的には、テキスト独立型の手法を用いる必要がある。

一般的にテキスト依存型、指定型の場合は、隠れマルコフモデル (HMM: hidden Markov model) [12]–[14] や動的時間伸縮 (DTW: dynamic time warping) [15], [16] のように、キーワードの時間変化パターンを表現できるモデルを用いて認識を行う。一方、テキスト独立型話者認識においては、HMM のような時間構造を考慮した話者モデルでは、話者特徴のみならず言語的時間構造までもモデル化されてしまうため、言語的変動が大きくなってしまう。そこで、話者の音声の短時間パターンをベクトル量子化 (VQ: vector quantization) [17]–[19] や混合ガウスモデル (GMM) [20], [21] などを用いてモデル化し、入力特徴ベクトルの順序に依存しない方法で認識を行う。

### 2.3 混合ガウスモデル (GMM)

混合ガウスモデル (GMM)  $\lambda$  は、図 1 で表されるように、ガウス基底関数の線形結合により表現される確率モデルで、 $I$  個のガウス基底要素密度 (Gaussian component density) それぞれの混合重み (mixture component weight), 平均ベクトル (mean vector), 共分散行列 (covariance matrix) により完全に表現できる。

$$\lambda = \{p_i, \mu_i, \Sigma_i \mid i = 1, \dots, I\} \quad (1)$$

ガウス基底関数の線形結合は、複雑な話者の特徴ベクトルの分布表現に優れているため、近年話者認識においてよく用いられるようになった。GMM に状態遷移機能を付加すれば、音声認識でよく用いられている HMM である。したがって、GMM は 1 状態の連続混合 HMM (continuous mixture density HMM) とみなすこともできる。

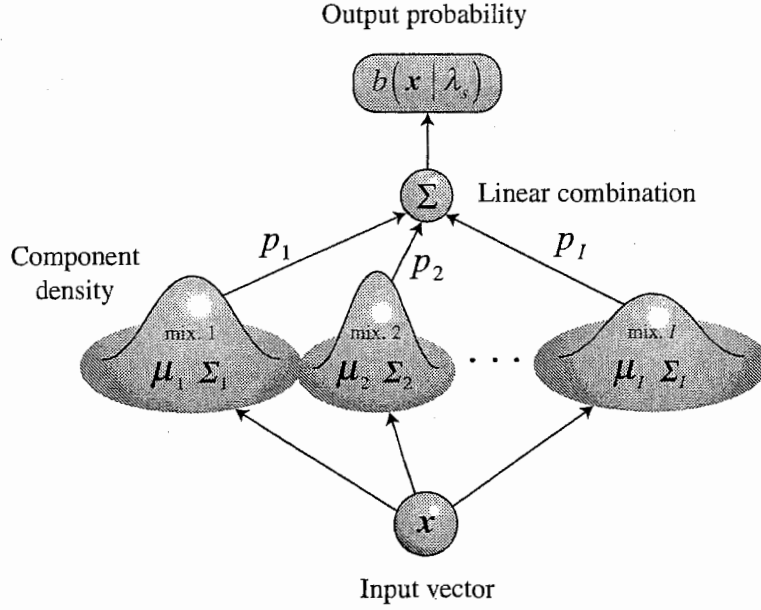


図1. Output probability of GMM.

モデル  $\lambda$  の観測ベクトル (observation vector)  $x$  に対する出力確率は次式で定義される。

$$b(x | \lambda) = \sum_{i=1}^I p_i \mathcal{N}(x | \mu_i, \Sigma_i) \quad (2)$$

ここで  $i$  は各混合要素のインデックス,  $I$  は混合数,  $p_i$  は要素  $i$  の混合要素重みで, 次式を満たす。

$$\sum_{i=1}^I p_i = 1 \quad (3)$$

$\mathcal{N}(x | \mu_i, \Sigma_i)$  は  $i$  番目の要素密度で, 次式のようにガウス基底関数で表される。

$$\mathcal{N}(x | \mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^D |\Sigma_i|}} \exp \left\{ -\frac{1}{2} (x - \mu_i)^T \Sigma_i^{-1} (x - \mu_i) \right\} \quad (4)$$

ここで,  $x$  は,  $D$  次元の任意のベクトルで,  $\mu_i$  は混合要素  $i$  の平均ベクトル,  $\Sigma_i$  は共分散行列である。  $|\Sigma_i|$ ,  $\Sigma_i^{-1}$  はそれぞれ共分散行列の行列式, 逆行列で,  $T$  は転置行列を表す。ここで共分散行列  $\Sigma_i$  は, 簡単のため, 通常的全共分散行列 (full covariance matrix) の代わりに対角成分以外が 0 である対角共分散行列 (diagonal covariance matrix) を用いる場合が多い。対角共分散を用いた場合, 全共分散行列に比べ自由パラメータ数が少なくなるため, 同じ混合数では性能が劣ってしまうが, これは混合数を多めに設定することで解決される [20]。  $\Sigma_i$  が対角共分散行列である場合, つまり,

$$\Sigma_i = \begin{bmatrix} \sigma_{i1}^2 & & 0 \\ & \sigma_{i2}^2 & \\ 0 & & \ddots \\ & & & \sigma_{iD}^2 \end{bmatrix} \quad (5)$$

であるとき, 式 (4) は式 (6) のように書き換えられる。

$$\mathcal{N}(x | \mu_i, \Sigma_i) = \frac{1}{\sqrt{(2\pi)^D \prod_{d=1}^D \sigma_{id}}} \exp \left\{ -\frac{1}{2} \sum_{d=1}^D \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right\} \quad (6)$$

## 2.4 GMMに基づく話者モデル

GMMに基づく話者認識システムにおいては, 各話者に対応する GMM, つまり, 式 (7) で表されるように, 平均ベクトル, 共分散行列, 混合重みの 3 つのパラメータから構成される  $I$  個のガウス混合要素が話者モデル

となる。

$$\lambda = \{p_i, \mu_i, \Sigma_i \mid i = 1, 2, \dots, I\} \quad (7)$$

認識では、与えられた観測ベクトル系列  $\mathbf{X}$  に対して  $S$  人の登録話者の GMM のモデルセット

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_S\} \quad (8)$$

の中から、式 (9) で表される最大の事後確率を出力する話者  $k$  を選ぶ。

$$k = \arg \max_s P(\lambda_s \mid \mathbf{X}) \quad (9)$$

$$= \arg \max_s \frac{P(\mathbf{X} \mid \lambda_s) P(\lambda_s)}{P(\mathbf{X})} \quad (10)$$

ここで、 $P(\lambda_s \mid \mathbf{X})$  は与えられた  $\mathbf{X}$  が話者  $s$  に属している事後確率、 $P(\lambda_s)$  は話者  $s$  の事前確率、 $P(\mathbf{X})$  は  $\mathbf{X}$  の生起確率、および、 $P(\mathbf{X} \mid \lambda_s)$  は話者  $s$  における  $\mathbf{X}$  の条件付き生起確率である。式 (10) は、式 (9) にベイズの定理を適用することで得られる。 $P(\mathbf{X})$  は  $\mathbf{X}$  が観測された後は定数となる。また、どの話者も同様に確からしい、つまり、 $P(\lambda_s) = 1/S$  であると仮定し、さらに対数をとると、式 (10) は、式 (11) のように簡単化できる。

$$k = \arg \max_s \log P(\mathbf{X} \mid \lambda_s) \quad (11)$$

したがって、図 2 に示すように、各登録話者のモデルにおけるテスト音声の特徴ベクトルに対する対数確率 (尤度) をそれぞれ計算し、この尤度が最大となる話者をテスト音声の発声話者であると判定すればよい。

入力音声から  $T$  フレームの特徴ベクトル

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \quad (12)$$

が得られた場合、モデル  $\lambda$  の  $\mathbf{X}$  に対する尤度  $P(\mathbf{X} \mid \lambda)$  は、次式で表される<sup>1</sup>。

$$P(\mathbf{X} \mid \lambda) = \prod_{t=1}^T b(\mathbf{x}_t \mid \lambda) \quad (13)$$

さらに、対数尤度  $\log P(\mathbf{X} \mid \lambda)$  は、次式となる。

$$\log P(\mathbf{X} \mid \lambda) = \sum_{t=1}^T \log b(\mathbf{x}_t \mid \lambda) \quad (14)$$

モデル  $\lambda$  に対する各特徴ベクトル  $\mathbf{x}_t$  の出力確率  $b(\mathbf{x}_t \mid \lambda)$  は、式 (2) により計算できる。

一般的に、GMM の作成には、EM アルゴリズムに代表される最尤推定法がよく用いられているが、本研究では最小分類誤り学習を主に扱っている。ただし、勾配探索に対する初期値設定には EM アルゴリズムを用いている。そこで、次に、EM アルゴリズムを用いた GMM の学習法について説明する。

## 2.5 EM アルゴリズムによる GMM の学習

ここでは、EM アルゴリズム (expectation maximization algorithm) を用いて対角共分散行列を用いた GMM を学習する方法について説明する。EM アルゴリズムを用いて、GMM の学習を行う場合、各混合要素の混合重み、平均ベクトル、共分散行列の 3 つのパラメータを推定する。

まず、モデルの各混合要素に混合重み、平均ベクトル、共分散行列の初期値を与える。これらの初期値としては、ベクトル量子化手法により、同じ学習データを用いて GMM の混合数と同じサイズのコードブックを作成し、各コードベクトルを GMM の平均ベクトルに、コードベクトルに属するデータ数の割合を混合重みに、コードベクトルに属するデータから計算した分散を共分散行列の対角要素にするなどの手法が考えられる。なお、本研究では、EM 学習前の GMM 各パラメータの初期値として、混合重みについては全要素均等の値 ( $1/I$ ) を与え、平均ベクトルは LBG アルゴリズム [22] を用いて作成したコードブックのコードベクトルを、分散は

<sup>1</sup> GMM に基づく確率過程モデルにおいては、ベクトル時系列  $\mathbf{x}_t$  は互いに統計的独立であると仮定される。



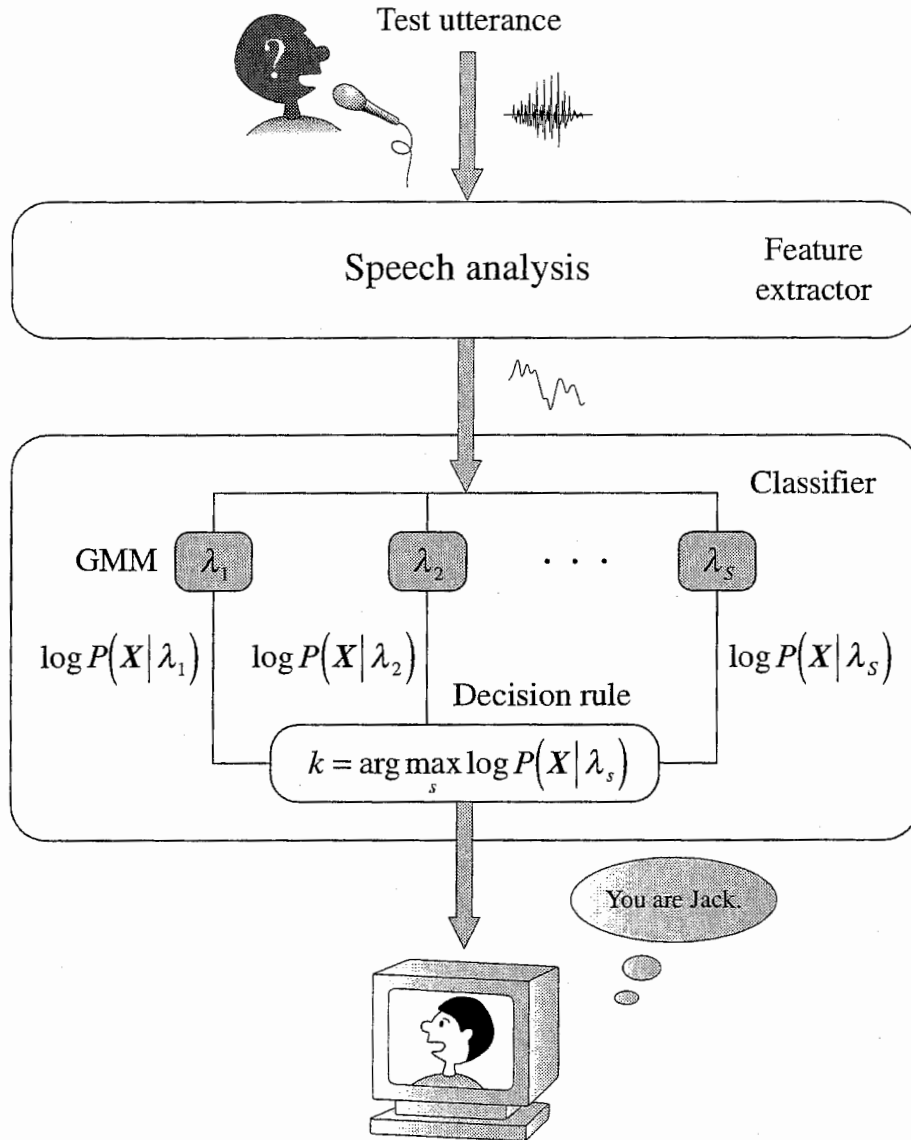


図 2. GMM based speaker identification system.

各コードベクトルに属する学習データの分散を与えている。次に以下の手続きを繰り返し、各混合要素のパラメータの更新を行う。

学習データベクトル系列

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T\} \quad (15)$$

が与えられた場合、 $\mathbf{x}_t$ が混合要素 $i$ に属している(事後)確率 $c(i | \mathbf{x}_t, \lambda)$ を、次式により計算する。

$$c(i | \mathbf{x}_t, \lambda) = \frac{p_i \mathcal{N}(\mathbf{x}_t | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_{j=1}^I p_j \mathcal{N}(\mathbf{x}_t | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)} \quad (16)$$

これを用いて各パラメータを更新する。まず、更新後の混合重み $p'_i$ は、

$$p'_i = \frac{1}{T} \sum_{t=1}^T c(i | \mathbf{x}_t, \lambda) \quad (17)$$

となり、更新後の平均ベクトル  $\mu'_i$  は、

$$\mu'_i = \frac{\sum_{t=1}^T c(i | \mathbf{x}_t, \lambda) \mathbf{x}_t}{\sum_{t=1}^T c(i | \mathbf{x}_t, \lambda)} \quad (18)$$

となる。また、対角共分散行列の各対角要素の更新式は、

$$\sigma'^2_{id} = \frac{\sum_{t=1}^T c(i | \mathbf{x}_t, \lambda) x_{td}^2}{\sum_{t=1}^T c(i | \mathbf{x}_t, \lambda)} - \mu'^2_{id}, \quad d = 1, 2, \dots, D \quad (19)$$

となる。ただし、 $\sigma'^2_{id}$  は対角共分散行列の  $d$  番目の対角要素、 $x_{td}$  は観測ベクトル  $\mathbf{x}_t$  の  $d$  次目の係数、 $\mu_{id}$  は平均ベクトル  $\mu_i$  の  $d$  次目の係数である。ここで、 $\sigma'^2_{id}$  が 0 になると、式 (6) において分母が 0 となるほか、逆行列も計算できないため、通常分散にフロアリング値 (flooring value) を設定し、分散がフロアリング値より小さくなった場合に、フロアリング値で置き換えるという操作をする。フロアリング値については、パラメータの種類などにより異なる値を経験的に設定する必要があるが、文献 [20] では、16 名を対象とした話者識別の予備実験から、メルケプストラム係数 (ここでは MFCC) を用いる場合、0.01–0.1 程度が適当であると述べている。

更新されたパラメータに対し、再び式 (16) を計算し、式 (17)–(19) を用いて各パラメータを更新する。この更新を繰り返すことにより、モデルの学習データに対する対数尤度

$$\log P(\mathbf{X} | \lambda) = \sum_{t=1}^T \log b(\mathbf{x}_t | \lambda) = \sum_{t=1}^T \log \left( \sum_{i=1}^I p_i \mathcal{N}(\mathbf{x}_t | \mu_i, \Sigma_i) \right) \quad (20)$$

が増化していく。この対数尤度の変化がある閾値以下になるまで更新を繰り返す。Reynolds は文献 [21] で、繰り返し回数は 10 回程度で十分であると述べている。

### 3 話者の特徴抽出 —メルケプストラム推定法—

本研究では、話者の特徴量としてメルケプストラム推定法によるケプストラム係数を用いる。一般にメルケプストラムと呼ばれるものには、線形予測分析 (linear predictive analysis) による LPC ケプストラム (LPCC) をメルスケールに変換したものや [23], [24], メルスケール上でのフィルタバンク分析 (filter-bank analysis) によるケプストラム (MFCC) [27] があるが、ここで扱うものは周波数変換されたケプストラムであり、これらとは異なる。なお、ここではメルケプストラム推定法と呼んでいるが、得られるケプストラム係数は必ずしもメルスケールとは限らない。

本節では、まずはじめに、従来の分析法である1次オールパス関数を周波数変換に用いた場合のメルケプストラム推定法 [28] について説明する。次に、周波数ワーピングの度合のみならず強調周波数をも調整可能な、2次オールパス関数による周波数変換を利用したメルケプストラム推定法について説明する。

#### 3.1 1次オールパス関数による周波数変換

まず、音声スペクトル  $H(e^{j\omega})$  を、 $M+1$  個のメルケプストラム係数  $c(m)$  を用いて次のように定義する。

$$H(z) = \exp \sum_{m=0}^M c(m) \tilde{z}^{-m} \quad (21)$$

ここで、システムの伝達関数は式 (22) で表される1次オールパス関数で、その周波数特性は式 (23) で表せる。

$$\tilde{z}^{-1} = A(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}, \quad |\alpha| < 1 \quad (22)$$

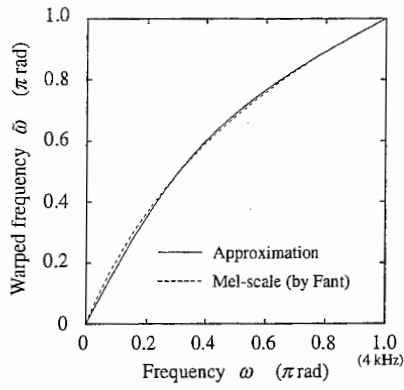
$$A(e^{j\omega}) = e^{-j\tilde{\omega}(\omega, \alpha)} \quad (23)$$

また、位相特性は式 (24) で表され、この位相特性において適切な  $\alpha$  の値を選択することによって、メルスケール [29] やバーク (Bark) スケール [30] などの人間の聴覚特性を高い精度で近似することができる。メルスケール (by G. Fant) およびバークスケール (by S. Seneff) を近似するための  $\alpha$  の値の例を表 1 に、そのときの位相特性を図 3 示す。

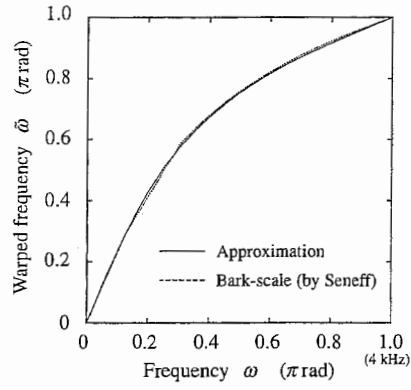
$$\begin{aligned} \tilde{\omega}(\omega, \alpha) &= -\arg [A(e^{j\omega})] \\ &= \tan^{-1} \frac{(1 - \alpha^2) \sin \omega}{(1 + \alpha^2) \cos \omega - 2\alpha} \end{aligned} \quad (24)$$

表 1. Examples of  $\alpha$  for approximating the mel and Bark scales.

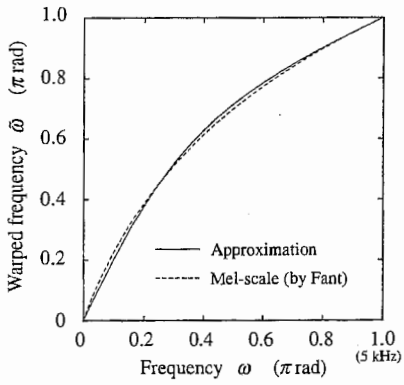
Sampling frequency	8kHz	10kHz	12kHz	16kHz
Mel scale	0.31	0.35	0.37	0.42
Bark scale	0.42	0.47	0.50	0.55



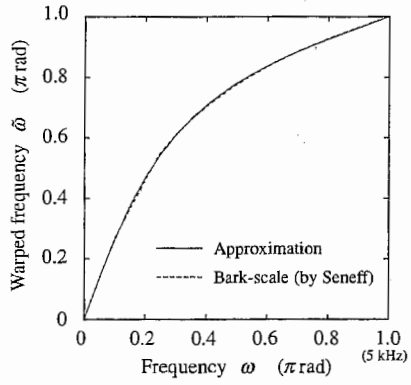
(a) Mel scale (8kHz)



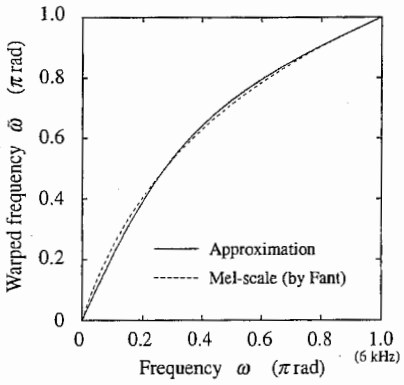
(b) Bark scale (8kHz)



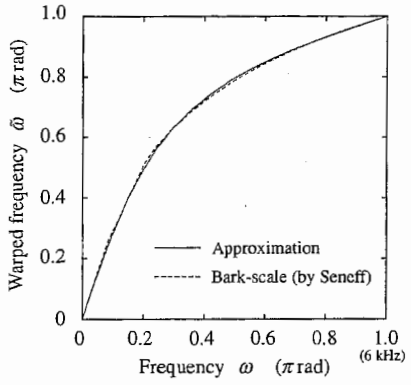
(c) Mel scale (10kHz)



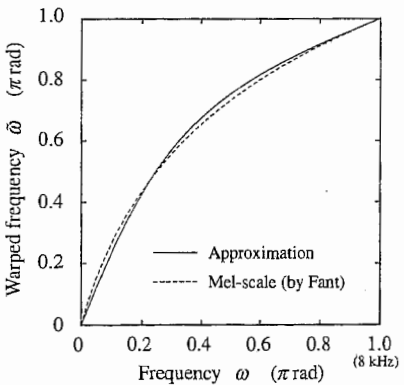
(d) Bark scale (10kHz)



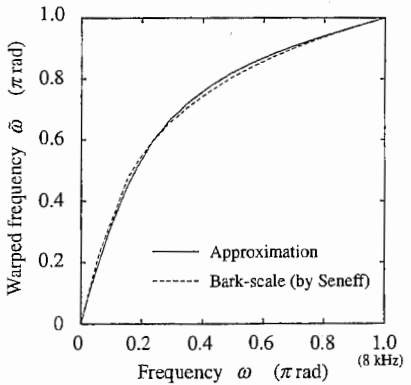
(e) Mel scale (12kHz)



(f) Bark scale (12kHz)



(g) Mel scale (16kHz)



(h) Bark scale (16kHz)

図 3. Examples of warping functions for approximating the mel and Bark scales.

### 3.2 2次オールパス関数による周波数変換

次に、2次オールパス関数による周波数変換を利用したメルケプストラム推定法について説明する。この分析法では、次式で表される2次オールパス関数の1/2乗  $A(z)$  をシステムの伝達関数として考える。

$$A(z) = \left( \frac{z^{-2} - 2\alpha \cos \theta z^{-1} + \alpha^2}{1 - 2\alpha \cos \theta z^{-1} + \alpha^2 z^{-2}} \right)^{\frac{1}{2}}, \quad |\alpha| < 1, \quad 0 \leq \theta \leq \pi \quad (25)$$

ここで、次のように  $\theta = 0$  と置くと、式 (25) の伝達関数  $A(z)$  は、式 (22) で表される1次オールパスフィルタの伝達関数に一致する。よって、この2次オールパス関数に基づく周波数変換は、1次オールパス関数による周波数変換を含んでいる。

$$\begin{aligned} A(z) &= \left( \frac{z^{-2} - 2\alpha \cos \theta z^{-1} + \alpha^2}{1 - 2\alpha \cos \theta z^{-1} + \alpha^2 z^{-2}} \right)^{\frac{1}{2}} \Bigg|_{\theta=0} \\ &= \left( \frac{z^{-2} - 2\alpha z^{-1} + \alpha^2}{1 - 2\alpha z^{-1} + \alpha^2 z^{-2}} \right)^{\frac{1}{2}} \\ &= \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}} \end{aligned} \quad (26)$$

このシステムの位相特性は次式で表される。

$$\begin{aligned} \tilde{\omega}(\omega, \theta, \alpha) &= -\arg [A(e^{j\omega})] \\ &= \omega + \tan^{-1} \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right) + \tan^{-1} \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right) \end{aligned} \quad (27)$$

ここで、パラメータ  $\alpha$  はワープさせる度合、 $\theta$  は強調する中心角周波数を表している。この関数  $\tilde{\omega}(\omega, \theta, \alpha)$  は、周波数  $\omega$  ( $0 \leq \omega \leq \pi$ ) を  $\tilde{\omega}$  ( $0 \leq \tilde{\omega} \leq \pi$ ) へ変換する単調な関数である、2次オールパス関数に基づくシステムの位相特性  $\tilde{\omega}(\omega, \theta, \alpha)$  の例を付録 A に示す。この位相特性は次の関係式を満たす。

$$\tilde{\omega}(\omega, \theta, \alpha) = \tilde{\omega}(\omega, \pi - \theta, -\alpha) \quad (28)$$

以上のような2次オールパス関数を利用して、音声スペクトル次式のように表現する。

$$\begin{aligned} |H(e^{j\omega})| &= \exp \sum_{m=0}^M c(m) \cos(\tilde{\omega}m) \\ &= \exp \sum_{m=0}^M c(m) \frac{A^m(e^{j\omega}) + A^m(e^{-j\omega})}{2} \end{aligned} \quad (29)$$

ここで、 $c(m)$  はメルケプストラム係数を表す。メルケプストラム推定法では、この音声のスペクトルモデルをもとに、入力音声に対して尤度を最大化するようにケプストラム係数を推定する。次に、メルケプストラム係数の推定法について説明する。

### 3.3 メルケプストラム係数の推定

メルケプストラム係数を求めるためには、式 (30) のようなスペクトル評価関数を用いる。

$$E = \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\exp R(\omega) - R(\omega) - 1\} d\omega \quad (30)$$

ここで、 $R(\omega)$  は式 (31) で表されるもので、 $I_N(\omega)$  は式 (32) のように入力音声  $s(n)$  と窓幅  $N$  の時間窓関数  $w(n)$  から計算されるピリオドグラムである。

$$\begin{aligned} R(\omega) &= \log \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \\ &= \log I_N(\omega) - \log |H(e^{j\omega})|^2 \end{aligned} \quad (31)$$

$$I_N(\omega) = \frac{\left| \sum_{n=0}^{N-1} w(n) s(n) e^{-j\omega n} \right|^2}{\sum_{n=0}^{N-1} w^2(n)} \quad (32)$$

ここで、入力信号

$$\mathbf{s} = [s(0) \ s(1) \ \cdots \ s(N-1)]^T \quad (33)$$

が与えられたとき、推定すべきメルケプストラム係数を

$$\mathbf{c} = [c(0) \ c(1) \ \cdots \ c(M)]^T \quad (34)$$

とする。メルケプストラム係数  $\mathbf{c}$  は、式 (30) で表されるスペクトル評価関数  $E$  を  $\mathbf{c}$  に関して最小化することで推定できる。なお、この最小化は、 $s(n)$  を平均 0 のガウス過程とみなした場合に、尤度  $P(\mathbf{s} | \mathbf{c})$  を  $\mathbf{c}$  に関して最大化することに等価である [31]。式 (21) で表されるスペクトルモデルを用いた場合、 $E$  は  $\mathbf{c}$  に関して下に凸であり、

$$\nabla_{\mathbf{c}} E = \frac{\partial E}{\partial \mathbf{c}} = \mathbf{0} \quad (35)$$

を満たすとき、大域的な最小値をとる。したがって、 $E$  の最小化問題は、ニュートン・ラフソン法 (Newton-Raphson method) などの繰り返し計算に基づいた最適化手法により解くことができる。ここで、2次オールパス関数により定義された基底関数を用いてスペクトル推定を行う場合についても、スペクトル評価関数  $E$  は  $\mathbf{c}$  に関して下に凸であり、1次オールパス関数の場合と同様に、大域的な一意解を最適化手法により容易に得られることがわかっている [10]。

そこで、ニュートン・ラフソン法の  $\tau$  番目の繰り返しステップにおいて、次の線形方程式を解く。

$$\mathbf{H} \Delta \mathbf{c}^{(\tau)} = -\nabla_{\mathbf{c}} E \Big|_{\mathbf{c}=\mathbf{c}^{(\tau)}} \quad (36)$$

ここで  $\mathbf{H}$  は、式 (37) で表されるような評価関数  $E$  の  $\mathbf{c}$  に関する 2 階微分で、ヘッセ行列 (Hessian matrix) と呼ばれる。 $\mathbf{H}$  は、 $E$  が  $\mathbf{c}$  に関して下に凸であることから、常に正定値行列 (positive definite matrix) となる。

$$\mathbf{H} = \nabla_{\mathbf{c}}^2 E = \frac{\partial^2 E}{\partial \mathbf{c} \partial \mathbf{c}^T} \quad (37)$$

$\mathbf{c}^{(\tau)}$  は、 $\tau$  番目の繰り返しにおいて得られる  $\mathbf{c}$  の推定値で、各繰り返しにおいて、線形方程式 (36) の解として

$$\Delta \mathbf{c}^{(\tau)} = [\Delta c^{(\tau)}(0) \ \Delta c^{(\tau)}(1) \ \cdots \ \Delta c^{(\tau)}(M)]^T \quad (38)$$

が得られ、これを用いて  $\mathbf{c}$  を次式のように更新する。

$$\mathbf{c}^{(\tau+1)} = \mathbf{c}^{(\tau)} + \Delta \mathbf{c}^{(\tau)} \quad (39)$$

メルケプストラム係数の初期値  $\mathbf{c}^{(0)}$  としては、FFT ケプストラムを文献 [32] の再帰式を用いて周波数変換したものや、1つ手前のフレームにおいて得られたメルケプストラム係数を用いることが考えられる。なお、この手法においては、数回の繰り返しで、十分な精度のメルケプストラム係数を推定できることがわかっている。

### 3.4 最小化問題の解法

線形方程式 (36) を解くためには、評価関数  $E$  の 1 階微分  $\nabla_{\mathbf{c}} E$  および 2 階微分  $\nabla_{\mathbf{c}}^2 E$  が必要である。そこで、まず 1 階微分を計算する。式 (29) より、対数パワースペクトルは

$$\begin{aligned} \log |H(e^{j\omega})|^2 &= \sum_{m=0}^M c(m) \{A^m(e^{j\omega}) + A^m(e^{-j\omega})\} \\ &= \sum_{m=0}^M c(m) \{e^{j\tilde{\omega}m} + e^{-j\tilde{\omega}m}\} \\ &= \mathbf{c}^T (\mathbf{z} + \mathbf{z}^*) \end{aligned} \quad (40)$$

となる。ここで、 $\mathbf{c}$  は  $M$  次のメルケプストラム係数、 $\mathbf{z}$ ,  $\mathbf{z}^*$  は次式で表される互いに共役なベクトルである。

$$\mathbf{c} = [c(0) \ c(1) \ \cdots \ c(M)]^T \quad (41)$$

$$\mathbf{z} = [1 \ e^{-j\tilde{\omega}} \ e^{-2j\tilde{\omega}} \ \cdots \ e^{-j\tilde{\omega}M}]^T \quad (42)$$

$$\mathbf{z}^* = [1 \ e^{j\tilde{\omega}} \ e^{2j\tilde{\omega}} \ \cdots \ e^{j\tilde{\omega}M}]^T \quad (43)$$

そこで、式 (40) を式 (30) で表されるスペクトル評価関数に代入すると次式となる。

$$\begin{aligned} E &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\exp R(\omega) - R(\omega) - 1\} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\exp \{\log I_N(\omega) - \log |H(e^{j\omega})|^2\} - \log I_N(\omega) + \log |H(e^{j\omega})|^2 - 1\} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \{\exp \{\log I_N(\omega) - \mathbf{c}^T (\mathbf{z} + \mathbf{z}^*)\} - \log I_N(\omega) + \mathbf{c}^T (\mathbf{z} + \mathbf{z}^*) - 1\} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \{I_N(\omega) \exp \{-\mathbf{c}^T (\mathbf{z} + \mathbf{z}^*)\} - \log I_N(\omega) + \mathbf{c}^T (\mathbf{z} + \mathbf{z}^*) - 1\} d\omega \end{aligned} \quad (44)$$

よって、 $E$  の  $\mathbf{c}$  に関する 1 階微分は次のようになる。

$$\begin{aligned} \nabla_{\mathbf{c}} E &= \nabla_{\mathbf{c}} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} \{I_N(\omega) \exp \{-\mathbf{c}^T (\mathbf{z} + \mathbf{z}^*)\} - \log I_N(\omega) + \mathbf{c}^T (\mathbf{z} + \mathbf{z}^*) - 1\} d\omega \right] \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \{I_N(\omega) \exp \{-\mathbf{c}^T (\mathbf{z} + \mathbf{z}^*)\} \{-(\mathbf{z} + \mathbf{z}^*)\} + (\mathbf{z} + \mathbf{z}^*)\} d\omega \\ &= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \{I_N(\omega) \exp \{-\mathbf{c}^T (\mathbf{z} + \mathbf{z}^*)\} - 1\} (\mathbf{z} + \mathbf{z}^*) d\omega \end{aligned} \quad (45)$$

$$\begin{aligned} &= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} (\mathbf{z} + \mathbf{z}^*) d\omega \\ &= -2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} \mathbf{z}^* d\omega \\ &= -2(\mathbf{r} - \mathbf{a}) \end{aligned} \quad (46)$$

ただし、 $\mathbf{r}$  および  $\mathbf{a}$  はそれぞれ次のように定義する。

$$\mathbf{r} = [r(0) \ r(1) \ \cdots \ r(M)]^T \quad (47)$$

$$\mathbf{a} = [a(0) \ a(1) \ \cdots \ a(M)]^T \quad (48)$$

$$r(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} e^{j\tilde{\omega}m} d\omega \quad (49)$$

$$a(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\tilde{\omega}m} d\omega \quad (50)$$

これを各係数ごとに表すと次式のようになる。

$$\begin{aligned}\frac{\partial E}{\partial c(m)} &= -2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} e^{j\omega m} d\omega \\ &= -2 \{r(m) - a(m)\}, \quad m = 0, 1, \dots, M\end{aligned}\quad (51)$$

さらに、 $E$ の $c$ に関する2階微分は次式となる。

$$\begin{aligned}H &= \nabla_c^2 E = \nabla_{c^T} \left[ -2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} z^* d\omega \right] \\ &= \nabla_{c^T} \left[ -2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ I_N(\omega) \exp \{-c^T(z + z^*)\} - 1 \right\} z^* d\omega \right] \\ &= -2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ I_N(\omega) \exp \{-c^T(z + z^*)\} z^* \left\{ -(z + z^*)^T \right\} \right\} d\omega \\ &= 2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ I_N(\omega) \exp \{-c^T(z + z^*)\} z^* (z + z^*)^T \right\} d\omega \\ &= 2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} z^* (z + z^*)^T d\omega \\ &= 2 \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} z^* z^T d\omega + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} z^* z^{*T} d\omega \right\} \\ &= 2(\mathbf{R} + \mathbf{Q})\end{aligned}\quad (52)$$

ただし、 $\mathbf{R}$ および $\mathbf{Q}$ は、次に示すような対称行列で、それぞれ Toeplitz 行列、Hankel 行列に対応する。

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & \cdots & r(M) \\ r(1) & r(0) & \cdots & r(M-1) \\ \vdots & \vdots & \ddots & \vdots \\ r(M) & r(M-1) & \cdots & r(0) \end{bmatrix}\quad (53)$$

$$\mathbf{Q} = \begin{bmatrix} r(0) & r(1) & \cdots & r(M) \\ r(1) & r(2) & \cdots & r(M+1) \\ \vdots & \vdots & \ddots & \vdots \\ r(M) & r(M+1) & \cdots & r(2M) \end{bmatrix}\quad (54)$$

また、2階微分の行列の各成分は次式で表せる。

$$\begin{aligned}\frac{\partial^2 E}{\partial c(m) \partial c(n)} &= 2 \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} e^{j\omega(m-n)} d\omega + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} e^{j\omega(m+n)} d\omega \right\} \\ &= 2 \{r(m-n) + r(m+n)\}, \quad m, n = 0, 1, \dots, M\end{aligned}\quad (55)$$

ここで、線形方程式(36)は、 $\mathbf{R}$ 、 $\mathbf{Q}$ 、 $\mathbf{r}$ 、 $\mathbf{a}$ を用いて次のように書き換えられる。

$$(\mathbf{R} + \mathbf{Q}) \Delta \mathbf{c}^{(\tau)} = \mathbf{r} - \mathbf{a} \Big|_{\mathbf{c} = \mathbf{c}^{(\tau)}}\quad (56)$$

対称な Toeplitz + Hankel 行列を用いたこの線形方程式は、例えば文献[33]における高速アルゴリズムを用いて解くことができる。あとは $r(m)$ 、 $a(m)$ の値さえ計算できればよい。そこで、まず以下のような逆フーリエ変換 $r'(m)$ 、 $a'(m)$ を行う。

$$\begin{aligned}r'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} e^{j\omega m} d\omega \\ &\approx \frac{1}{N} \sum_{k=0}^{N-1} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} e^{j \frac{2\pi}{N} km} = \mathcal{F}_{(m)}^{-1} \left[ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \right]\end{aligned}\quad (57)$$



$$\begin{aligned}
 a'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega m} d\omega \\
 &\approx \frac{1}{N} \sum_{k=0}^{N-1} e^{j\frac{2\pi}{N} km} = \mathcal{F}_{(m)}^{-1}[1]
 \end{aligned} \tag{58}$$

$r(m)$  および  $a(m)$  は,  $r'(m)$  および  $a'(m)$  の値をそれぞれある行列を用いて変換することで求められる. 詳しい変換方法については, 後に付録 B(実偶関数の場合) にて紹介する.

なお, 本研究では, メルケプストラム係数の初期値としてケプストラム係数をメルスケールに周波数変換したものをを用いるが, 2次オールパス関数を用いる場合, 1次オールパス関数を用いる場合と同様の再帰式によって変換を行うことができないため, これについても行列を用いて変換する必要がある. また, メルケプストラム推定の各繰り返しにおいてパワースペクトルの値を計算する必要があるため, 逆にメルケプストラム係数からケプストラム係数に変換する操作も必要となる. ケプストラム係数  $c'(m)$  とそれに対応するメルケプストラム係数を  $c(m)$  は,

$$c(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |H(e^{j\omega})|^2 e^{j\omega m} d\omega \tag{59}$$

$$\begin{aligned}
 c'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \log |H(e^{j\omega})|^2 e^{j\omega m} d\omega \\
 &\approx \frac{1}{N} \sum_{k=0}^{N-1} \log |H(e^{j\omega})|^2 e^{j\frac{2\pi}{N} km} = \mathcal{F}_{(m)}^{-1} \left[ \log |H(e^{j\omega})|^2 \right]
 \end{aligned} \tag{60}$$

で与えられる. ここで, 先ほどの  $r(m)$  や  $a(m)$  の式では, 最後の部分が  $d\omega$  であったのに対し,  $c(m)$  の最後の部分は  $d\omega$  となっているため,  $c'(m)$  から  $c(m)$  を求める場合には,  $r'(m)$  および  $a'(m)$  から  $r(m)$  および  $a(m)$  を求めるときに用いたものと同じ行列を用いて変換することはできない. そこで,  $c'(m)$  と  $c(m)$  の双方向の変換については, 付録 C にて説明する.

最後に, メルケプストラム推定法によるスペクトル推定例を示す. 図4の左側は, 上からリニアスケール, メルスケール,  $(\theta, \alpha) = (0.5\pi, 0.4)$ ,  $(\theta, \alpha) = (0.5\pi, 0.6)$  の周波数伸縮関数を表しており, その右側に左側の周波数伸縮関数に対応したスペクトル推定例を示している. 図のように, 周波数伸縮関数が急勾配となる周波数部分では, スペクトルも細かく推定されていることがわかる.

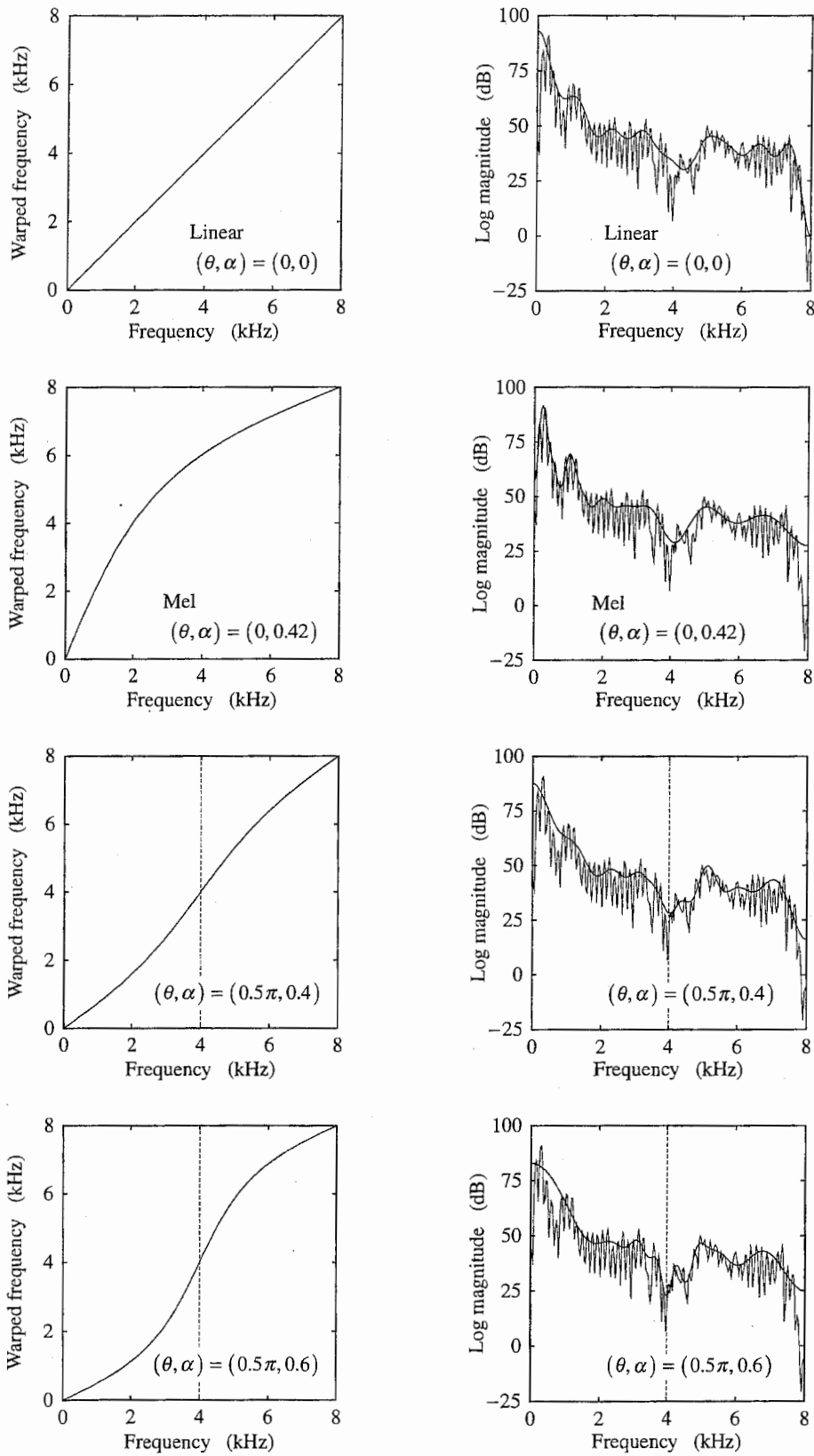


図 4. Left: Warping functions, Right: Spectral estimats.

## 4 最小分類誤りのための識別的特徴抽出法

### 4.1 識別的特徴抽出法 (DFE)

多くのパターン認識システムは、特徴抽出器 (feature extractor) と分類器 (classifier) とで構成されている。従来、特徴抽出のためのパラメータ  $\Theta$  については、分類器のパラメータ  $\Lambda$  と独立に経験的な方法で定められることが多かった。そこで、最小分類誤り/一般化確率の降下法 (MCE/GPD: minimum classification error/generalized probabilistic descent) の拡張として、特徴抽出のためのパラメータ  $\Theta$  についても分類器のパラメータ  $\Lambda$  と合わせて、 $\Phi = \{\Theta, \Lambda\}$  というパラメータセットとして最小分類誤り規準にて学習する識別的特徴抽出法 (DFE: discriminative feature extraction) が提案されている。ここでは、話者識別システムのための DFE 学習について説明する。

### 4.2 話者識別のための DFE 学習則

ここでは、2次オールパス関数による周波数変換を利用したメルケプストラムを特徴量として用いた、GMM 話者モデルに基づく話者識別システムにおいて、GMM のパラメータセットおよび周波数ワーピングパラメータを識別的特徴抽出法に基づき学習する方法について述べる。

DFE で最適化を行うべきパラメータセット  $\Phi$  は、特徴抽出のためのパラメータセット  $\Theta$  および識別部のパラメータセット  $\Lambda$  である。

$$\Phi = \{\Theta, \Lambda\} \quad (61)$$

ここで、 $\Theta$  は周波数ワーピング関数の強調する中心周波数  $\theta$  および伸縮する度合  $\alpha$  の 2 つである。 $\Lambda$  は  $S$  人の識別対象話者の GMM のモデルのパラメータセットである。

$$\Theta = \{\theta, \alpha\} \quad (62)$$

$$\Lambda = \{\lambda_1, \lambda_2, \dots, \lambda_S\} \quad (63)$$

音声データ  $S$  が与えられたとき、特徴抽出のためのパラメータ  $\Theta$  によって、特徴ベクトル系列  $X$  が得られたとする。

$$X = \mathcal{E}_\Theta(S) \quad (64)$$

この特徴ベクトル  $X$  に対する話者  $s$  (クラス  $C_s$ ) の判別関数 (discriminant function) は、話者  $s$  の GMM  $\lambda_s$  の平均対数尤度 (対数尤度を特徴ベクトル系列  $X$  のベクトル数  $T$  で割って、1 特徴ベクトル当たりの平均したもの) として次式のように定義する。

$$g_s(X; \Lambda) = \frac{1}{T} \log P(X | \lambda_s) \quad (65)$$

次に、次式で表される誤分類測度 (misclassification measure) を導入する。

$$d_k(X; \Lambda) = -g_k(X; \Lambda) + G_k(X; \Lambda), \quad X \in C_k \quad (66)$$

ただし、入力ベクトル  $X$  は、話者  $k$  (クラス  $C_k$ ) に属するものとする。式 (66) の右辺第 2 項は、antidiscriminant function と呼ばれるもので、 $k$  以外の話者の尤度を用いて、 $L_p$  ノルム形式で次式のように表せる。

$$G_k(X; \Lambda) = \log \left[ \frac{1}{S-1} \sum_{s \neq k} \exp \{g_s(X; \Lambda)\eta\} \right]^{\frac{1}{\eta}}, \quad \eta > 0 \quad (67)$$

ここで、 $\eta$  は正の定数で、 $S$  は話者 (クラス) の総数を表す。誤分類測度  $d_k(X; \Lambda)$  の値が正である場合、 $X$  が間違っ て分類されたことを意味し、負の場合、正しく分類されたことを意味する。すなわち、この誤分類測度によって、分類則における判別関数の比較操作をスカラーの値の大小比較に置き換えることができたわけである。  $\eta$  の値を無限大に設定すると、

$$\eta \rightarrow \infty \implies G_k(X; \Lambda) \rightarrow \max_{s \neq k} g_s(X; \Lambda) \quad (68)$$

となるので、式 (65) の誤分類測度は

$$d_k(\mathbf{X}; \Lambda) = -g_k(\mathbf{X}; \Lambda) + g_j(\mathbf{X}; \Lambda) \quad (69)$$

となる。ただし、クラス  $C_j$  は、クラス  $C_k$  以外で最大の判別関数の値を示すクラスを表す。つまり、

$$j = \arg \max_{s \neq k} g_s(\mathbf{X}; \Lambda) \quad (70)$$

である。このとき、分類判断における比較操作は、 $X$  が属するクラス  $C_k$  と、それ以外でもっとも可能性の高い、つまり判別関数の値が大きいクラス  $C_j$  との間のみで行われる。この操作によって、計算量が減ることは明らかである。性能面においても、一対多型の誤分類測度を用いた場合と比べてそれほど大きな差はないと言われており、本研究においても  $\eta$  を無限大とし、一対一型の誤分類測度を用いることにする。判別関数が定まったので、あとは、ある平滑な損失を選択し、その損失からなる経験的損失 (empirical loss) の最小化を行えばよい。損失関数は、例えば、

$$\ell_k(\mathbf{X}; \Lambda) = \begin{cases} 1, & \text{misclassification} \\ 0, & \text{otherwise} \end{cases} \quad (71)$$

のように、正しく分類された場合に 0、誤って分類された場合に 1 をを選ぶのが自然であるが、これでは明らかに非平滑である。そこで、これを近似するような滑らかな関数、例えば、次式で表されるようなシグモイド関数 (sigmoid function) を損失関数 (loss function) として用いる。

$$\ell_k(\mathbf{X}; \Lambda) = \ell_k(d_k) = \frac{1}{1 + \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) - \beta)\}}, \quad \gamma > 0 \quad (72)$$

ここで、 $\gamma$  はシグモイド関数の傾きを表し、値が小さくなるほど学習の収束も遅くなる。 $\beta$  はシグモイド関数のオフセットに対応する。シグモイド関数の例を図 5 に示す。

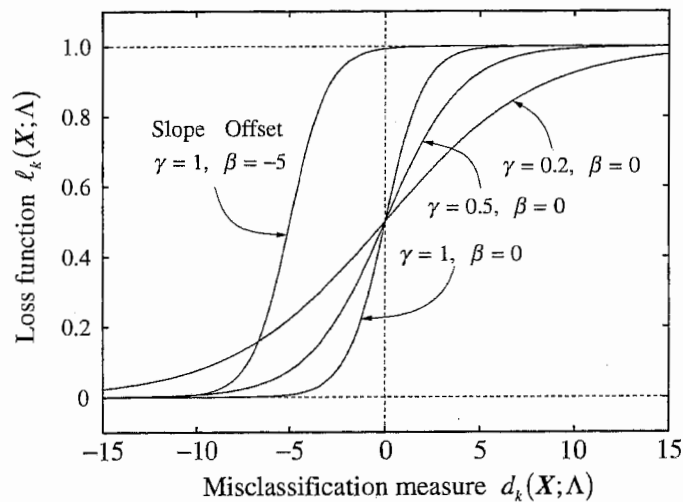


図 5. Examples of sigmoid loss functions.

そこで、学習音声データセット

$$S = \{S_1, S_2, \dots, S_H\} \quad (73)$$

が与えられ、特徴抽出パラメータ  $\Theta$  によって、

$$\mathcal{X} = \mathcal{E}_\Theta(S) \quad (74)$$

と特徴抽出が行われ、

$$\mathcal{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_H\} \quad (75)$$

が得られたとすると、経験的損失は、式 (72) の損失関数を用いて次のように定義できる。

$$L(\mathcal{X}; \Lambda) = \frac{1}{H} \sum_{h=1}^H \sum_{s=1}^S \ell_s(\mathbf{X}_h; \Lambda) \cdot 1(\mathbf{X}_h \in C_s) \quad (76)$$

ここで、 $1(\cdot)$  は indicator 関数と呼ばれるもので、次式のように、引数が真となるとき 1 をとる関数である。

$$1(S) = \begin{cases} 1, & \text{if } S \text{ is true} \\ 0, & \text{otherwise} \end{cases} \quad (77)$$

式 (72) の損失関数の場合、式 (76) の  $L(\mathcal{X}; \Lambda)$  は、分類誤り率を十分に近似する。パラメータセット  $\Phi$  は、勾配法に基づき次式のように更新される。

$$\Phi^{(\tau+1)} = \Phi^{(\tau)} - \varepsilon^{(\tau)} V^{(\tau)} \nabla_{\Phi} L(\mathcal{X}; \Lambda) \Big|_{\Phi = \Phi^{(\tau)}} \quad (78)$$

ここで、 $\varepsilon^{(\tau)}$  は学習係数 (learning rate あるいは learning step size) で、正の値をとり、学習の繰り返し回数に伴い単調減少させる。 $V^{(\tau)}$  は、学習行列 (learning matrix) と呼ばれる正定値行列で、ニュートン・ラフソン法を適用するためには 2 階微分の逆行列であるが、損失関数の 2 階微分の計算が困難な場合は、対角共分散行列の逆行列や単位行列で代用したり、あるいは、学習係数をパラメータ別に調整する場合もある。学習係数としてよく用いられるものの一つは、次に示すような、 $\varepsilon_0$  を初期値として学習回数に従い線形的に単調減少する関数である。本研究でもこの関数を用いている。

$$\varepsilon^{(\tau)} = \varepsilon_0 \left(1 - \frac{\tau}{T}\right) \quad (79)$$

ここで、 $T$  は総学習回数で、 $\varepsilon_0, T$  については経験的に定める。

ところで、MCE/GPD や DFE の学習についても、ニューラルネットの学習などと同様に、一括型 (batch mode あるいは epoch mode) の学習方法と、逐次型 (sample by sample mode) の学習方法がある。一括型とは、式 (78) で表されるように、全ての学習データを用いて更新量を累積し、各エポックの最後に全パラメータについて一度にまとめて更新する方法である。また逐次型とは、学習データを 1 つずつ用いて逐次にパラメータを更新する方法で、全データ ( $\mathbf{X} \in \mathcal{X}$ ) について次式を用いて一通り更新を行い、これを 1 エポックとする方法である。

$$\Phi^{(\rho+1)} = \Phi^{(\rho)} - \varepsilon^{(\rho)} V^{(\rho)} \nabla_{\Phi} \ell_k(\mathbf{X}; \Lambda) \Big|_{\Phi = \Phi^{(\rho)}} \quad (80)$$

なお、逐次型の学習を行う場合は、学習データを用いる順番が固定であると、エポックの終わりの方で使われるクラスが優勢となり、学習に偏りが生じるため、学習データは各エポックごと無作為に並べ換える必要がある。

DFE では、特徴抽出パラメータセット  $\Theta$  とモデルパラメータセット  $\Lambda$  の 2 つを勾配法に基づき更新すると述べたが、元々性質の異なる 2 つのパラメータセットを同じ学習係数を用いて同時に更新してもうまく学習できるとは限らない。この問題を解決するための方法として、特徴抽出のパラメータおよびモデルパラメータについて異なる学習係数を設定したり、一方を固定で交互に学習する方法などが挙げられる。ここでは、式 (81)、(82) のように、 $\Theta$  と  $\Lambda$  について別々の学習係数を設け、交互に学習することにする。

$$\Theta^{(\rho+1)} = \Theta^{(\rho)} - \varepsilon^{(\rho)} V^{(\rho)} \nabla_{\Theta} \ell_k(\mathbf{X}; \Lambda) \Big|_{\Theta = \Theta^{(\rho)}} \quad (81)$$

$$\Lambda^{(\rho+1)} = \Lambda^{(\rho)} - \epsilon^{(\rho)} U^{(\rho)} \nabla_{\Lambda} \ell_k(\mathbf{X}; \Lambda) \Big|_{\Lambda = \Lambda^{(\rho)}} \quad (82)$$

### 4.3 損失関数の微分

話者識別システムを DFE にて学習するためには、損失関数を周波数ワーピングパラメータ  $(\theta, \alpha)$  および各話者の GMM の各パラメータについて微分する必要がある。そこで、ここでは損失関数の微分の計算方法について述べる。

まず、損失関数のパラメータセットによる微分は、連鎖律 (chain rule) によって次式のように書き換えられる。

$$\nabla_{\Phi} \ell_k(\mathbf{X}; \Lambda) = \frac{\partial \ell_k(\mathbf{X}; \Lambda)}{\partial d_k(\mathbf{X}; \Lambda)} \sum_{s=1}^S \left\{ \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} \cdot \nabla_{\Phi} g_s(\mathbf{X}; \Lambda) \right\} \quad (83)$$

ここで、損失関数の誤分類測度に関する微分は、以下の通りである。

$$\begin{aligned} \frac{\partial \ell_k(\mathbf{X}; \Lambda)}{\partial d_k(\mathbf{X}; \Lambda)} &= \frac{\partial}{\partial d_k(\mathbf{X}; \Lambda)} \left[ \frac{1}{1 + \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) + \beta)\}} \right] \\ &= -\frac{-\gamma \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) + \beta)\}}{\left[1 + \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) + \beta)\}\right]^2} \\ &= \gamma \frac{1}{1 + \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) + \beta)\}} \left[ 1 - \frac{1}{1 + \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) + \beta)\}} \right] \\ &= \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \end{aligned} \quad (84)$$

誤分類測度の判別関数に関する微分は、式 (66) で表される一対多型の誤分類測度を用いた場合、

$$\frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} = \frac{\partial}{\partial g_s(\mathbf{X}; \Lambda)} \left[ -g_k(\mathbf{X}; \Lambda) + \log \left[ \frac{1}{S-1} \sum_{i \neq k} \exp\{g_i(\mathbf{X}; \Lambda)\eta\} \right]^{\frac{1}{\eta}} \right] \quad (85)$$

を計算することになるが、ここでは場合分けが必要となる。まず、 $s$  がデータ  $\mathbf{X}$  の持ち主  $k$  であるとき、つまり  $s = k$  のとき、

$$\frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} = \frac{\partial}{\partial g_k(\mathbf{X}; \Lambda)} [-g_k(\mathbf{X}; \Lambda)] = -1 \quad (86)$$

となる。次に  $s \neq k$  の場合、

$$\begin{aligned} \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} &= \frac{\partial}{\partial g_s(\mathbf{X}; \Lambda)} \left[ \log \left[ \frac{1}{S-1} \sum_{i \neq k} \exp\{g_i(\mathbf{X}; \Lambda)\eta\} \right]^{\frac{1}{\eta}} \right] \\ &= \frac{\partial}{\partial g_s(\mathbf{X}; \Lambda)} \left[ \frac{1}{\eta} \log \frac{1}{S-1} \sum_{i \neq k} \exp\{g_i(\mathbf{X}; \Lambda)\eta\} \right] \\ &= \frac{1}{\eta} \frac{S-1}{\sum_{i \neq k} \exp\{\eta g_i(\mathbf{X}; \Lambda)\}} \frac{\eta}{S-1} \exp\{\eta g_s(\mathbf{X}; \Lambda)\} \\ &= \frac{\exp\{\eta g_s(\mathbf{X}; \Lambda)\}}{\sum_{i \neq k} \exp\{\eta g_i(\mathbf{X}; \Lambda)\}} \end{aligned} \quad (87)$$

となる。一方、式 (69) で表される一対一型の誤分類測度を用いる場合は次式のようになる。

$$\begin{aligned} \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} &= \frac{\partial}{\partial g_s(\mathbf{X}; \Lambda)} [-g_k(\mathbf{X}; \Lambda) + g_j(\mathbf{X}; \Lambda)] \\ &= \begin{cases} -1, & s = k \\ 1, & s = j \\ 0, & s \neq k, j \end{cases} \end{aligned} \quad (88)$$

式 (83) においてまだ計算していない部分は、判別関数のパラメータセットに関する微分  $\nabla_{\Phi} g_s(\mathbf{X}; \Lambda)$  である。これは、さらに次のように書き下すことができる。

$$\begin{aligned}
 \nabla_{\Phi} g_s(\mathbf{X}; \Lambda) &= \nabla_{\Phi} \left[ \frac{1}{T} \log P(\mathbf{X} | \lambda_s) \right] \\
 &= \nabla_{\Phi} \left[ \frac{1}{T} \sum_{t=1}^T \log b(\mathbf{x}_t | \lambda_s) \right] \\
 &= \frac{1}{T} \sum_{t=1}^T \nabla_{\Phi} \left[ \log b(\mathbf{x}_t | \lambda_s) \right] \\
 &= \frac{1}{T} \sum_{t=1}^T \frac{1}{b(\mathbf{x}_t | \lambda_s)} \nabla_{\Phi} b(\mathbf{x}_t | \lambda_s)
 \end{aligned} \tag{89}$$

あとは GMM 出力確率のパラメータセットに関する微分  $\nabla_{\Phi} b(\mathbf{x} | \lambda)$  を計算すればよい。これは、特徴抽出パラメータ  $\Theta$  に関する微分、GMM パラメータ  $\Lambda$  に関する微分に分けられる。

$$\nabla_{\Phi} b(\mathbf{x}_t | \lambda_s) = \left[ \nabla_{\Theta} b(\mathbf{x}_t | \lambda_s) \quad \nabla_{\Lambda} b(\mathbf{x}_t | \lambda_s) \right]^T \tag{90}$$

#### 4.4 GMM 出力確率のワーピングパラメータに関する微分

ここでは、GMM 出力確率をワーピングパラメータに関して微分するための計算手順を述べる。本研究では、特徴抽出に 2 次オールパス関数に基づく周波ワーピングを利用したメルケプストラム推定法を用いるので、特徴抽出パラメータは強調中心周波数  $\theta$  と伸縮する度合  $\alpha$  の 2 つになる。そこで、次のように  $\theta, \alpha$  それぞれについて微分を行う。以下、簡単のため入力ベクトル番号  $t$  および話者番号  $s$  は省くことにする。

$$\nabla_{\Theta} b(\mathbf{x} | \lambda) = \left[ \frac{\partial b(\mathbf{x} | \lambda)}{\partial \theta} \quad \frac{\partial b(\mathbf{x} | \lambda)}{\partial \alpha} \right]^T \tag{91}$$

これらの微分はそれぞれ次式のように変形できる。

$$\frac{\partial b(\mathbf{x} | \lambda)}{\partial \theta} = \left[ \nabla_{\mathbf{x}} b(\mathbf{x} | \lambda) \right]^T \frac{\partial \mathbf{x}}{\partial \theta} \tag{92}$$

$$\frac{\partial b(\mathbf{x} | \lambda)}{\partial \alpha} = \left[ \nabla_{\mathbf{x}} b(\mathbf{x} | \lambda) \right]^T \frac{\partial \mathbf{x}}{\partial \alpha} \tag{93}$$

まず、出力確率の特徴ベクトルに関する微分を次のように各次元に分割して計算する。

$$\nabla_{\mathbf{x}} b(\mathbf{x} | \lambda) = \left[ \frac{\partial b(\mathbf{x} | \lambda)}{\partial x_1} \quad \frac{\partial b(\mathbf{x} | \lambda)}{\partial x_2} \quad \dots \quad \frac{\partial b(\mathbf{x} | \lambda)}{\partial x_D} \right]^T \tag{94}$$

出力確率の特徴ベクトルの  $d$  次目の値  $x_d$  に関する微分は、次のように計算できる。

$$\begin{aligned}
 \frac{\partial b(\mathbf{x} | \lambda)}{\partial x_d} &= \frac{\partial}{\partial x_d} \left[ \sum_{i=1}^I p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \right] \\
 &= \frac{\partial}{\partial x_d} \left[ \sum_{i=1}^I p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \right] \\
 &= \sum_{i=1}^I p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \frac{\partial}{\partial x_d} \left[ \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \right] \\
 &= \sum_{i=1}^I p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \frac{\partial}{\partial x_d} \left[ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right] \\
 &= \sum_{i=1}^I \left\{ -\frac{1}{2} p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \frac{\partial}{\partial x_d} \left[ \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right] \right\}
 \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^I \left\{ -\frac{1}{2\sigma_{id}^2} p_i \mathcal{N}(x | \mu_i, \Sigma_i) \frac{\partial}{\partial x_d} \left[ (x_d - \mu_{id})^2 \right] \right\} \\
&= \sum_{i=1}^I \left\{ -\frac{1}{2\sigma_{id}^2} p_i \mathcal{N}(x | \mu_i, \Sigma_i) \cdot 2(x_d - \mu_{id}) \right\} \\
&= -\sum_{i=1}^I \frac{x_d - \mu_{id}}{\sigma_{id}^2} p_i \mathcal{N}(x | \mu_i, \Sigma_i) \tag{95}
\end{aligned}$$

次に、特徴ベクトルの周波数ワーピングパラメータに関する微分  $\partial x / \partial \theta$ ,  $\partial x / \partial \alpha$  の計算を行う。ここで、特徴ベクトルはメルケプストラム係数であるので、以下、 $x$  に代わって  $c$  と表記する。メルケプストラム係数は、スペクトル評価関数が最小となるように推定されたもので、単純に  $\theta, \alpha$  の関数として表すことはできない。そこで、陰関数の定理を用いて次のように変形する。

$$\frac{\partial x}{\partial \theta} = \frac{\partial c}{\partial \theta} = -[\nabla_c^2 E]^{-1} \cdot \frac{\partial}{\partial \theta} [\nabla_c E] = -\mathbf{H}^{-1} \frac{\partial}{\partial \theta} [\nabla_c E] \tag{96}$$

$$\frac{\partial x}{\partial \alpha} = \frac{\partial c}{\partial \alpha} = -[\nabla_c^2 E]^{-1} \cdot \frac{\partial}{\partial \alpha} [\nabla_c E] = -\mathbf{H}^{-1} \frac{\partial}{\partial \alpha} [\nabla_c E] \tag{97}$$

さらに、式 (96), (97) は、Toeplitz + Hankel 行列を用いた線形方程式として次の形に書き換えられる。

$$\mathbf{H} \frac{\partial c}{\partial \kappa} = -\frac{\partial}{\partial \kappa} [\nabla_c E], \quad \kappa = \theta | \alpha \tag{98}$$

スペクトル評価関数  $E$  の  $c$  に関する 1 階微分および 2 階微分については、メルケプストラムを推定する際に式 (46), (52) にて計算済みであるので、あとは 1 階微分の  $\theta, \alpha$  に関する微分を行い、メルケプストラム推定法で用いた方法と同じ方法で線形方程式を解けば、その解として  $\partial c / \partial \kappa$  が得られる。1 階微分の  $\kappa$  ( $\theta$  もしくは  $\alpha$ ) に関する微分は、式 (45) を代入し、さらに連鎖律を適用して次のように書き換えられる。

$$\begin{aligned}
\frac{\partial}{\partial \kappa} [\nabla_c E] &= \frac{\partial}{\partial \kappa} \left[ -\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ I_N(\omega) \exp \{ -c^T (z + z^*) \} - 1 \right\} (z + z^*) d\omega \right] \\
&= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial}{\partial \kappa} \left[ \left\{ I_N(\omega) \exp \{ -c^T (z + z^*) \} - 1 \right\} (z + z^*) \right] d\omega \\
&= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \nabla_{\psi} \left[ \left\{ I_N(\omega) \exp (-c^T \psi) - 1 \right\} \psi \right] \cdot \frac{\partial \psi}{\partial \tilde{\omega}} \cdot \frac{\partial \tilde{\omega}}{\partial \kappa} d\omega \\
&= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \psi \left\{ I_N(\omega) \exp (-c^T \psi) \cdot (-c^T) \right\} + \left\{ I_N(\omega) \exp (-c^T \psi) - 1 \right\} \mathbf{I} \right\} \frac{\partial \psi}{\partial \tilde{\omega}} \cdot \frac{\partial \tilde{\omega}}{\partial \kappa} d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \psi c^T - \left( \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right) \mathbf{I} \right\} \frac{\partial \psi}{\partial \tilde{\omega}} \cdot \frac{\partial \tilde{\omega}}{\partial \kappa} d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} c^T \frac{\partial \psi}{\partial \tilde{\omega}} \cdot \frac{\partial \tilde{\omega}}{\partial \kappa} \psi d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} \frac{\partial \psi}{\partial \tilde{\omega}} \cdot \frac{\partial \tilde{\omega}}{\partial \kappa} d\omega \tag{99}
\end{aligned}$$

ただし、 $\mathbf{I}$  は  $(M+1) \times (M+1)$  の単位行列、 $\psi$  は次式のように定義する。

$$\begin{aligned}
\psi &= z + z^* \\
&= [\psi_0 \ \psi_1 \ \cdots \ \psi_M]^T \\
&= [2 e^{-j\tilde{\omega}} + e^{j\tilde{\omega}} \ e^{-j\tilde{\omega} \cdot 2} + e^{j\tilde{\omega} \cdot 2} \ \cdots \ e^{-j\tilde{\omega} M} + e^{j\tilde{\omega} M}]^T \tag{100}
\end{aligned}$$

$\psi$  の位相特性  $\tilde{\omega}$  に関する微分は次式となる。

$$\begin{aligned}
\frac{\partial \psi_m}{\partial \tilde{\omega}} &= \frac{\partial}{\partial \tilde{\omega}} \left[ e^{-j\tilde{\omega} m} + e^{j\tilde{\omega} m} \right] \\
&= j m \left( -e^{-j\tilde{\omega} m} + e^{j\tilde{\omega} m} \right) \\
&= -2 m \sin(m\tilde{\omega}) \tag{101}
\end{aligned}$$



そこで、式 (99) を成分ごとに記述すると、次式のようになる。

$$\begin{aligned}
\frac{\partial^2 E}{\partial c(m) \partial \kappa} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \mathbf{c}^T \frac{\partial \psi}{\partial \tilde{\omega}} \psi_m d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} \frac{\partial \psi_m}{\partial \tilde{\omega}} d\omega \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \{-2n \sin(n\tilde{\omega})\} \right\} \psi_m d\omega \\
&\quad - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} j m (-e^{-j\tilde{\omega}m} + e^{j\tilde{\omega}m}) d\omega \\
&= -2 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \cdot n \sin(n\tilde{\omega}) \right\} (e^{-j\tilde{\omega}m} + e^{j\tilde{\omega}m}) d\omega \\
&\quad - m \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} j \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} (-e^{-j\tilde{\omega}m} + e^{j\tilde{\omega}m}) d\omega \\
&= -4 \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \cdot n \sin(n\tilde{\omega}) \right\} e^{j\tilde{\omega}m} d\omega \\
&\quad - 2m \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} j \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} e^{j\tilde{\omega}m} d\omega \tag{102}
\end{aligned}$$

ここで、関数  $\tilde{\omega}(\omega, \theta, \alpha)$  の  $\theta, \alpha$  に関する微分は、それぞれ式 (104), (105) となる。ただし、これらの微分を計算するためには、微分公式

$$\frac{d}{dx} [\tan^{-1} x] = \frac{1}{1+x^2} \tag{103}$$

を用いる。

$$\begin{aligned}
\frac{\partial \tilde{\omega}(\omega, \theta, \alpha)}{\partial \theta} &= \frac{\partial}{\partial \theta} \left[ \omega + \tan^{-1} \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right) + \tan^{-1} \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right) \right] \\
&= \left\{ 1 + \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right)^2 \right\}^{-1} \left\{ \frac{-\alpha \cos(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} - \frac{\alpha \sin(\omega - \theta) \cdot \{-\alpha \sin(\omega - \theta)\}}{\{1 - \alpha \cos(\omega - \theta)\}^2} \right\} \\
&\quad + \left\{ 1 + \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right)^2 \right\}^{-1} \left\{ \frac{\alpha \cos(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} - \frac{\alpha \sin(\omega + \theta) \cdot \alpha \sin(\omega + \theta)}{\{1 - \alpha \cos(\omega + \theta)\}^2} \right\} \\
&= \frac{\{1 - \alpha \cos(\omega - \theta)\}^2}{1 + \alpha^2 - 2\alpha \cos(\omega - \theta)} \frac{\alpha^2 - \alpha \cos(\omega - \theta)}{\{1 - \alpha \cos(\omega - \theta)\}^2} + \frac{\{1 - \alpha \cos(\omega + \theta)\}^2}{1 + \alpha^2 - 2\alpha \cos(\omega + \theta)} \frac{-\alpha^2 + \alpha \cos(\omega + \theta)}{\{1 - \alpha \cos(\omega + \theta)\}^2} \\
&= \frac{\alpha^2 - \alpha \cos(\omega - \theta)}{1 + \alpha^2 - 2\alpha \cos(\omega - \theta)} - \frac{\alpha^2 - \alpha \cos(\omega + \theta)}{1 + \alpha^2 - 2\alpha \cos(\omega + \theta)} \tag{104}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \tilde{\omega}(\omega, \theta, \alpha)}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \left[ \omega + \tan^{-1} \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right) + \tan^{-1} \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right) \right] \\
&= \left\{ 1 + \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right)^2 \right\}^{-1} \left\{ \frac{\sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} - \frac{\alpha \sin(\omega - \theta) \cdot \{-\cos(\omega - \theta)\}}{\{1 - \alpha \cos(\omega - \theta)\}^2} \right\} \\
&\quad + \left\{ 1 + \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right)^2 \right\}^{-1} \left\{ \frac{\sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} - \frac{\alpha \sin(\omega + \theta) \cdot \{-\cos(\omega + \theta)\}}{\{1 - \alpha \cos(\omega + \theta)\}^2} \right\} \\
&= \frac{\{1 - \alpha \cos(\omega - \theta)\}^2}{1 + \alpha^2 - 2\alpha \cos(\omega - \theta)} \frac{\sin(\omega - \theta)}{\{1 - \alpha \cos(\omega - \theta)\}^2} + \frac{\{1 - \alpha \cos(\omega + \theta)\}^2}{1 + \alpha^2 - 2\alpha \cos(\omega + \theta)} \frac{\sin(\omega + \theta)}{\{1 - \alpha \cos(\omega + \theta)\}^2} \\
&= \frac{\sin(\omega - \theta)}{1 + \alpha^2 - 2\alpha \cos(\omega - \theta)} + \frac{\sin(\omega + \theta)}{1 + \alpha^2 - 2\alpha \cos(\omega + \theta)} \tag{105}
\end{aligned}$$

ここで,  $\Re(m)$ ,  $\Im(m)$  を

$$\Re(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \cdot n \sin(n\tilde{\omega}) \right\} e^{j\tilde{\omega}m} d\omega \quad (106)$$

$$\Im(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} j \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} e^{j\tilde{\omega}m} d\omega \quad (107)$$

とおくと, 式 (102) は次式になる.

$$\frac{\partial^2 E}{\partial c(m) \partial \kappa} = -4 \cdot \Re(m) - 2m \cdot \Im(m) \quad (108)$$

$\Re(m)$ ,  $\Im(m)$  を計算するためには, まず, 以下のような逆フーリエ変換  $\Re'(m)$ ,  $\Im'(m)$  を計算する.

$$\begin{aligned} \Re'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \cdot n \sin(n\tilde{\omega}) \right\} e^{j\tilde{\omega}m} d\omega \\ &\approx \sum_{n=0}^{N-1} \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \cdot n \sin(n\tilde{\omega}) \right\} e^{j\frac{2\pi}{N}km} \\ &= \mathcal{F}_{(m)}^{-1} \left[ \frac{\partial \tilde{\omega}}{\partial \kappa} \frac{I_N(\omega)}{|H(e^{j\omega})|^2} \left\{ \sum_{n=0}^M c(n) \cdot n \sin(n\tilde{\omega}) \right\} \right] \end{aligned} \quad (109)$$

$$\begin{aligned} \Im'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} j \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} e^{j\tilde{\omega}m} d\omega \\ &\approx \sum_{n=0}^{N-1} j \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} e^{j\frac{2\pi}{N}km} \\ &= \mathcal{F}_{(m)}^{-1} \left[ j \frac{\partial \tilde{\omega}}{\partial \kappa} \left\{ \frac{I_N(\omega)}{|H(e^{j\omega})|^2} - 1 \right\} \right] \end{aligned} \quad (110)$$

ここで,  $\Re'(m)$  は, 奇×偶×奇, つまり実偶関数の逆フーリエ変換である. 一方,  $\Im'(m)$  は,  $j$  (奇×偶), つまり純虚奇関数の逆フーリエ変換となっている. これらの逆フーリエ変換の結果はともに実数となり, それぞれ付録 A に示す実偶関数の逆フーリエ変換値のための変換行列および純虚奇関数の逆フーリエ変換値のための変換行列を用いて  $\Re(m)$ ,  $\Im(m)$  に変換することができる.

以上に述べた方法で, 特徴抽出パラメータに関する微分計算を行うことができるが, 実際にパラメータ  $\theta$ ,  $\alpha$  を更新することを考えると, これらのパラメータのとり得る値の範囲が問題となる. つまり,

$$0 \leq \theta \leq \pi \quad (111)$$

$$-1 < \alpha < 1 \quad (112)$$

を常に満たさなければならない. そこで, 式 (113), (114) の左側に示すような  $\bar{\theta}$ ,  $\bar{\alpha}$  という新しいパラメータのスケールにおいて更新を行うことにする. また, 元のスケールに戻す際は右側の式を用いる. ただし,  $\theta$  に関しては, 0 から  $\pi$  の間の値をとるシグモイド関数で変換すると,  $\theta$  の値が 0 をとり得なくなるため,  $\theta$  については,  $-\pi$  から  $\pi$  のシグモイド関数を用いて有効範囲を広げている.  $\theta$  の値が負になった場合, ワーピング関数は  $\theta$  に関して偶関数であるので, 正の値に変換すればよい. また,  $\theta$  は  $\pi$  を取り得ないが, 関係式 (28) より,  $\theta = 0$  のとき  $\alpha$  の符号を逆転すれば, 同じワーピングが得られる. 他の解決策として,  $\theta$  が  $\pi$  を越えた場合に,  $-\pi$  に戻して循環させる方法も考えられるが, 今回は, シグモイド関数で値を変換する方法を用いた.

$$\bar{\theta} = \log \frac{\pi + \theta}{\pi - \theta} \iff \theta = \pi \frac{e^{\bar{\theta}} - 1}{e^{\bar{\theta}} + 1} \quad (113)$$

$$\bar{\alpha} = \log \frac{1 + \alpha}{1 - \alpha} \iff \alpha = \frac{e^{\bar{\alpha}} - 1}{e^{\bar{\alpha}} + 1} \quad (114)$$

ここで, 元のパラメータ  $\theta$  の新しいパラメータ  $\bar{\theta}$  に関する微分は次式となる.

$$\frac{\partial \theta}{\partial \bar{\theta}} = \frac{\partial}{\partial \bar{\theta}} \left[ \pi \frac{e^{\bar{\theta}} - 1}{e^{\bar{\theta}} + 1} \right]$$

$$\begin{aligned}
&= \frac{\partial}{\partial \bar{\theta}} \left[ \pi \left( 1 - \frac{2}{e^{\bar{\theta}} + 1} \right) \right] \\
&= \pi \frac{2e^{\bar{\theta}}}{(e^{\bar{\theta}} + 1)^2} \\
&= \frac{\pi}{2} \frac{4e^{\bar{\theta}}}{(e^{\bar{\theta}} + 1)^2} \\
&= \frac{\pi}{2} \frac{(e^{\bar{\theta}} + 1)^2 - (e^{\bar{\theta}} - 1)^2}{(e^{\bar{\theta}} + 1)^2} \\
&= \frac{1}{2\pi} \left\{ \pi^2 - \left( \pi \frac{e^{\bar{\theta}} - 1}{e^{\bar{\theta}} + 1} \right)^2 \right\} \\
&= \frac{\pi^2 - \theta^2}{2\pi} \tag{115}
\end{aligned}$$

また、元のパラメータ  $\alpha$  の新しいパラメータ  $\bar{\alpha}$  に関する微分は次式となる。

$$\begin{aligned}
\frac{\partial \alpha}{\partial \bar{\alpha}} &= \frac{\partial}{\partial \bar{\alpha}} \left[ \frac{e^{\bar{\alpha}} - 1}{e^{\bar{\alpha}} + 1} \right] \\
&= \frac{\partial}{\partial \bar{\alpha}} \left[ 1 - \frac{2}{e^{\bar{\alpha}} + 1} \right] \\
&= \frac{2e^{\bar{\alpha}}}{(e^{\bar{\alpha}} + 1)^2} \\
&= \frac{1}{2} \frac{4e^{\bar{\alpha}}}{(e^{\bar{\alpha}} + 1)^2} \\
&= \frac{1}{2} \frac{(e^{\bar{\alpha}} + 1)^2 - (e^{\bar{\alpha}} - 1)^2}{(e^{\bar{\alpha}} + 1)^2} \\
&= \frac{1}{2} \left\{ 1 - \left( \frac{e^{\bar{\alpha}} - 1}{e^{\bar{\alpha}} + 1} \right)^2 \right\} \\
&= \frac{1 - \alpha^2}{2} \tag{116}
\end{aligned}$$

したがって、各特徴パラメータの新しいパラメータに関する微分は、連鎖律より以下の式で与えられる。

$$\frac{\partial \mathbf{c}}{\partial \bar{\theta}} = \frac{\partial \mathbf{c}}{\partial \theta} \frac{\partial \theta}{\partial \bar{\theta}} = \frac{\pi^2 - \theta^2}{2\pi} \frac{\partial \mathbf{c}}{\partial \theta} \tag{117}$$

$$\frac{\partial \mathbf{c}}{\partial \bar{\alpha}} = \frac{\partial \mathbf{c}}{\partial \alpha} \frac{\partial \alpha}{\partial \bar{\alpha}} = \frac{1 - \alpha^2}{2} \frac{\partial \mathbf{c}}{\partial \alpha} \tag{118}$$

以上をまとめると、各繰り返しにおける  $\bar{\theta}$ ,  $\bar{\alpha}$  の更新量 (一対一型の誤分類測度の場合) は次のようになる。

$$\begin{aligned}
\Delta \bar{\theta}^{(\rho)} &= \epsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \\
&\quad \cdot \sum_{s=1}^S \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} \frac{1}{T} \sum_{t=1}^T \frac{1}{b_s(\mathbf{c}_t)} \sum_{m=0}^M \frac{\partial c_t(m)}{\partial \theta^{(\rho)}} \frac{\partial \theta^{(\rho)}}{\partial \bar{\theta}^{(\rho)}} \left\{ - \sum_{i=1}^I \frac{c_t(m) - \mu_{sim}}{\sigma_{sim}^2} p_{si} \mathcal{N}(\mathbf{c}_t | \boldsymbol{\mu}_{si}, \boldsymbol{\Sigma}_{si}) \right\} \\
&= \epsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \frac{\pi^2 - \theta^{2(\rho)}}{2\pi} \\
&\quad \cdot \frac{1}{T} \sum_{t=1}^T \left\{ \frac{1}{b_k(\mathbf{c}_t)} \sum_{m=0}^M \frac{\partial c_t(m)}{\partial \theta^{(\rho)}} \sum_{i=1}^I \frac{c_t(m) - \mu_{kim}}{\sigma_{kim}^2} p_{ki} \mathcal{N}(\mathbf{c}_t | \boldsymbol{\mu}_{ki}, \boldsymbol{\Sigma}_{ki}) \right. \\
&\quad \left. - \frac{1}{b_j(\mathbf{c}_t)} \sum_{m=0}^M \frac{\partial c_t(m)}{\partial \theta^{(\rho)}} \sum_{i=1}^I \frac{c_t(m) - \mu_{jim}}{\sigma_{jim}^2} p_{ji} \mathcal{N}(\mathbf{c}_t | \boldsymbol{\mu}_{ji}, \boldsymbol{\Sigma}_{ji}) \right\} \tag{119}
\end{aligned}$$

$$\begin{aligned}
\Delta \bar{\alpha}^{(\rho)} &= \epsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \\
&\quad \cdot \sum_{s=1}^S \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} \frac{1}{T} \sum_{t=1}^T \frac{1}{b_s(\mathbf{c}_t)} \sum_{m=0}^M \frac{\partial c_t(m)}{\partial \alpha^{(\rho)}} \frac{\partial \alpha^{(\rho)}}{\partial \bar{\alpha}^{(\rho)}} \left\{ - \sum_{i=1}^I \frac{c_t(m) - \mu_{sim}}{\sigma_{sim}^2} p_{si} \mathcal{N}(\mathbf{c}_t | \boldsymbol{\mu}_{si}, \boldsymbol{\Sigma}_{si}) \right\}
\end{aligned}$$

$$\begin{aligned}
&= \epsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \frac{1 - \alpha^{2(\rho)}}{2} \\
&\quad \cdot \frac{1}{T} \sum_{t=1}^T \left\{ \frac{1}{b_k(\mathbf{c}_t)} \sum_{m=0}^M \frac{\partial c_t(m)}{\partial \alpha^{(\rho)}} \sum_{i=1}^I \frac{c_t(m) - \mu_{kim}}{\sigma_{kim}^2} p_{ki} \mathcal{N}(\mathbf{c}_t | \boldsymbol{\mu}_{ki}, \boldsymbol{\Sigma}_{ki}) \right. \\
&\quad \left. - \frac{1}{b_j(\mathbf{c}_t)} \sum_{m=0}^M \frac{\partial c_t(m)}{\partial \alpha^{(\rho)}} \sum_{i=1}^I \frac{c_t(m) - \mu_{jim}}{\sigma_{jim}^2} p_{ji} \mathcal{N}(\mathbf{c}_t | \boldsymbol{\mu}_{ji}, \boldsymbol{\Sigma}_{ji}) \right\} \quad (120)
\end{aligned}$$

最終的に、次式を用いて更新を行う。

$$\bar{\theta}^{(\rho+1)} = \bar{\theta}^{(\rho)} + \Delta \bar{\theta}^{(\rho)} \quad (121)$$

$$\bar{\alpha}^{(\rho+1)} = \bar{\alpha}^{(\rho)} + \Delta \bar{\alpha}^{(\rho)} \quad (122)$$

#### 4.5 GMM 出力確率の GMM パラメータに関する微分

GMM 出力確率の GMM のパラメータセットに関する微分は、次のように話者ごとに分割して考える。

$$\nabla_{\Lambda} b(\mathbf{x}_t | \lambda_s) \rightarrow \nabla_{\lambda_s} b(\mathbf{x}_t | \lambda_s), \quad s = 1, 2, \dots, S \quad (123)$$

以下、簡単のため入力ベクトル番号  $t$  および話者番号  $s$  は省く。さらに、モデルパラメータに関する微分は、重み、平均、分散それぞれのスカラー値に関する微分に分けられる。

$$\begin{aligned}
\nabla_{\lambda} b(\mathbf{x}_t | \lambda) &\rightarrow \frac{\partial b(\mathbf{x} | \lambda)}{\partial p_i}, \frac{\partial b(\mathbf{x} | \lambda)}{\partial \mu_{id}}, \frac{\partial b(\mathbf{x} | \lambda)}{\partial \sigma_{id}} \\
&\quad i = 1, 2, \dots, I, \quad d = 1, 2, \dots, D \quad (124)
\end{aligned}$$

混合重み  $p_i$  に関する微分は次のようになる。

$$\begin{aligned}
\frac{\partial b(\mathbf{x} | \lambda)}{\partial p_i} &= \frac{\partial}{\partial p_i} \left[ \sum_{j=1}^I p_j \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \right] \\
&= \frac{\partial}{\partial p_i} \left[ p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \right] \\
&= \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (125)
\end{aligned}$$

平均ベクトル要素  $\mu_{id}$  に関する微分は次のように計算できる。

$$\begin{aligned}
\frac{\partial b(\mathbf{x} | \lambda)}{\partial \mu_{id}} &= \frac{\partial}{\partial \mu_{id}} \left[ \sum_{j=1}^I p_j \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \right] \\
&= \frac{\partial}{\partial \mu_{id}} \left[ p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \right] \\
&= \frac{\partial}{\partial \mu_{id}} \left[ p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \right] \\
&= p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \frac{\partial}{\partial \mu_{id}} \left[ \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \right] \\
&= p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \frac{\partial}{\partial \mu_{id}} \left[ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right] \\
&= -\frac{1}{2} p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \frac{\partial}{\partial \mu_{id}} \left[ \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right] \\
&= -\frac{1}{2\sigma_{id}^2} p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \frac{\partial}{\partial \mu_{id}} \left[ (x_d - \mu_{id})^2 \right]
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{2\sigma_{id}^2} p_i \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \left\{ -2(x_d - \mu_{id}) \right\} \\
&= \frac{x_d - \mu_{id}}{\sigma_{id}^2} p_i \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)
\end{aligned} \tag{126}$$

対角共分散の対角要素  $\sigma_{id}^2$  に関する微分は次式となる.

$$\begin{aligned}
\frac{\partial b(\mathbf{x} \mid \lambda)}{\partial \sigma_{id}^2} &= \left[ \frac{\partial}{\partial \sigma_{id}} [\sigma_{id}^2] \right]^{-1} \frac{\partial b(\mathbf{x} \mid \lambda)}{\partial \sigma_{id}} \\
&= \frac{1}{2\sigma_{id}} \frac{\partial}{\partial \sigma_{id}} \left[ \sum_{j=1}^I p_j \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) \right] \\
&= \frac{1}{2\sigma_{id}} \frac{\partial}{\partial \sigma_{id}} \left[ p_i \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \right] \\
&= \frac{1}{2\sigma_{id}} \frac{\partial}{\partial \sigma_{id}} \left[ p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \right] \\
&= \frac{1}{2\sigma_{id}} p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l \neq d} \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l \neq d} \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \\
&\quad \cdot \frac{\partial}{\partial \sigma_{id}} \left[ \frac{1}{\sigma_{id}} \exp \left\{ -\frac{1}{2} \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right\} \right] \\
&= \frac{1}{2\sigma_{id}} p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l \neq d} \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l \neq d} \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \\
&\quad \cdot \left[ -\frac{1}{\sigma_{id}^2} \exp \left\{ -\frac{1}{2} \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right\} \right. \\
&\quad \quad \left. + \frac{1}{\sigma_{id}} \exp \left\{ -\frac{1}{2} \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right\} \left( -\frac{1}{2} \right) \cdot (-2) \frac{(x_d - \mu_{id})^2}{\sigma_{id}^3} \right] \\
&= \frac{1}{2\sigma_{id}} p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l \neq d} \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l \neq d} \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \\
&\quad \cdot \frac{1}{\sigma_{id}^2} \exp \left\{ -\frac{1}{2} \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right\} \cdot \left\{ -1 + \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 \right\} \\
&= \frac{1}{2\sigma_{id}^2} p_i \frac{1}{\sqrt{(2\pi)^D} \prod_{l=1}^D \sigma_{il}} \exp \left\{ -\frac{1}{2} \sum_{l=1}^D \left( \frac{x_l - \mu_{il}}{\sigma_{il}} \right)^2 \right\} \cdot \left\{ \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 - 1 \right\} \\
&= \frac{1}{2\sigma_{id}^2} \left\{ \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 - 1 \right\} p_i \mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)
\end{aligned} \tag{127}$$

ただし、パラメータを更新する際、混合重みおよび分散の値は常に正の値であるという制約が必要である。そこで、これらの値は対数領域において更新する。また、今回は学習行列を省略するため、平均ベクトルについては標準偏差で正規化した値において更新することにする。そこで、各パラメータを次の左側の式で示すような  $\bar{p}_i, \bar{\mu}_{id}, \bar{\sigma}_{id}$  という新しいパラメータのスケールにおいて更新を行う。逆に、変換後は次に示す右側の式を用いて元のスケールに戻す。

$$\bar{p}_i = \log p_i \iff p_i = \exp \bar{p}_i \tag{128}$$

$$\bar{\mu}_{id} = \frac{\mu_{id}}{\sigma_{id}} \iff \mu_{id} = \sigma_{id} \bar{\mu}_{id} \tag{129}$$

$$\bar{\sigma}_{id}^2 = \log \sigma_{id}^2 \iff \sigma_{id}^2 = \exp \bar{\sigma}_{id}^2 \tag{130}$$

ここで、元の各パラメータの新しいスケールのパラメータに関する微分は、

$$\frac{\partial p_i}{\partial \bar{p}_i} = \frac{\partial}{\partial \bar{p}_i} [\exp \bar{p}_i] = \exp \bar{p}_i = p_i \tag{131}$$

$$\frac{\partial \mu_{id}}{\partial \bar{\mu}_{id}} = \frac{\partial}{\partial \bar{\mu}_{id}} [\sigma_{id} \bar{\mu}_{id}] = \sigma_{id} \quad (132)$$

$$\frac{\partial \sigma_{id}^2}{\partial \bar{\sigma}_{id}^2} = \frac{\partial}{\partial \bar{\sigma}_{id}^2} [\exp \bar{\sigma}_{id}^2] = \exp \bar{\sigma}_{id}^2 = \sigma_{id}^2 \quad (133)$$

となり、これらと式 (125), (126), (127) から、連鎖律により、GMM の出力確率の新しいパラメータに関する微分は次のようになる。

$$\frac{\partial b(\mathbf{x} | \lambda)}{\partial \bar{p}_i} = \frac{\partial b(\mathbf{x} | \lambda)}{\partial p_i} \frac{\partial p_i}{\partial \bar{p}_i} = p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (134)$$

$$\frac{\partial b(\mathbf{x} | \lambda)}{\partial \bar{\mu}_{id}} = \frac{\partial b(\mathbf{x} | \lambda)}{\partial \mu_{id}} \frac{\partial \mu_{id}}{\partial \bar{\mu}_{id}} = \frac{x_d - \mu_{id}}{\sigma_{id}} p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (135)$$

$$\frac{\partial b(\mathbf{x} | \lambda)}{\partial \bar{\sigma}_{id}^2} = \frac{\partial b(\mathbf{x} | \lambda)}{\partial \sigma_{id}^2} \frac{\partial \sigma_{id}^2}{\partial \bar{\sigma}_{id}^2} = \frac{1}{2} \left\{ \left( \frac{x_d - \mu_{id}}{\sigma_{id}} \right)^2 - 1 \right\} p_i \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \quad (136)$$

以上をまとめると、各繰り返しにおける、話者の  $s$  の  $\bar{p}_{si}$ ,  $\bar{\mu}_{sid}$ ,  $\bar{\sigma}_{sid}^2$  それぞれの更新量は次のようになる。ただし、話者  $k$  の学習データ  $\mathbf{X}$  が与えられた場合とする。

$$\Delta \bar{p}_{si}^{(\rho)} = \varepsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} \frac{1}{T} \sum_{t=1}^T \frac{1}{b(\mathbf{x}_t | \lambda_s)} p_{si} \mathcal{N}(\mathbf{x}_t | \boldsymbol{\mu}_{si}, \boldsymbol{\Sigma}_{si}) \quad (137)$$

$$\Delta \bar{\mu}_{sid}^{(\rho)} = \varepsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} \frac{1}{T} \sum_{t=1}^T \frac{1}{b(\mathbf{x}_t | \lambda_s)} \frac{x_{td} - \mu_{sid}}{\sigma_{sid}} p_{si} \mathcal{N}(\mathbf{x}_t | \boldsymbol{\mu}_{si}, \boldsymbol{\Sigma}_{si}) \quad (138)$$

$$\Delta \bar{\sigma}_{sid}^{2(\rho)} = \varepsilon^{(\rho)} \gamma \ell_k(\mathbf{X}; \Lambda) \{1 - \ell_k(\mathbf{X}; \Lambda)\} \frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} \frac{1}{T} \sum_{t=1}^T \frac{1}{b(\mathbf{x}_t | \lambda_s)} \frac{1}{2} \left\{ \left( \frac{x_{td} - \mu_{sid}}{\sigma_{sid}} \right)^2 - 1 \right\} p_{si} \mathcal{N}(\mathbf{x}_t | \boldsymbol{\mu}_{si}, \boldsymbol{\Sigma}_{si}) \quad (139)$$

ここで、今回は一対一型の誤分類測度を用いるので、誤分類測度の判別関数に関する微分は、

$$\frac{\partial d_k(\mathbf{X}; \Lambda)}{\partial g_s(\mathbf{X}; \Lambda)} = \begin{cases} -1, & s = k \\ 1, & s = j \\ 0, & s \neq k, j \end{cases} \quad (140)$$

となる。つまり、本人  $k$  の各モデルパラメータおよび本人以外で一番尤度の高い話者  $j$  の各モデルパラメータを次の更新式によって逆方向に更新し、他のモデルについては更新しなくてよい。

$$\bar{p}_{si}^{(\rho+1)} = \bar{p}_{si}^{(\rho)} - \Delta \bar{p}_{si}^{(\rho)} \quad (141)$$

$$\bar{\mu}_{sid}^{(\rho+1)} = \bar{\mu}_{sid}^{(\rho)} - \Delta \bar{\mu}_{sid}^{(\rho)} \quad (142)$$

$$\bar{\sigma}_{sid}^{2(\rho+1)} = \bar{\sigma}_{sid}^{2(\rho)} - \Delta \bar{\sigma}_{sid}^{2(\rho)} \quad (143)$$

なお、混合重みは常に

$$\sum_{i=1}^I p_i = 1 \quad (144)$$

を満たさなければならない。正確にはこの制約を GPD 学習に組み込むのが望ましいが、今回は簡単のため、次式を用いて、更新都度和が 1 になるように調整している。

$$p_i = \frac{\exp \bar{p}_i}{\sum_{j=1}^I \exp \bar{p}_j} \quad (145)$$

## 5 実験

### 5.1 識別タスクの設定

発声時期差が考慮された日本語による話者認識用のデータベースはほとんど作られていないのが現状である。そこで、まず、使用可能なデータベースを用いて、識別タスク設定のための予備実験を行う。今回は、NTT VR データベース (発声時期差あり) および ATR データベースの C set (時期差なし) を用いて予備実験を行う。予備実験では、メルスケールにおけるメルケプストラム推定により得られたケプストラム係数を特徴量として用い、EM アルゴリズムで ML 学習した GMM によりどの程度の認識率が得られるかを調べる。

このときの音声の分析条件を表 2 に、GMM の学習条件を表 3 に示す。ここで、話者の特徴は高域に現れるという説があるため、高い周波数もパラメータ  $\theta$  を用いて強調することができるように、サンプリング周波数は 16kHz と高めに設定している。また、NTT データベースおよび ATR データベースを用いた場合の実験タスクをそれぞれ表 4, 5 に示す。なお、NTT データベースのタスクにおいては、文章を無音区間を除く 0.5 秒前後の単語長程度に分割して用いている。

これらの実験条件において、16, 32, 64 混合の GMM を用いた場合に得られた結果をそれぞれ表 6, 7 に示す。表のように、ATR データベースでは、発声時期差がないこともあり、小さな GMM においてもかなり低い識別誤り率が得られており、提案手法により改善する余地がほとんどないと考えられる。NTT データベースに関しても低い誤り率となっているが、ATR データベースと比較すると誤り率は比較的高く、特にサイズの小さな GMM の場合については、まだ改善の余地は残っていると考えられるため、今回は NTT データベースを用いて実験を行うこととする。

表 2. Analysis conditions for preliminary experiments.

Sampling frequency	16kHz
Window function	Blackman window
Frame length	30ms (480pt)
Frame period	8ms (128pt)
Analysis	18-th order mel-cepstral analysis
Warping parameters	$(\theta, \alpha) = (0, 0.42)$
Feature parameter	$c(1) - c(18)$

表 3. GMM training conditions.

GMM	GMM with diagonal covariance matrix
Number of Gaussians	16, 32, 64
Initialization	LBG codebook
Minimum variance	0.001
Minimum training iteration	50 times
Maximum training iteration	100 times

表4. Speaker identification task (NTT).

Database	NTT VR database
Speakers	13 females
Training data	90.8, 90.9, 90.12, 91.3 5 sentences (B set) / session
Test data	91.6 15 sentences (A + B set)
Data length	0.48–0.6 second

表5. Speaker identification task (ATR).

Database	ATR database
Speakers	20 females (f101–f120)
Training data	Phonetically balanced words (216 words)
Test data	Common words (520 words)
Data length	word length

表6. NTT database results in error rate (%).

Mixture	16	32	64
Training data	0.95 (12/1264)	0.24 (3/1264)	0.079 (1/1264)
Test data	8.42 (81/962)	7.17 (69/962)	7.80 (75/962)

表7. ATR database results in error rate (%).

Mixture	16	32	64
Training data	0.21 (9/4320)	0.093 (4/4320)	0.023 (1/4320)
Test data	4.41 (459/10400)	3.58 (372/10400)	3.05 (317/10400)



## 5.2 ML 学習による GMM を用いた識別実験

ここでは、2次オールパス関数に基づくメルケプストラムを用いることにより、どの程度識別率に変化があるかを、ML 学習による GMM を用いて調べる。この実験における諸条件は予備実験とほぼ同じで、分析条件、GMM の学習条件、識別タスクは、それぞれ表 8、表 9、表 10 に示す通りである。ただし、 $\theta$  については 0 から  $\pi$  まで  $0.05\pi \sim 0.1\pi$  刻みに、 $\alpha$  については 0 から 1 まで 0.05 刻みに変化させ、得られたケプストラム係数を用いて GMM を ML 学習にて作成し識別を行う。また、女性 13 名による識別実験に加えて、男性 22 名の実験、男女 35 名全員による実験も行う。

表 8. Analysis conditions.

Sampling frequency	16kHz
Window function	Blackman window
Frame length	30ms (480pt)
Frame period	8ms (128pt)
Analysis	18-th order mel-cepstral analysis
Feature parameter	$c(1) - c(18)$

表 9. GMM training conditions.

GMM	GMM with diagonal covariance matrix
Training algorithm	EM algorithm
Number of Gaussians	4, 8, 16, 32
Initialization	LBG codebook
Minimum variance	0.001
Minimum training iteration	50 times
Maximum training iteration	100 times

表 10. Speaker identification task (NTT).

Database	NTT VR database
Speakers	13 females, 22 males, 35 (13+22) speakers
Training data	90.8, 90.9, 90.12, 91.3 5 sentences (B set) / session
Test data	91.6 15 sentences (A + B set)
Data length	0.48 - 0.6 second

4, 8, 16, 32, 64 混合にて識別を行った結果、32 混合と 64 混合では識別率に大きな差は見られなかった。そこで、4 から 32 混合までの結果を付録 D に示す。女性 13 名の結果を表 14 から表 17 に、男性 22 名の結果を表 18 から表 21 に、男女混合の結果を表 22 から表 25 に示す。各表は、それぞれ上側が学習データ、下側がテストデータに対する誤り率 (%) を示している。ただし、2次オールパス関数による周波数伸縮を行った場合、 $\alpha$  の値をあまり大きくしても結果が良くならなかったため、ここでは  $\alpha$  が 0.5 までの結果のみを掲載する。

表より、学習データに関しては、混合数が小さい場合は、 $(\theta, \alpha)$  の値によって誤り率にも差が見られるが、32 混合においては、どのワーピングにおいても識別誤りがほとんどなくなるため、誤り率にあまり差は見られない。一方、テストデータに関しては、2次オールパス関数を用いて  $(\theta, \alpha)$  の組を適切に調節すれば、女性、男性、男女混合どの場合においても、従来のリニア、メル、バークの 3 つのどのスケールよりも低い誤り率が得られることがわかる。以上のことから、2次オールパス関数に基づくメルケプストラム推定法を話者識別システムの特徴抽出手法として導入し、識別的特徴抽出手法に基づき周波数伸縮関数の最適化を行うことによって、認識性能が改善されることが大いに期待できる。

### 5.3 GPD 学習による GMM パラメータの最適化

#### (5.3.1) 学習のチューニングパラメータの設定

ここでは、GPD 学習法に基づき GMM のパラメータのみの最適化を行う。GPD 学習のためのチューニングパラメータがいくつか存在するが、これらはモデルのサイズや認識タスクによって調整する必要がある。そこで、経験的にパラメータの値を振って実際に学習を行い、損失の値がほぼ順調に下がるパラメータの値を採用した。表 11 に実際に用いたパラメータの値を示す。

表 11. Adjustable parameters for MCE/GPD training.

$\gamma$ (Slope of sigmoid function)	1.5 (fixed)
$\beta$ (Offset of sigmoid function)	0 (fixed)
$\epsilon_0$ (Initial learning step size)	0.2 (sample by sample mode) 100.0 (batch mode)
$\mathcal{T}$ (Trainig iteration)	20-30 epoch

#### (5.3.2) 損失関数の検討

今回損失関数として用いたシグモイド関数

$$\ell_k(\mathbf{X}; \Lambda) = \ell(d_k) = \frac{1}{1 + \exp\{-\gamma(d_k(\mathbf{X}; \Lambda) - \beta)\}} \quad (146)$$

を微分すると次式となる。

$$\frac{\partial \ell(d_k)}{\partial d_k} = \gamma \ell(d_k) \{1 - \ell(d_k)\} \quad (147)$$

このロジスティック分布 (logistic distribution) は、次式で表されるガウス分布に近い形をしている。

$$\mathcal{N}(x | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2}\right\} \quad (148)$$

ここで、式 (147), (148) において、 $x = d_k$  とし、 $\mu$  を  $d_k$  の分布の平均、 $\sigma^2$  を  $d_k$  の分布の分散と考えれば、

$$\beta = \mu \quad (149)$$

$$\gamma = \frac{4}{\sqrt{2\pi\sigma^2}} \quad (150)$$

とおくと、図 6 に示すように、この 2 つの分布の形はほぼ一致する。そこで、学習の繰り返しごとに誤分類測度  $d_k$  の分布の平均、分散を計算し、式 (150), (150) を用いてシグモイド関数の傾き  $\gamma$  とオフセット  $\beta$  を設定して学習する方法を試みた。しかし、学習がうまく進まなかったため、最終的には表 11 に示す定数を用いた。

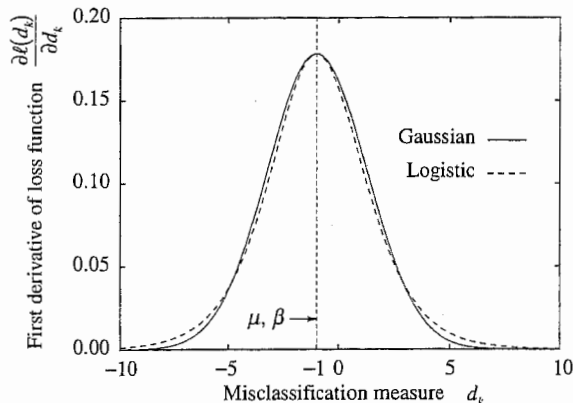


図 6. Logistic distribution and Gaussian distribution.

## (5.3.3) 損失と識別率の変化

次に、GPD 学習にて GMM のパラメータを更新した際の損失の値と識別誤り率の変化の様子を示す。例として、女性 13 名の場合において、メルスケールで分析した特徴量を用い、4 混合から 32 混合の GMM を学習した場合の損失の変化を図 7 に、またそのときの学習データおよびテストデータに対する誤り率の変化をそれぞれ図 8、図 9 に示す。図 7 の損失の変化をみると、テストデータに関して、4 混合の GMM の一括型学習において一部不安定に変化している箇所もあるが、学習データに関してはすべてほぼ滑らかに単調減少している。学習データに関して不安定な変化が生じてしまう場合は、その繰り返しステップにおける学習係数の値が多過ぎるためであり、そのような場合、一次関数の代わりに二次関数を用いたり、学習係数を適応しながら自動調整する方法（例えば Quick Propagation 法）などを採用して、学習係数の調整を行う必要がある。図より、20 エポックという同じ学習繰り返し回数においては、一括型より逐次型の方が損失の低下が大きく、図 8、図 9 の識別率についても逐次型の方が良い結果を得ているが、一括型に関しても、学習繰り返し回数を増やすなどの措置をとれば、さらに経験的損失を減少させることもできると考えられる。誤り率に関しても、学習データ、テストデータともに繰り返しにともなってほぼ単調に減少し、例えば逐次型学習を行った GMM では、テストデータに関して、4 混合で 46.4%、8 混合で 38.0%、16 混合で 17.3%、32 混合で 11.6% と、特に小さな GMM における誤り率が大きく改善されている。

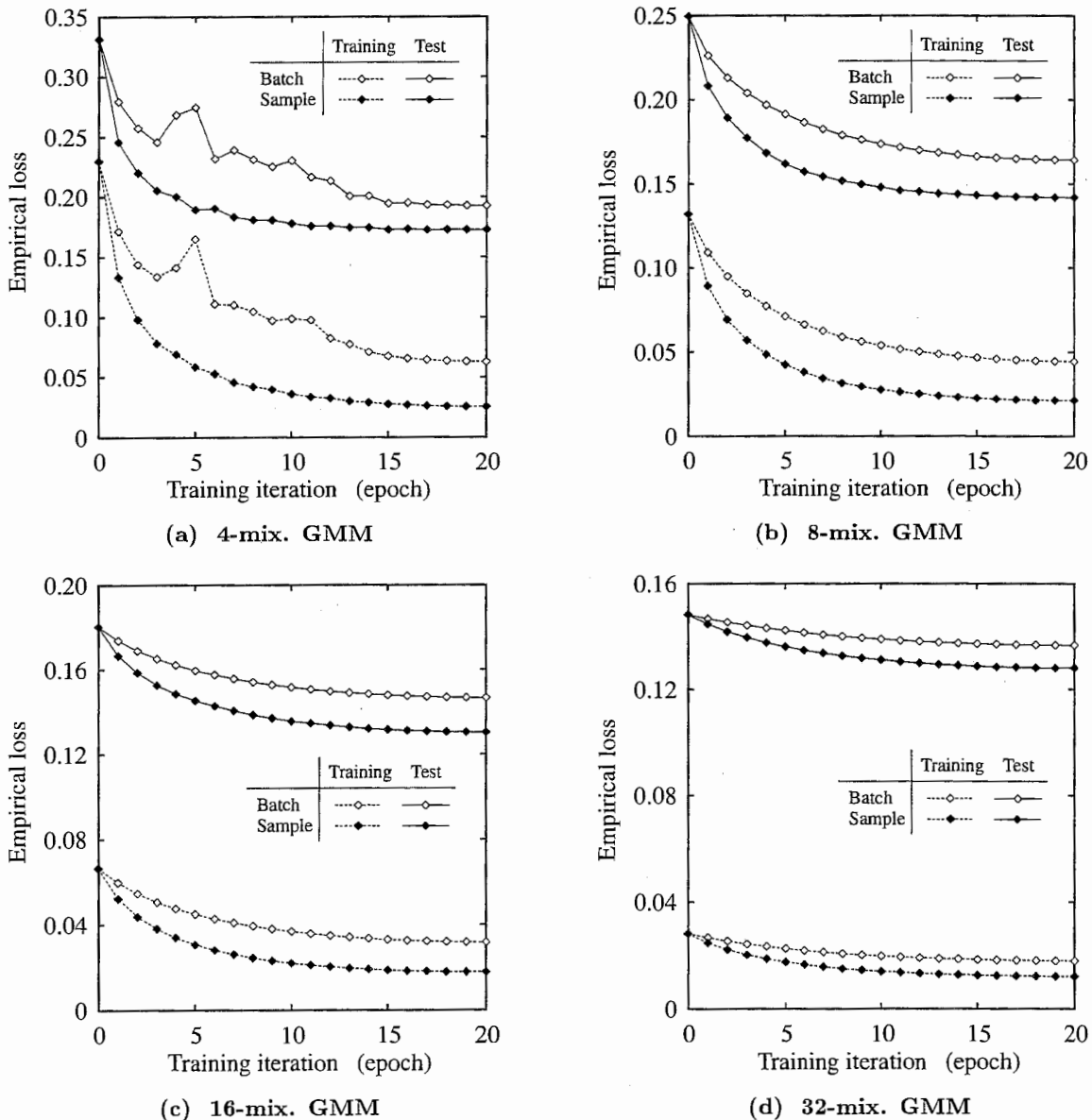
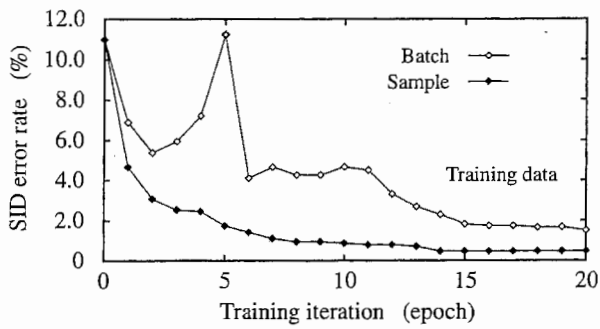
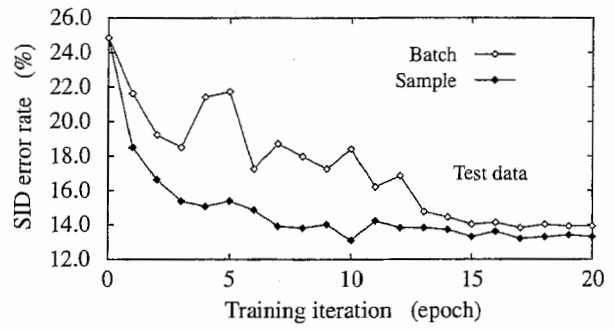


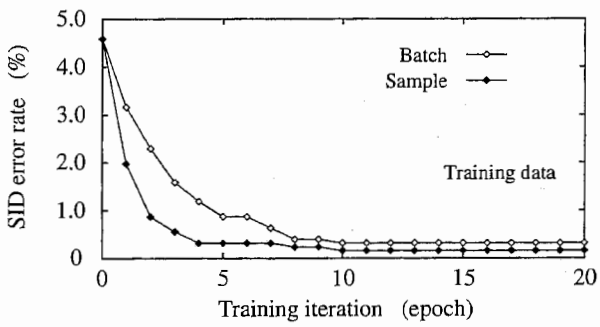
図 7. Change of empirical loss.



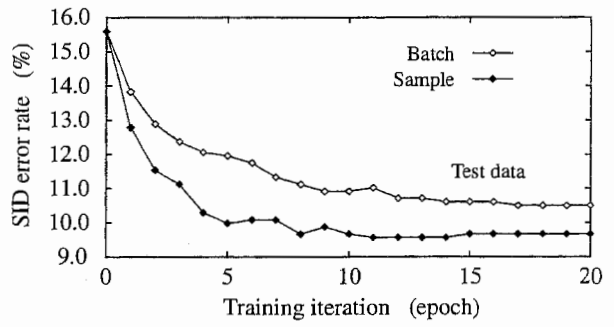
(a) 4-mix. GMM



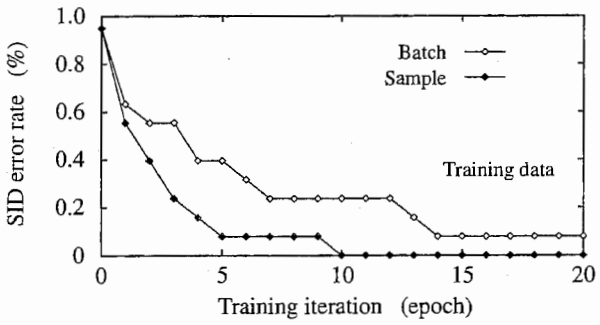
(a) 4-mix. GMM



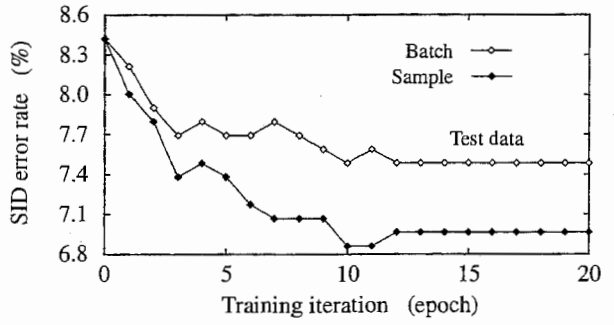
(b) 8-mix. GMM



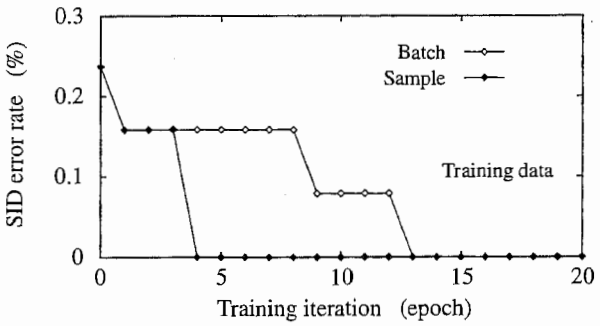
(b) 8-mix. GMM



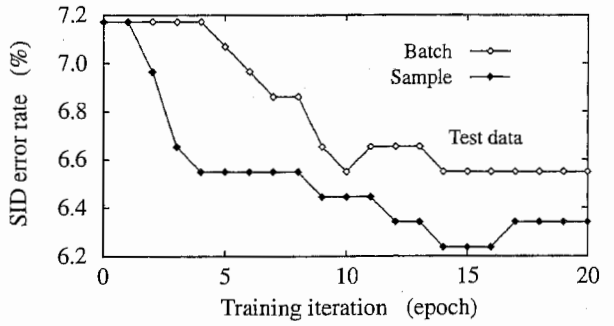
(c) 16-mix. GMM



(c) 16-mix. GMM



(d) 32-mix. GMM



(d) 32-mix. GMM

図 8. Change of error rate (Training data).

図 9. Change of error rate (Test data).

## 5.4 DFE による認識システムの最適化

### (5.4.1) ワーピングパラメータの最適化に関する検討

特徴抽出のためのパラメータセット  $\Theta = \{\theta, \alpha\}$  についても GMM のパラメータセットと同様の枠組で、式 (81) に基づき更新を行った。はじめに、GMM パラメータを固定し、 $\theta, \alpha$  のみ逐次型学習にて更新を行ってみた。しかし、2つのパラメータともにほとんど値の変化がなく、学習係数を大きくしても不安定な動きをするのみで、値が一向に収束する傾向がなかったため、その原因を調べることにした。

まず、GMM の全パラメータは固定であると考え、 $\theta, \alpha$  それぞれに関して損失がどのような形をしているかを調べた。すると、 $\theta, \alpha$  ともに、GMM の ML 学習時に用いた  $\theta, \alpha$  の値とほぼ同じ値において損失が最小値をとり、下に凸の形をしていることがわかった。例えば、 $(\theta, \alpha) = (0, 0)$  で ML 学習した GMM の場合に、 $\theta$  を固定と考えたとき、損失が  $\alpha$  に関してどのような形をしているのかを図 10 に示す。ただし、ここで示す例は女性 13 名の場合である。図のように、どの混合数においても、損失は  $\alpha = 0$  の近傍で最小となっている。次に  $\theta$  に関して損失がどのような形をしているかを示す。 $\alpha = 0$  においては、 $\theta$  の値を振っても、ワーピング関数の形は変わらないので、別のワーピング  $(\theta, \alpha) = (0.3\pi, 0.2)$  で GMM を ML 学習した場合の例を図 11 に示す。 $\theta$  についても、多少のずれはあるが、 $\theta$  の値がほぼ  $0.3\pi$  において最小値をとっている。つまり、更新を行う以前に、既に  $\theta, \alpha$  の値は、GMM のパラメータに対して最適となっていたのである。

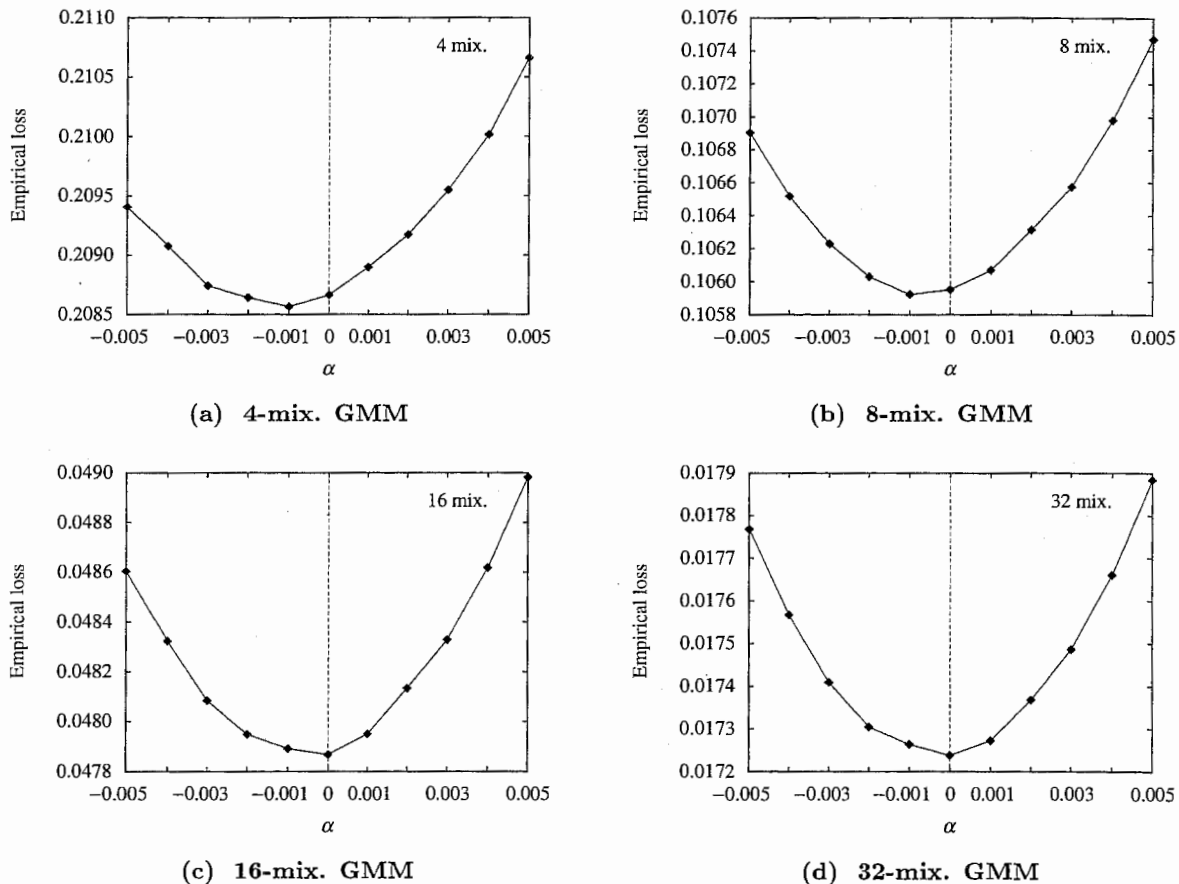
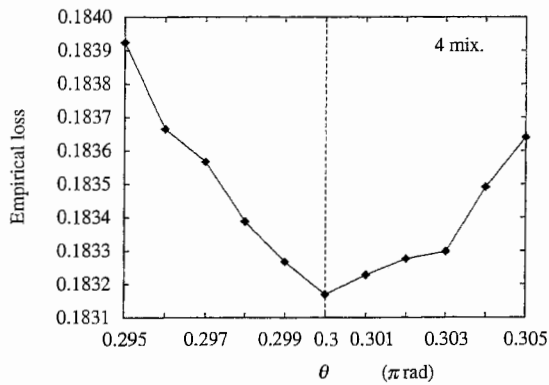
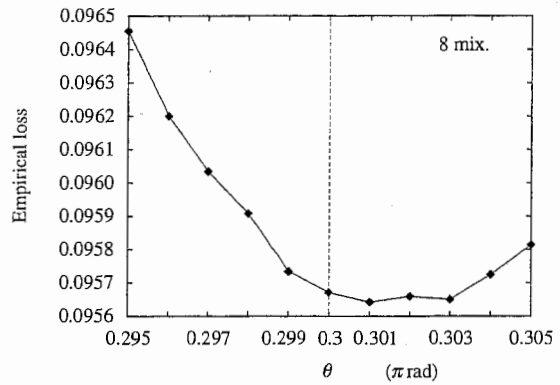


図 10. Empirical loss on training data as a function of  $\alpha$ .

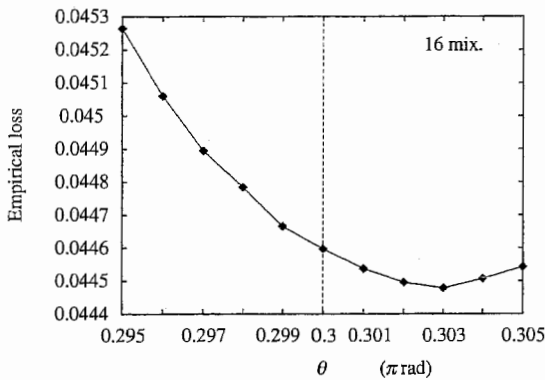
次に、式 (81) を用いて  $\theta, \alpha$  を更新した際に、損失の値が正しく最小値の方向に向かっていくかどうかをチェックした。これを調べるためには、GMM の ML 学習時の  $\theta, \alpha$  において損失が最小になっていることを利用して、 $\theta, \alpha$  の初期値を ML 学習時に用いた値からかけ離れた値に設定し、実際に更新を行って ML 学習時の  $\theta, \alpha$  の値に近づくかどうかを見ればよい。そこで、例として、 $(\theta, \alpha) = (0.3\pi, 0.2)$  で ML 学習した GMM において、 $(0.1\pi, 0.42)$  を初期値として更新した場合に、損失が最小となる  $(0.3\pi, 0.2)$  に近づくかどうかを調べた。その様子を図 12 (a) に示す。図のように、 $(\theta, \alpha)$  の値は順調に損失の最小値に向かって動いていることがわかる。また、同様に  $(0.4\pi, 0.5)$  で ML 学習した GMM において  $(0.29\pi, 0.2)$  を初期値として更新した例を図 12 (b) に示



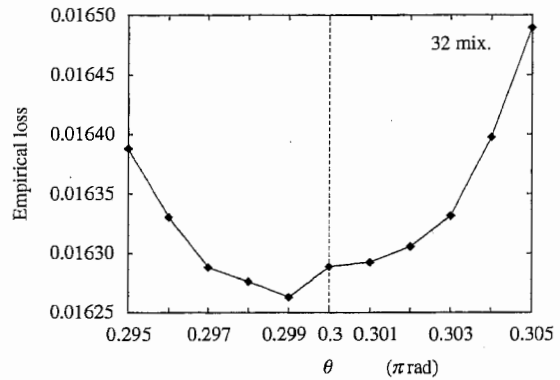
(a) 4-mix. GMM



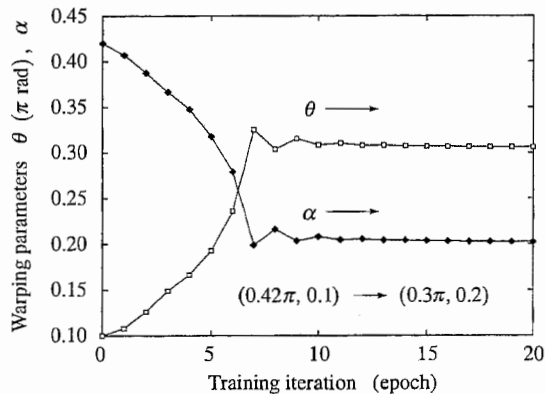
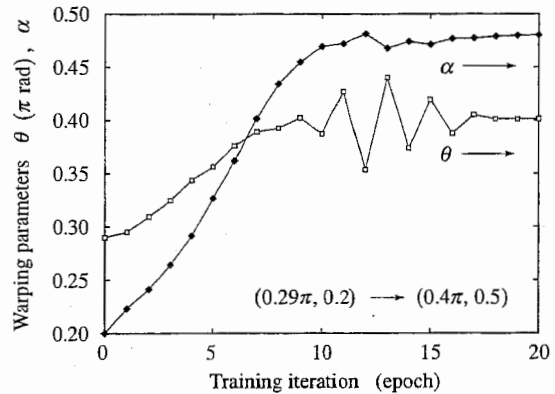
(b) 8-mix. GMM



(c) 16-mix. GMM



(d) 32-mix. GMM

図 11. Empirical loss on training data as a function of  $\theta$ .(a)  $(0.1\pi, 0.42) \rightarrow (0.3\pi, 0.2)$ (b)  $(0.29\pi, 0.2) \rightarrow (0.4\pi, 0.5)$ 図 12. Changes of  $\alpha$  and  $\theta$ .

す。これについても  $(0.4\pi, 0.5)$  に向かって動いており、GMM を固定した場合においては、その GMM に対して最適な  $\theta, \alpha$  の値を自動探索することが可能であるとわかった。

次に、特徴抽出のパラメータと GMM パラメータの同時最適化を行うために、これらのパラメータを交互に更新した。ワーピングパラメータの初期値を変えていくつか実験を行ったが、ワーピングが既に GMM に対して最適値となっており、GMM のパラメータに強く制約を受け、ワーピングパラメータについては初期値からほとんど更新されなかった。そこで、学習係数の初期値を大きめに設定し、更新とともにその値を徐々に小さくしたが、最終的にたどり着くワーピングは、ワーピングパラメータの初期値や学習係数の初期値によってさまざまな値にたどりついてしまった。例えば、学習係数の初期値  $\varepsilon_0$  の値を 0.1 から 0.5 まで変化させて、GMM、ワーピングパラメータの順に交互に 30 ずつ更新を行った場合のワーピングパラメータ値の変化の様子を図 13 に示す。

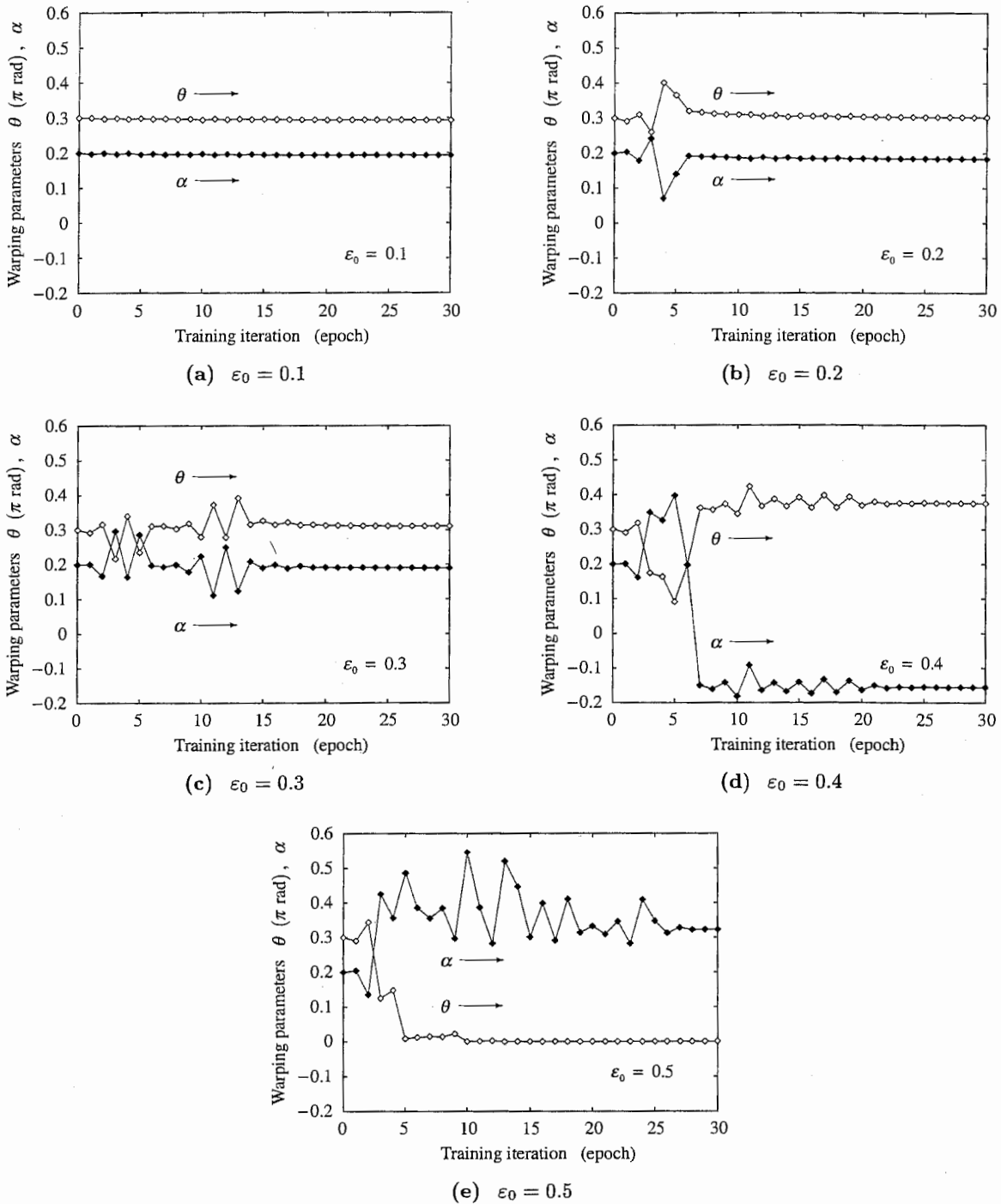


図13. Training of alternate GMM parameters and Warping parameters.

図のように、学習が終わりに近づいたところにたまたま行き着いていたワーピングの値に落ち着くという状況になってしまった。結果的に、この方法では、GMMのパラメータを固定としたとき、そのGMMに対して最適な $\theta, \alpha$ を自動探索することは可能であっても、システム全体を通してワーピングパラメータおよびGMMパラメータの最適値を勾配法により一貫して自動探索することはできなかった。これは、GMMのパラメータ数(例えば女性13名のGMMが32混合の場合15392)に比較し、特徴抽出のパラメータ数は2つのみと自由度が少なく、GMMのパラメータの制約を受けて局所解に陥り易く、大域的な解を得ることができなかったためと考えられる。そこで今回は、最適な $\theta, \alpha$ については全探索することとした。ワーピングパラメータに関する全探索は、パラメータ数が少ないため容易に行うことができる。ただし、この点に関しては、他の自由度の高いワーピング関数を導入するなど、今後さらに検討する必要がある。

## (5.4.2) 最適ワーピングパラメータの全探索

ワーピングパラメータに関しては勾配法による自動探索ができなかったため、今回は全探索を行う。 $\theta$ については0から $\pi$ まで0.05 $\pi$ ~0.1 $\pi$ 刻みに、 $\alpha$ については0から1まで0.05刻みに変化させて分析を行い、得られたケプストラム係数を用いてGMMをGPD学習にて学習し、識別を行った。結果を付録Eに掲載する。

まず、女性13名の識別タスクにおいて、4, 8, 16, 32混合のGMMを逐次型学習にて最適化した場合の結果を表26から表29に、一括型学習にて最適化した場合の結果を表30から表33に示す。また、逐次型のGPD学習に関しては、男性22名および男女35名(13+22)全員に対する実験も行った。そこで、男性に対する結果を、表34から表37に、男女混合の結果を、表38から表41に示す。各表は、それぞれ上側が学習データ、下側がテストデータに対する誤り率(%)を示している。 $\alpha$ の値を0.5より大きくしても誤り率は高くなる傾向にあったため、 $\alpha=0.5$ までの結果のみを掲載している。

ここで、最適なワーピングを選ぶ方法として、2つの方法が考えられる。一つは、学習データに対する誤り率が最小となるワーピングを最適なワーピングとする方法である。しかし、GPD学習でGMMの最適化を行った場合、逐次型、一括型どちらの学習法においても、16, 32混合のようなサイズの大きなGMMに関しては、学習データに対する誤り率が0%となる $(\theta, \alpha)$ の組が多数存在するため、学習データに対する誤り率の比較より最適なワーピングを選択することは、この場合不可能である。

そこで、もう一つの方法を考える。もう一つは、学習データに対する平均損失の値が最小となるワーピングを最適なワーピングと考える方法である。平均損失はシステムの性能をスカラー値で表したものであり、この損失が小さいほどシステムの性能も良いと考えられる。図14は、32混合のGMMを用いた男女35名の識別実験における平均損失の値を示したものである。グラフ中の各線はそれぞれ $\alpha$ が0.1から0.35の場合を示しており、横軸は強調する中心周波数 $\theta$ (kHz)に相当する。この場合、平均損失値を最小とするワーピングは、 $(\theta, \alpha) = (0.2\pi, 0.15)$ の場合であった。ここで、この最適なワーピング関数および従来のリニア、メル、バークスケールのワーピング関数の傾きを図15にて比較する。この傾きは周波数が引き伸ばされる度合に相当する。図のように、最適なワーピング関数の傾きは、従来用いられているどのワーピングの傾きとも明らかに形が異なっている。

この最適なワーピング関数を用いた提案システムの識別誤り率を、従来のシステム結果と比べたものが、表12である。これは、従来のリニア、メルおよびバークスケールを用いてML学習またはGPD学習をしたシステムと、本手法で得られた最適なワーピングを用いたシステムを比較している。

表のように、GPD学習を行ったシステムは、ML学習によるシステムと比べ、いずれのワーピングの場合も誤り率が減少し、GPD学習の効果が確認できる。また、GMMパラメータのみならず特徴抽出部のワーピングパラメータの最適化も行った提案システムにおいてはさらに性能が向上し、32混合で4.87%という識別誤り率が得られた。この結果は、リニア、メル、バークスケールのML学習によるシステムと比べ、それぞれ31%、

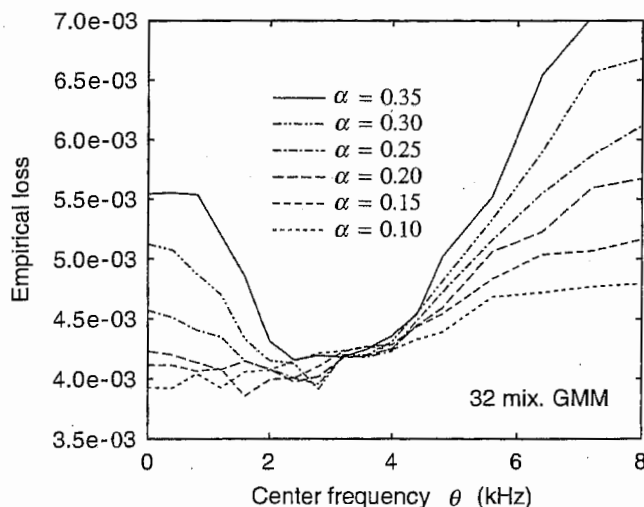


図14. Comparison of empirical losses after GPD training.



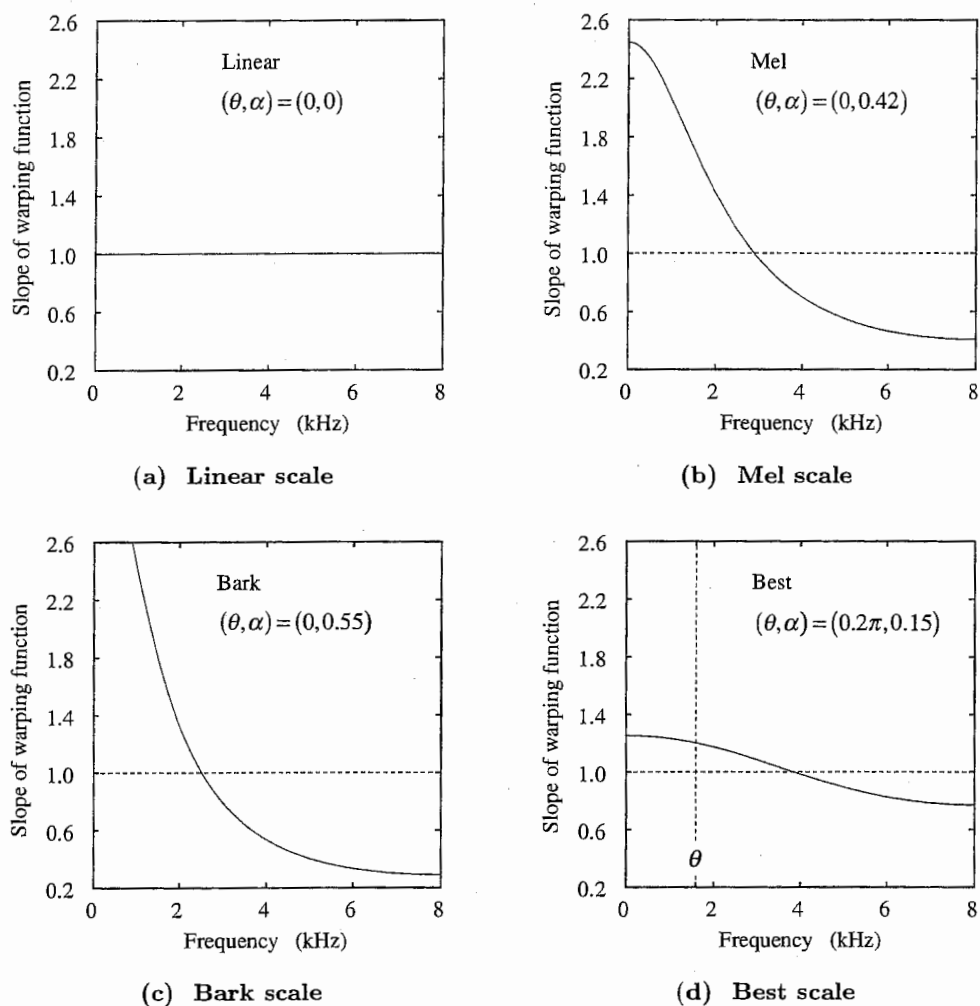


図 15. Comparison of slopes of warping functions.

表 12. Speaker identification error rates obtained by ML, GPD and DFE systems (%).

	Linear	Mel	Bark
ML-based system	7.05	5.93	7.15
GPD-based system	6.90	5.20	5.98
Proposed DFE-based system	4.87 ( $0.2\pi, 0.15$ )		

18%, 32%の誤り率改善に相当する。このことから、最適な認識システムの構築には、識別器に加え特徴抽出器も最小誤り規準で設計することが重要であると言える。

なお、本研究においては、話者の特徴が多いと言われる 5kHz 付近の高周波数帯域を強調しても、識別率はかえって低下してしまった。この点に関しては、他のワーピング関数を採用するなどして、さらに検討する必要がある。

## (5.4.3) 他の分析法との比較

本研究では、ケプストラム係数の推定方法としてメルケプストラム推定法を用いている。ここでは、このメルケプストラム推定法を他の3つの分析手法と比較する。まず18次の線形予測分析により得られた線形予測係数から18次のケプストラム係数(LPCC)を抽出した[23]。さらに、これをメルスケールに周波数伸縮しメルケプストラム係数(LPMCC)を求めた[24]。次に、18次のメル線形予測分析により得られた線形予測係数からメルケプストラム係数(MLPCC)を抽出した[25],[26]。最後に、24チャンネルのメルスケールのフィルタバンク分析により18次のメルケプストラム係数(MFCC)を抽出した[27]。なお、実験においては、どの場合も $c(0)$ を除く18個のケプストラム係数を特徴量として用いる。

男女35名の識別タスクにおいて、32混合のGMMをGPD学習する前(ML)およびGPD学習した後(GPD)の識別誤り率を表13に示す。表より、LPCCを用いた場合に比較し、リニアスケールのメルケプストラム推定法によるケプストラム係数を用いた場合の方がやや良い結果が得られることがわかる。また、メルスケールのメルケプストラム推定法についても同じメルスケールのLPMCC, MLPCC, MFCCに比べ良い結果が得られている。

これまで、音声認識において、メルケプストラム推定法が他の分析法に比べ良い結果が得られることがわかってきたが、これらの実験結果より、話者識別においてもメルケプストラム推定法が有効であることがわかった。さらに、2次オールパス関数を周波数伸縮関数として導入し、最小誤り規準にて特徴抽出器の最適化を行った提案システム(DFE)では、他のどのシステムより良い認識性能が得られている。

表13. Comparison of mel-cepstral estimation technique with other cepstral analysis techniques (%).

Cepstral coefficients		ML	GPD
LP analysis	LPCC	7.15	7.00
	LPMCC	8.11	7.81
Mel-LP analysis	MLPCC	7.35	6.44
Filter-bank analysis	MFCC	6.80	5.87
Mel-cepstral estimation	(linear)	7.05	6.90
	(mel)	5.93	5.20
	(best)	4.87	(DFE)

## 6 あとがき

本研究では、話者認識システムの構築において、特徴抽出器および分類器を識別的特徴抽出手法に基づき、最小誤り規準で設計する手法を提案した。ここでは、GMM 話者モデルを分類器とするテキスト独立型話者識別システムにおいて、2次オールパス関数に基づくメルケプストラム推定法を特徴抽出器として適用し、男女計35名による識別実験において、周波数ワーピングパラメータおよびGMMパラメータの同時最適化を行い、従来のシステムとの比較を行った。

特徴抽出器および分類器のパラメータの最適値を勾配法により同時に自動探索することはできなかったが、損失を最小とする周波数ワーピングを全探索することにより、最小分類誤りという一貫した規準で特徴抽出器と分類器の同時最適化を行うことができた。その結果、従来用いられているリニア、メル、パークスケールとは異なる周波数ワーピングが得られた。また、この最適な周波数ワーピングを用いた提案システムでは、従来の特徴量を用いてML学習あるいはGPD学習にて分類器のみ最適化したシステムと比較し、高い識別性能が得られることを確認した。

本研究において得られた最適な周波数ワーピングは、後続の識別器に依存するものであり、すべての話者認識システムにおいて不偏なものとは言えないが、これらの実験結果より、最適な話者認識システムの構築のためには、識別器のみならず特徴抽出器も一貫して最小誤り規準にて最適化することが重要であることが明らかにされた。

今後の課題としては、最適な特徴抽出のためのパラメータの自動探索法に関する検討、他のワーピング関数の導入、DMD (discriminativemetric design) [34] による話者依存の特徴抽出パラメータの採用などが挙げられる。

## 7 謝辞

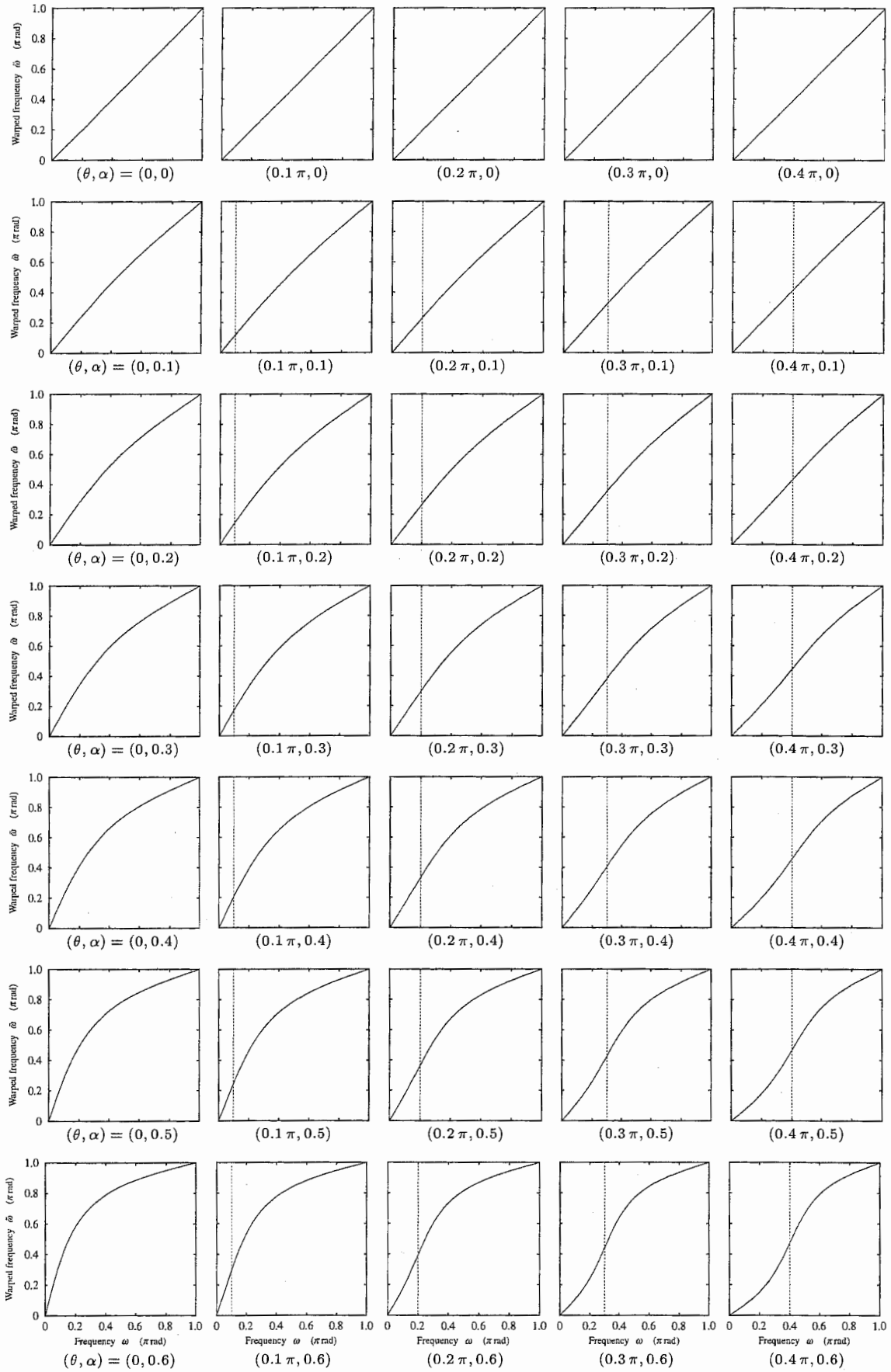
本研究を進めるにあたり、話者認識用データベースを提供して頂いた元 NTT ヒューマンインタフェース研究所古井特別研究室室長古井貞熙氏(現東工大大学院教授)、松井知子氏(現 ATR 音声翻訳研究所研究員)、ご指導、ご支援いただいた ATR 人間情報通信研究所第一研究室の諸氏に深謝いたします。さらに実務訓練の機会を与えて下さった ATR 人間情報通信研究所一ノ瀬裕社長に心から感謝いたします。

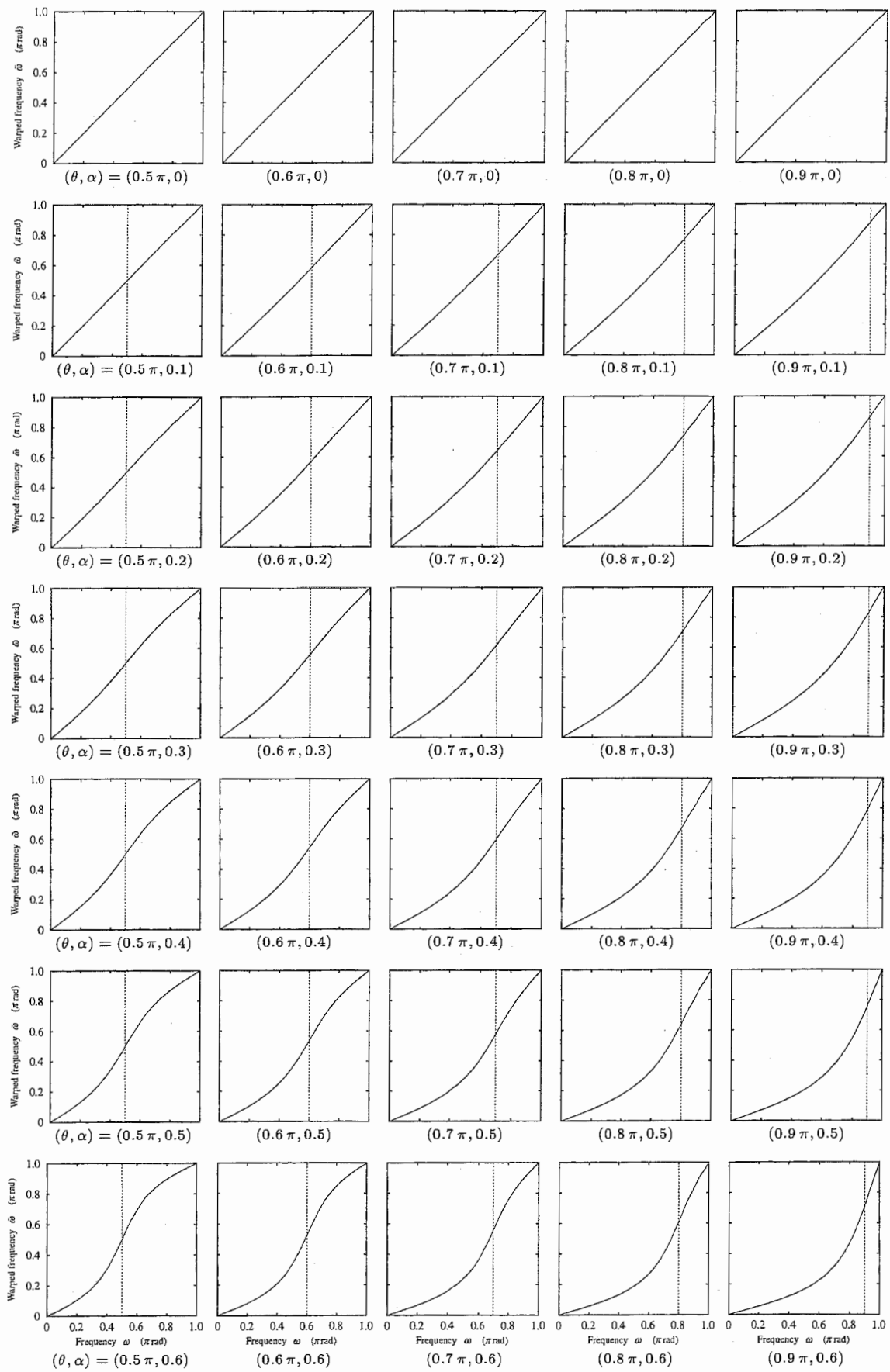
## 参考文献

- [1] 田頭 茂明, 西島 政幸, 有木 康雄: “話者部分空間への写像による話者認識と話者正規化”, 信学技報, SP95-28, pp. 25–32, Jun. 1995.
- [2] S. Hayakawa and F. Itakura: “Text-dependent speaker recognition using the information in the higher frequency band,” *Proc. ICASSP-94*, vol. 1, pp. 137–140, Apr. 1994.
- [3] S. Hayakawa and F. Itakura: “The influence of noise on the speaker recognition performance using the higher frequency band,” *Proc. ICASSP-95*, vol. 1, pp. 321–324, May 1995.
- [4] A. Biem, S. Katagiri, and B.-H. Juang: “Discriminative feature extraction for speech recognition,” *Neural Networks Signal Processing*, pp. 392–401, Sep. 1993.
- [5] S. Katagiri, C.-H. Lee, and B.-H. Juang: “New discriminative training algorithms based on the generalized probabilistic descent method,” *Neural Networks Signal Processing*, pp. 299–308, Sep. 1991.
- [6] B.-H. Juang and S. Katagiri: “Discriminative learning for minimum error classification,” *IEEE Trans. Signal Processing*, vol. 40, pp. 3043–3054, Dec. 1992.
- [7] W. Chou, B.-H. Juang, and C.-H. Lee: “Segmental GPD training of HMM based speech recognizer,” *Proc. ICASSP-92*, pp. 473–476, Mar. 1992.
- [8] C.-S. Liu, C.-H. Lee, W. Chou, B.-H. Juang, and A. E. Rosenberg: “A study on minimum error discriminative training for speaker recognition,” *J. Acoust. Soc. Am.*, vol. 97, no. 1, pp. 637–648, Jan. 1995.
- [9] C.-S. Liu: “A general framework of feature extraction: Application to speaker recognition,” *ICASSP-96*, vol. 2, pp. 669–672, May 1996.
- [10] 徳田 恵一, 益子 貴史, 小林 隆夫, 北村 正: “2次オールパス関数による周波数変換を利用したメルケプストラム分析”, 日本音響学会講演論文集, 3-7-13, Mar. 1998.
- [11] 松井知子, 古井貞熙: “テキスト指定型話者認識”, 電子情報通信学会論文誌, J79-D-II, no. 5, pp. 647–656, May 1996.
- [12] L. R. Rabiner: “A tutorial on hidden Markov models and selected applications in speech recognition system,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [13] X. D. Huang, Y. Ariki, and M. A. Jack: “Hidden Markov models for speech recognition,” Edinburgh Information Technology Series 7, Edinburgh University Press., Redwood Press LTD, 1990.
- [14] 中川 聖一: “確率モデルによる音声認識; 電子情報通信学会編”, コロナ社, 1988.
- [15] T. Vintsyuk: “Speech recognition by dynamic programming,” *Kybernetika*, 4, 1, pp. 81–88, 1968.
- [16] 迫江 博昭, 千葉 成美: “動的計画法を利用した音声の時間正規化に基づく連続音声認識”, 日本音響学会誌, 27, 9, pp. 483–500, 1971.
- [17] F. K. Soong, A. E. Rosenberg, L. R. Rabiner, and B. H. Juang: “A vector quantization approach to speaker recognition,” *Proc. ICASSP-85*, pp. 387–390, 1985.
- [18] 松井 知子, 古井 貞熙: “VQ, 離散/連続 HMM によるテキスト独立型話者認識法の比較検討”, 信学技報, SP91-89, pp. 65–70, Dec. 1991.
- [19] T. Matsui and S. Furui: “Comparison of text-independent speaker recognition methods using VQ-distortion and discrete/continuous HMMs,” *Proc. ICASSP-92*, vol. 2, pp. 157–164, Mar. 1992.

- [20] D. A. Reynolds and R. C. Rose: "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 1, pp. 72–83, Jan. 1995.
- [21] D. A. Reynolds: "Speaker identification and verification using Gaussian mixture speaker models," *Speech Communication*, vol. 17, no. 1-2, pp. 91–108, Aug. 1995.
- [22] Y. Linde, A. Buzo, and R. M. Gray: "An algorithm for vector quantization design," *IEEE Trans. Communications*, vol. 28, no. 1, pp. 84–95, Jan. 1980.
- [23] B. S. Atal: "Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification," *J. Acoust. Soc. Am.*, vol. 55, no. 6, pp. 1304–1312, Jun. 1974.
- [24] K. Shikano: "Evaluation of LPC spectral matching measures for phonetic unit recognition," *CMU Technical Report*, CMU-CS-86-108, Computer Science Department, 1986.
- [25] H. W. Strube: "Linear prediction on a warped frequency scale," *J. Acoust. Soc. Am.*, vol. 68, no. 4, pp. 1071–1076, Oct. 1980.
- [26] E. Krüger and H. W. Strube: "Linear prediction on a warped frequency scale," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-36, pp. 1529–1531, Sep. 1988.
- [27] S. B. Davis and P. Mermelstein: "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-28, pp. 357–366, Aug. 1986.
- [28] 今井 聖: "音声信号処理", 森北出版, 1996.
- [29] S. S. Stevens: "Psychophysics," John Wiley & Sons, New York, 1975.
- [30] E. Zwicker, G. Flottorp, and S. S. Stevens: "Critical bandwidth in loudness summation," *J. Acoust. Soc. Am.*, 29, pp. 548–557, 1957.
- [31] K. Tokuda, T. Kobayashi, and S. Imai: "Speech coding based on adaptive mel-cepstral analysis," *IEEE Trans. Speech and Audio Processing*, vol. 3, no. 6, pp. 481–486, Nov. 1995.
- [32] A. V. Oppenheim and D. H. Johnson: "Discrete representation of signals," *Proc. IEEE*, vol. 60, no. 7, pp. 681–691, Jun. 1972.
- [33] A. E. Yagle: "New analogs of split algorithms for arbitrary Toeplitz-plus-Hankel matrices," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-39, pp. 2457–2463, Nov. 1991.
- [34] H. Watanabe, T. Yamaguchi, and S. Katagiri: "Discriminative metric design for robust pattern recognition," *IEEE Trans. Signal Processing*, vol. 45, no. 11, pp. 2655–2662, Nov. 1997.

付録A 2次オールパス関数による周波数ワーピング







## 付録B 逆フーリエ変換値の変換行列

ここでは、次のような非直線周波数目盛 $\tilde{\omega}$ における逆フーリエ変換 $x(m)$ の値を、直線周波数目盛 $\omega$ における逆フーリエ変換の値 $x'(m)$ を用いて計算する方法について説明する。

$$x(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\tilde{\omega}m} d\omega \quad (151)$$

$$x'(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\omega m} d\omega \quad (152)$$

$$\approx \frac{1}{N} \sum_{k=0}^{N-1} X(\omega) e^{j\frac{2\pi}{N}km} = \mathcal{F}_{(m)}^{-1}[X(\omega)]$$

まず、式(151)は次のように書き換えられる。

$$\begin{aligned} x(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\tilde{\omega}m} \frac{d\omega}{d\tilde{\omega}} d\tilde{\omega} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) \frac{d\omega}{d\tilde{\omega}} e^{j\tilde{\omega}m} d\tilde{\omega} \end{aligned} \quad (153)$$

そこで、式(153)および(152)をそれぞれ両辺フーリエ変換すると、

$$\sum_{m=-\infty}^{\infty} x(m) e^{-j\tilde{\omega}m} = X(\omega) \frac{d\omega}{d\tilde{\omega}} \quad (154)$$

$$\sum_{m=-\infty}^{\infty} x'(m) e^{-j\omega m} = X(\omega) \quad (155)$$

となり、式(155)を式(154)へ代入すると、

$$\sum_{m=-\infty}^{\infty} x(m) e^{-j\tilde{\omega}m} = \sum_{m=-\infty}^{\infty} x'(m) \frac{d\omega}{d\tilde{\omega}} e^{-j\omega m} \quad (156)$$

という関係になる。上式の両辺をワープされた周波数 $\tilde{\omega}$ に関して逆フーリエ変換すると次式が得られる。

$$\begin{aligned} x(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \sum_{n=-\infty}^{\infty} x'(n) \frac{d\omega}{d\tilde{\omega}} e^{-j\omega n} \right\} e^{j\tilde{\omega}m} d\tilde{\omega} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega n} e^{j\tilde{\omega}m} \frac{d\omega}{d\tilde{\omega}} d\tilde{\omega} \right\} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega n} e^{j\tilde{\omega}m} d\omega \right\} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\tilde{\omega}m} e^{j\omega n} d\omega \right\} \\ &= \sum_{n=-\infty}^{\infty} t(m, n) x'(n) \end{aligned} \quad (157)$$

行列で表現すると次式となる。

$$\begin{bmatrix} \vdots \\ x(-1) \\ x(0) \\ x(1) \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & & \\ & t(-1, -1) & t(-1, 0) & t(-1, 1) & \\ \cdots & t(0, -1) & t(0, 0) & t(0, 1) & \cdots \\ & t(1, -1) & t(1, 0) & t(1, 1) & \\ & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ x'(-1) \\ x'(0) \\ x'(1) \\ \vdots \end{bmatrix} \quad (158)$$

ただし  $t(m, n)$  は,

$$\begin{aligned} t(m, n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\tilde{\omega}m} e^{j\tilde{\omega}n} d\tilde{\omega} \\ &\approx \frac{1}{N'} \sum_{k=0}^{N'-1} e^{-j\tilde{\omega}m} e^{j\frac{2\pi}{N'}kn} = \mathcal{F}_{(n)}^{-1} [e^{-j\tilde{\omega}m}] \end{aligned} \quad (159)$$

で与えられる。よって、この行列は逆フーリエ変換によって計算できる。ただし、この逆フーリエ変換の結果の減衰の度合を考慮すると、 $N'$  の値は  $N$  に比べて十分に大きな値にする必要がある。さらに、 $X(\omega)$  が実偶関数および純虚奇関数の場合、逆フーリエ変換の結果の対称性を用いて、以下のように計算を簡略化できる。

### B.1 実偶関数の場合

$X(\omega)$  が実偶関数の場合、つまり

$$X(-\omega) = X(\omega) = X^*(\omega) \quad (160)$$

の場合、 $x'(m)$  も実偶関数となり、

$$x'(-m) = x'(m) = x'^*(m) \quad (161)$$

を満たす。この対称性を利用すると、式 (157) は次式のように書き換えられる。

$$\begin{aligned} x(m) &= \sum_{n=-\infty}^{-1} t(m, n) x'(n) + t(m, 0) x'(0) + \sum_{n=1}^{\infty} t(m, n) x'(n) \\ &= t(m, 0) x'(0) + \sum_{n=1}^{\infty} \{t(m, n) + t(m, -n)\} x'(n) \\ &= \sum_{n=0}^{\infty} t^+(m, n) x'(n) \end{aligned} \quad (162)$$

ただし、 $t^+(m, n)$  は次式で与えられる。

$$t^+(m, n) = \begin{cases} t(m, 0), & n = 0 \\ t(m, n) + t(m, -n), & n > 0 \end{cases} \quad (163)$$

### B.2 純虚奇関数の場合

一方、 $X(\omega)$  が純虚奇関数の場合、つまり

$$X(-\omega) = -X(\omega) = X^*(\omega) \quad (164)$$

の場合、 $x'(m)$  は実奇関数となり、

$$x'(-m) = -x'(m) = x'^*(m) \quad (165)$$

を満たす。この対称性を利用すると、式 (157) は次式のように書き換えられる。

$$\begin{aligned} x(m) &= \sum_{n=-\infty}^{-1} t(m, n) x'(n) + t(m, 0) x'(0) + \sum_{n=1}^{\infty} t(m, n) x'(n) \\ &= t(m, 0) x'(0) + \sum_{n=1}^{\infty} \{t(m, n) - t(m, -n)\} x'(n) \\ &= \sum_{n=0}^{\infty} t^-(m, n) x'(n) \end{aligned} \quad (166)$$

ただし、 $t^-(m, n)$  は次式で与えられる。

$$t^-(m, n) = \begin{cases} t(m, 0), & n = 0 \\ t(m, n) - t(m, -n), & n > 0 \end{cases} \quad (167)$$

## 付録 C ケプストラム係数の周波数変換行列

## C.3 ケプストラムからメルケプストラムへの変換

まずはじめに、以下のような非直線周波数目盛 $\tilde{\omega}$ における逆フーリエ変換 $x(m)$ の値(ただし、付録 B の式 (151) において  $d\omega$  であった部分が、ここでは  $d\tilde{\omega}$  となっていることに注意する。)を、直線周波数目盛 $\omega$ における逆フーリエ変換の値 $x'(m)$ を用いて計算する方法について述べる。この変換手法は、メルケプストラム分析においてメルケプストラム係数の初期値を設定する際に、ケプストラム係数 $c'(m)$ からメルケプストラム係数 $c(m)$ を求めるために用いる。

$$x(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\tilde{\omega}m} d\tilde{\omega} \quad (168)$$

$$\begin{aligned} x'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega) e^{j\omega m} d\omega \quad (169) \\ &\approx \frac{1}{N} \sum_{k=0}^{N-1} X(\omega) e^{j\frac{2\pi}{N}km} = \mathcal{F}_{(m)}^{-1}[X(\omega)] \end{aligned}$$

式 (168) および (169) をそれぞれ両辺フーリエ変換すると、

$$\sum_{m=-\infty}^{\infty} x(m) e^{-j\tilde{\omega}m} = X(\omega) \quad (170)$$

$$\sum_{m=-\infty}^{\infty} x'(m) e^{-j\omega m} = X(\omega) \quad (171)$$

となる。式 (170), (171) より、

$$\sum_{m=-\infty}^{\infty} x(m) e^{-j\tilde{\omega}m} = \sum_{m=-\infty}^{\infty} x'(m)' e^{-j\omega m} \quad (172)$$

が得られる。上式の両辺をワープされた周波数 $\tilde{\omega}$ に関して逆フーリエ変換すると次式となる。

$$\begin{aligned} x(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \sum_{n=-\infty}^{\infty} x'(n) e^{-j\omega n} \right\} e^{j\tilde{\omega}m} d\tilde{\omega} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega n} e^{j\tilde{\omega}m} d\tilde{\omega} \right\} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega n} e^{-j\tilde{\omega}m} d\tilde{\omega} \right\} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j\omega n} e^{-j\tilde{\omega}m} \frac{d\tilde{\omega}}{d\omega} d\omega \right\} \\ &= \sum_{n=-\infty}^{\infty} x'(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\tilde{\omega}m} \frac{d\tilde{\omega}}{d\omega} e^{j\omega n} d\omega \right\} \\ &= \sum_{n=-\infty}^{\infty} u(m, n) x'(n) \quad (173) \end{aligned}$$

また、行列で表現すると次のようになる。

$$\begin{bmatrix} \vdots \\ x(-1) \\ x(0) \\ x(1) \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & & \\ & u(-1, -1) & u(-1, 0) & u(-1, 1) & \\ \cdots & u(0, -1) & u(0, 0) & u(0, 1) & \cdots \\ & u(1, -1) & u(1, 0) & u(1, 1) & \\ & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ x'(-1) \\ x'(0) \\ x'(1) \\ \vdots \end{bmatrix} \quad (174)$$

ただし,  $u(m, n)$  は,

$$\begin{aligned} u(m, n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\tilde{\omega}m} \frac{d\tilde{\omega}}{d\omega} e^{j\omega n} d\omega \\ &\approx \frac{1}{N'} \sum_{k=0}^{N'-1} e^{-j\tilde{\omega}m} \frac{d\tilde{\omega}}{d\omega} e^{j\frac{2\pi}{N'}kn} = \mathcal{F}_{(n)}^{-1} \left[ e^{-j\tilde{\omega}m} \frac{d\tilde{\omega}}{d\omega} \right] \end{aligned} \quad (175)$$

で与えられる. 式の通り, この行列は逆フーリエ変換によって計算できる. ただし, この逆フーリエ変換の結果の減衰の度合を考慮すると,  $N'$  は十分に大きな値にする必要がある. なお, 関数  $\tilde{\omega}(\omega, \theta, \alpha)$  の  $\omega$  に関する微分は微分公式 (103) より以下のようになる.

$$\begin{aligned} \frac{\partial \tilde{\omega}(\omega, \theta, \alpha)}{\partial \omega} &= \frac{\partial}{\partial \omega} \left[ \omega + \tan^{-1} \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right) + \tan^{-1} \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right) \right] \\ &= 1 + \left\{ 1 + \left( \frac{\alpha \sin(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} \right)^2 \right\}^{-1} \left\{ \frac{\alpha \cos(\omega - \theta)}{1 - \alpha \cos(\omega - \theta)} - \frac{\alpha \sin(\omega - \theta) \cdot \alpha \sin(\omega - \theta)}{\{1 - \alpha \cos(\omega - \theta)\}^2} \right\} \\ &\quad + \left\{ 1 + \left( \frac{\alpha \sin(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} \right)^2 \right\}^{-1} \left\{ \frac{\alpha \cos(\omega + \theta)}{1 - \alpha \cos(\omega + \theta)} - \frac{\alpha \sin(\omega + \theta) \cdot \alpha \sin(\omega + \theta)}{\{1 - \alpha \cos(\omega + \theta)\}^2} \right\} \\ &= 1 + \frac{\{1 - \alpha \cos(\omega - \theta)\}^2}{1 + \alpha^2 - 2\alpha \cos(\omega - \theta)} \frac{\alpha \cos(\omega - \theta) - \alpha^2}{\{1 - \alpha \cos(\omega - \theta)\}^2} \\ &\quad + \frac{\{1 - \alpha \cos(\omega + \theta)\}^2}{1 + \alpha^2 - 2\alpha \cos(\omega + \theta)} \frac{\alpha \cos(\omega + \theta) - \alpha^2}{\{1 - \alpha \cos(\omega + \theta)\}^2} \\ &= 1 + \frac{\alpha \cos(\omega - \theta) - \alpha^2}{1 - 2\alpha \cos(\omega - \theta) + \alpha^2} + \frac{\alpha \cos(\omega + \theta) - \alpha^2}{1 - 2\alpha \cos(\omega + \theta) + \alpha^2} \end{aligned} \quad (176)$$

ここで,  $x'(m)$  はケプストラム係数で実偶関数となるため, その対称性を利用すると, 式 (173) は次式のように書き換えられる.

$$\begin{aligned} x(m) &= \sum_{n=-\infty}^{-1} u(m, n) x'(n) + u(m, 0) x'(0) + \sum_{n=1}^{\infty} u(m, n) x'(n) \\ &= u(m, 0) x'(0) + \sum_{n=1}^{\infty} \{u(m, n) + u(m, -n)\} x'(n) \\ &= \sum_{n=0}^{\infty} u^+(m, n) x'(n) \end{aligned} \quad (177)$$

ただし,  $u^+(m, n)$  は次式で与えられる.

$$u^+(m, n) = \begin{cases} u(m, 0), & n = 0 \\ u(m, n) + u(m, -n), & n > 0 \end{cases} \quad (178)$$

#### C.4 メルケプストラムからケプストラムへの変換

ここでは, 逆に  $x(m)$  から  $x'(m)$  に変換する手法について説明する. この手法は, メルケプストラム分析の各繰り返しにおいて, メルケプストラム係数  $c(m)$  からケプストラム係数  $c'(m)$  に変換する際に用いる.

式 (172) の両辺を, 今度は  $\omega$  に関して逆フーリエ変換すると,

$$\begin{aligned} x'(m) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\{ \sum_{n=-\infty}^{\infty} x(n) e^{-j\tilde{\omega}n} \right\} e^{j\omega m} d\omega \\ &= \sum_{n=-\infty}^{\infty} x(n) \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\tilde{\omega}n} e^{j\omega m} d\omega \right\} \\ &= \sum_{n=-\infty}^{\infty} v(m, n) x(n) \end{aligned} \quad (179)$$

となる。ただし、 $v(m, n)$ は、

$$\begin{aligned} v(m, n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega n} e^{j\omega m} d\omega \\ &\approx \frac{1}{N'} \sum_{k=0}^{N'-1} e^{-j\omega n} e^{j\frac{2\pi}{N'} km} = \mathcal{F}_{(m)}^{-1} [e^{-j\omega n}] \\ &= t(n, m) \end{aligned} \quad (180)$$

で与えられる。行列で表現すると次式となる。

$$\begin{aligned} \begin{bmatrix} \vdots \\ x(-1) \\ x(0) \\ x(1) \\ \vdots \end{bmatrix} &= \begin{bmatrix} \ddots & & & & \\ & v(-1, -1) & v(-1, 0) & v(-1, 1) & \\ \cdots & v(0, -1) & v(0, 0) & v(0, 1) & \cdots \\ & v(1, -1) & v(1, 0) & v(1, 1) & \\ & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ x'(-1) \\ x'(0) \\ x'(1) \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} \ddots & & & & \\ & t(-1, -1) & t(0, -1) & t(1, -1) & \\ \cdots & t(-1, 0) & t(0, 0) & t(1, 0) & \cdots \\ & t(-1, 1) & t(0, 1) & t(1, 1) & \\ & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ x'(-1) \\ x'(0) \\ x'(1) \\ \vdots \end{bmatrix} \end{aligned} \quad (181)$$

つまり、式 (159) の変換行列  $t(m, n)$  を求める際に横方向に埋めた逆フーリエ変換の値を、ここでは縦方向に埋めることになる。

ここで、 $x(m)$  はメルケプストラム係数で実偶関数となるため、その対称性を利用すると、式 (179) は次式のように書き換えられる。

$$\begin{aligned} x'(m) &= \sum_{n=-\infty}^{-1} v(m, n) x(n) + v(m, 0) x(0) + \sum_{n=1}^{\infty} v(m, n) x(n) \\ &= v(m, 0) x(0) + \sum_{n=1}^{\infty} \{v(m, n) + v(m, -n)\} x(n) \\ &= \sum_{n=0}^{\infty} v^+(m, n) x(n) \end{aligned} \quad (182)$$

ただし、 $v^+(m, n)$  は次式で与えられる。

$$v^+(m, n) = \begin{cases} v(m, 0), & n = 0 \\ v(m, n) + v(m, -n), & n > 0 \end{cases} \quad (183)$$

以上のような変換行列の値は、 $\theta$  および  $\alpha$  が固定であれば定数となる。したがって、実際にはメルケプストラム分析の前にこの行列をあらかじめ計算し、入力音声の全フレームについて分析が終るまで行列の値を記憶しておけばよいので、何度も行列を計算し直す必要はなく、計算量も膨大とはならない。

## 付録D ML学習によるシステムの識別誤り率

表14. SID error rates (ML, 4-mix., 13-females) (%)

(a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	11.16	11.00	16.30							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	10.05	7.83	9.81	8.62	10.44	10.84	11.00	9.81	12.50	12.82
0.05 $\pi$	9.89	8.07	9.65	8.47	10.60	11.08	11.79	8.62	12.10	12.10
0.10 $\pi$	9.97	8.31	9.10	8.39	10.60	11.00	12.66	10.52	10.52	11.16
0.15 $\pi$	9.97	8.54	8.39	8.54	8.86	10.44	11.08	11.79	10.44	9.89
0.20 $\pi$	10.13	8.86	8.07	9.41	8.78	9.02	10.13	10.84	11.47	10.84
0.25 $\pi$	10.36	9.10	8.07	8.86	8.54	8.70	9.49	9.26	9.97	9.97
0.30 $\pi$	10.28	9.26	8.70	7.67	9.26	7.12	7.52	9.34	9.57	9.89
0.35 $\pi$	10.28	10.21	8.54	8.62	8.31	8.47	8.23	7.83	7.91	10.05
0.40 $\pi$	9.89	10.05	9.97	10.13	9.49	9.73	9.26	9.02	9.18	9.89
0.45 $\pi$	10.84	10.28	8.78	10.28	10.13	10.36	10.28	8.62	7.91	8.39
0.50 $\pi$	11.16	11.16	9.34	8.94	8.94	9.10	9.65	10.52	10.68	10.05
0.55 $\pi$	11.16	10.84	9.49	9.34	8.94	9.02	9.97	9.18	10.05	10.76
0.60 $\pi$	11.00	9.41	9.81	9.89	10.84	12.18	10.68	11.63	11.31	11.39
0.70 $\pi$	11.31	10.60	10.60	12.58	12.18	12.90	13.21	12.03	11.63	11.55
0.80 $\pi$	10.52	9.57	11.63	12.42	13.69	12.58	12.34	12.26	12.42	12.26
0.90 $\pi$	10.36	11.55	12.10	12.58	12.74	12.42	11.31	12.18	12.03	12.42
$\pi$	10.60	11.71	12.66	12.82	12.34	12.26	11.55	11.39	12.66	13.37

(b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	26.20	24.84	30.46							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	26.20	23.18	25.57	25.05	24.64	24.01	22.45	23.49	25.57	25.99
0.05 $\pi$	26.30	23.39	25.16	25.16	24.32	23.91	22.56	22.35	26.30	25.88
0.10 $\pi$	26.51	23.28	24.84	25.36	26.20	24.12	24.01	22.35	25.57	24.12
0.15 $\pi$	26.61	21.21	23.18	25.36	24.64	26.09	25.05	24.01	23.18	24.53
0.20 $\pi$	26.61	22.35	22.87	23.60	24.74	24.01	24.32	23.39	24.32	25.57
0.25 $\pi$	26.72	23.80	22.45	22.97	24.64	24.43	22.97	23.91	22.97	23.60
0.30 $\pi$	25.88	26.09	21.83	22.77	23.49	20.89	21.62	22.97	22.25	23.70
0.35 $\pi$	26.61	26.51	25.36	22.56	22.04	21.93	21.10	20.79	21.52	23.18
0.40 $\pi$	26.20	25.78	26.51	26.30	25.16	24.32	21.62	22.35	22.35	23.18
0.45 $\pi$	26.40	26.51	26.09	26.92	26.30	24.74	23.49	23.08	23.08	23.80
0.50 $\pi$	26.09	26.09	25.36	25.36	24.43	23.08	22.97	24.53	26.09	26.61
0.55 $\pi$	25.88	26.51	25.57	25.78	25.36	26.09	25.99	26.40	24.43	24.32
0.60 $\pi$	26.51	25.88	26.82	26.30	27.44	29.42	29.52	28.17	28.17	26.82
0.70 $\pi$	26.20	25.78	28.38	30.35	27.86	29.94	29.52	29.73	31.70	31.60
0.80 $\pi$	26.20	28.17	28.07	28.38	28.48	28.38	29.73	29.63	28.48	28.59
0.90 $\pi$	26.09	28.38	28.17	28.48	27.55	29.00	28.38	27.65	27.96	28.90
$\pi$	25.99	28.48	28.27	28.48	27.96	29.00	28.07	27.03	28.79	28.48

表15. SID error rates (ML, 8-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	3.64		4.59		8.47					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	2.29	3.09	3.32	3.16	3.24	4.51	4.19	3.64	5.22	6.25
0.05 $\pi$	2.29	2.93	3.24	3.16	3.48	4.11	3.72	3.48	4.75	6.17
0.10 $\pi$	2.37	2.77	3.01	3.09	3.09	4.35	3.72	3.48	4.19	4.43
0.15 $\pi$	2.14	2.77	3.01	2.85	3.09	3.32	3.96	3.96	3.88	4.27
0.20 $\pi$	2.14	2.61	2.69	3.09	3.09	3.40	3.80	3.64	4.83	4.51
0.25 $\pi$	2.37	2.53	3.01	2.77	2.53	2.77	3.16	3.01	3.56	4.19
0.30 $\pi$	2.85	2.53	2.29	2.45	2.93	2.45	2.22	2.93	3.80	3.32
0.35 $\pi$	2.93	2.29	2.69	2.69	2.22	3.09	3.32	2.85	3.01	3.48
0.40 $\pi$	3.01	2.77	2.29	2.45	2.61	3.01	2.93	2.61	2.69	3.88
0.45 $\pi$	3.01	3.09	2.69	2.69	2.77	2.69	2.77	2.77	3.09	2.61
0.50 $\pi$	3.56	3.80	3.01	2.61	3.24	3.40	3.32	4.19	3.88	4.03
0.55 $\pi$	3.24	3.24	3.16	3.16	2.93	2.85	3.09	2.85	4.11	3.88
0.60 $\pi$	3.09	3.56	3.80	3.01	3.01	3.24	2.93	3.16	3.09	3.64
0.70 $\pi$	3.24	3.56	3.09	3.16	3.56	3.40	3.72	3.56	4.67	3.32
0.80 $\pi$	3.40	3.16	3.01	3.96	4.43	4.83	4.75	4.03	4.43	4.83
0.90 $\pi$	3.56	3.09	3.48	4.59	4.43	4.27	4.43	3.96	3.88	4.75
$\pi$	3.64	3.32	3.80	4.43	3.96	5.30	4.03	4.27	4.19	4.91

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	16.11		15.59		21.62					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	14.24	11.12	15.38	14.24	14.14	13.93	14.14	15.70	15.70	16.94
0.05 $\pi$	14.35	11.43	15.07	13.83	13.93	14.24	13.72	15.18	16.84	16.94
0.10 $\pi$	14.24	11.12	14.45	13.51	14.66	15.18	13.72	14.03	15.07	16.63
0.15 $\pi$	14.14	11.64	13.72	13.20	14.03	14.55	14.76	14.76	15.38	14.76
0.20 $\pi$	14.35	12.06	11.64	13.83	12.89	13.41	14.03	14.97	15.70	17.46
0.25 $\pi$	14.86	12.58	11.33	13.51	14.24	12.89	13.62	14.66	15.18	14.76
0.30 $\pi$	16.22	14.55	11.75	11.23	12.99	13.72	13.31	13.62	14.55	15.70
0.35 $\pi$	15.49	14.86	13.20	12.99	12.27	13.31	12.99	13.62	13.93	15.49
0.40 $\pi$	15.70	16.11	15.90	15.38	15.18	14.86	13.31	13.93	13.72	16.22
0.45 $\pi$	15.70	16.11	15.18	16.11	15.90	15.38	15.18	15.70	15.70	16.42
0.50 $\pi$	16.11	16.11	16.22	15.28	16.32	17.26	17.15	15.90	16.84	17.57
0.55 $\pi$	16.94	16.32	15.38	15.38	14.86	16.94	17.88	17.67	17.78	18.19
0.60 $\pi$	16.94	16.32	15.90	15.28	16.94	17.67	17.67	18.61	19.02	18.40
0.70 $\pi$	16.11	15.28	16.22	18.30	17.88	18.92	21.41	21.62	21.73	19.44
0.80 $\pi$	16.53	15.59	17.46	18.50	19.65	20.06	20.06	18.61	19.23	20.27
0.90 $\pi$	16.01	15.18	18.61	19.33	20.69	19.02	20.48	19.44	19.96	19.85
$\pi$	15.90	15.59	18.40	18.92	20.58	19.13	18.19	18.61	20.37	19.23

表 16. SID error rates (ML, 16-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.55	0.95	2.69							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.55	0.32	0.71	0.79	0.71	0.63	0.79	0.71	1.42	2.14
0.05 $\pi$	0.47	0.32	0.63	0.79	0.79	0.87	0.95	0.95	1.03	2.29
0.10 $\pi$	0.40	0.40	0.55	0.87	0.63	0.71	0.71	0.95	1.27	1.58
0.15 $\pi$	0.47	0.47	0.63	0.87	0.79	0.79	0.95	1.19	0.79	1.27
0.20 $\pi$	0.79	0.24	0.40	0.71	0.95	0.79	0.87	0.87	0.87	0.71
0.25 $\pi$	0.71	0.63	0.24	0.71	0.40	0.63	1.03	1.27	1.11	1.19
0.30 $\pi$	0.79	0.32	0.16	0.32	0.63	0.71	0.63	0.71	0.87	0.87
0.35 $\pi$	0.87	0.63	0.40	0.16	0.16	0.40	0.71	0.87	0.63	0.55
0.40 $\pi$	0.87	0.79	0.79	0.47	0.40	0.40	0.55	0.63	0.40	0.63
0.45 $\pi$	0.63	0.63	0.40	0.55	0.47	0.40	0.55	0.55	0.71	0.71
0.50 $\pi$	0.79	0.55	0.55	0.47	0.55	0.55	0.32	0.87	0.95	0.79
0.55 $\pi$	0.63	0.55	0.87	0.55	0.63	0.55	0.55	0.40	0.47	0.71
0.60 $\pi$	0.71	0.71	0.87	0.63	0.71	0.47	0.71	0.71	1.03	0.71
0.70 $\pi$	0.79	0.71	0.71	0.71	0.63	0.71	0.95	1.03	0.79	1.03
0.80 $\pi$	0.63	0.95	0.79	1.03	1.03	1.11	1.03	0.40	0.87	1.19
0.90 $\pi$	0.79	0.79	1.11	1.03	1.19	0.87	0.55	0.79	0.87	1.03
$\pi$	0.79	0.95	1.03	0.79	1.11	0.79	0.71	0.87	0.79	0.95

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	11.33	8.42	15.38							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	9.56	7.90	8.63	8.42	9.15	8.52	9.88	9.04	9.25	11.12
0.05 $\pi$	9.98	7.80	8.84	8.73	8.63	8.63	9.25	9.46	9.15	10.71
0.10 $\pi$	9.25	8.32	7.69	8.94	8.52	8.21	9.25	8.42	9.46	9.15
0.15 $\pi$	9.98	8.63	7.17	9.46	8.32	9.56	8.73	9.98	9.15	10.29
0.20 $\pi$	9.88	7.90	7.59	8.52	9.56	8.42	8.84	9.25	10.81	10.60
0.25 $\pi$	9.98	8.73	7.80	8.11	9.04	8.94	7.69	9.25	10.50	10.08
0.30 $\pi$	11.02	8.94	7.80	8.63	7.69	8.84	7.48	7.90	8.63	10.50
0.35 $\pi$	10.71	10.19	9.77	9.46	8.32	9.25	9.15	8.11	9.56	9.98
0.40 $\pi$	10.81	10.60	10.91	9.88	10.39	9.88	9.15	9.56	10.39	10.71
0.45 $\pi$	10.91	11.02	11.43	11.64	12.47	11.75	11.54	11.64	11.12	10.91
0.50 $\pi$	11.54	11.75	11.23	11.02	10.39	11.43	12.37	12.37	11.02	11.95
0.55 $\pi$	11.12	10.19	11.12	11.23	11.02	11.43	12.06	11.43	11.95	12.89
0.60 $\pi$	10.50	10.71	10.50	11.43	11.95	12.79	12.79	12.47	13.41	12.99
0.70 $\pi$	10.39	11.85	11.85	12.68	13.41	14.14	15.07	14.86	13.62	13.93
0.80 $\pi$	11.12	11.64	12.27	14.03	14.86	15.07	13.83	12.68	12.99	14.66
0.90 $\pi$	11.02	11.23	13.72	13.62	14.97	13.20	12.99	13.20	13.62	15.07
$\pi$	11.12	11.43	13.31	13.31	14.45	13.10	12.16	13.10	13.31	13.72



表17. SID error rates (ML, 32-mix, 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.08	0.24	0.71							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.16	0	0.16	0.16	0.24	0.16	0.24	0.24	0.47	0.32
0.05 $\pi$	0.16	0.08	0.24	0.16	0.16	0.16	0.32	0.32	0.40	0.24
0.10 $\pi$	0.08	0.08	0.08	0.16	0.08	0.16	0.32	0.40	0.24	0.16
0.15 $\pi$	0.16	0.08	0	0.08	0.16	0.08	0.24	0.32	0.47	0.16
0.20 $\pi$	0.08	0	0	0.24	0.08	0.08	0.08	0.32	0.24	0.16
0.25 $\pi$	0.24	0.08	0.08	0.08	0.16	0.16	0.24	0.16	0.16	0.08
0.30 $\pi$	0.16	0.16	0.08	0.16	0.08	0.08	0.16	0.24	0.24	0
0.35 $\pi$	0.16	0	0.08	0	0	0.08	0.08	0.08	0.08	0.16
0.40 $\pi$	0.16	0.16	0.16	0	0.24	0.08	0.16	0	0.16	0.08
0.45 $\pi$	0.16	0.16	0.16	0	0	0	0.08	0.08	0	0.08
0.50 $\pi$	0.08	0.16	0.16	0.16	0.24	0.16	0.08	0	0	0
0.55 $\pi$	0.16	0.16	0.32	0.08	0.24	0.08	0.24	0.16	0.08	0.08
0.60 $\pi$	0.16	0.08	0.24	0.24	0.24	0.08	0.08	0.08	0.08	0
0.70 $\pi$	0.16	0.08	0.16	0.08	0.08	0	0.08	0	0.16	0
0.80 $\pi$	0.08	0.08	0.24	0.16	0.08	0	0.24	0	0.08	0.08
0.90 $\pi$	0.08	0.08	0.16	0.24	0	0.16	0.08	0.16	0	0
$\pi$	0.16	0.16	0.08	0.08	0.08	0.24	0	0	0.16	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	8.63	7.17	9.98							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.42	5.51	6.65	6.13	6.55	5.82	6.96	7.69	7.07	7.90
0.05 $\pi$	7.90	6.55	6.76	6.24	6.55	7.17	7.69	7.38	8.11	7.38
0.10 $\pi$	8.21	6.55	6.96	6.44	7.07	7.17	6.34	6.76	7.59	8.11
0.15 $\pi$	8.52	6.44	6.44	6.55	6.03	6.44	6.55	7.59	8.00	7.80
0.20 $\pi$	8.11	7.17	5.30	6.34	5.72	6.13	6.65	6.55	6.13	7.17
0.25 $\pi$	8.94	7.38	6.13	6.44	5.72	6.03	5.72	5.93	6.44	6.96
0.30 $\pi$	8.94	8.32	7.07	6.24	7.17	6.65	5.72	6.03	6.76	8.00
0.35 $\pi$	8.94	8.84	8.63	7.59	6.65	7.59	6.55	6.34	6.24	7.17
0.40 $\pi$	8.73	9.77	8.73	9.36	8.42	7.90	7.07	8.11	8.21	9.04
0.45 $\pi$	9.36	9.88	9.67	9.56	8.73	8.73	8.94	9.15	8.73	8.94
0.50 $\pi$	9.04	9.15	8.84	8.52	9.25	8.84	9.15	9.36	9.88	10.08
0.55 $\pi$	9.56	9.67	9.56	9.77	8.52	9.88	8.73	9.56	9.56	9.77
0.60 $\pi$	9.36	9.04	9.77	9.56	9.77	9.46	10.60	9.77	9.67	10.39
0.70 $\pi$	9.25	10.08	8.94	9.46	10.39	10.91	11.33	11.33	12.27	11.75
0.80 $\pi$	8.84	9.04	9.88	9.88	10.60	10.81	11.64	11.33	10.71	11.02
0.90 $\pi$	10.08	8.84	10.29	10.19	9.67	10.19	11.02	10.60	11.23	11.64
$\pi$	9.67	8.84	9.46	10.71	10.29	11.02	10.71	10.08	11.54	12.06

表18. SID error rates (ML, 4-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	3.91		4.63		5.34					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	3.27	3.56	3.06	2.99	3.27	3.70	3.49	4.41	4.63	4.98
0.05 $\pi$	3.49	3.56	3.06	2.92	3.13	3.84	3.42	4.48	4.48	4.84
0.10 $\pi$	3.42	3.42	2.92	3.06	2.49	3.42	3.20	3.70	4.41	4.56
0.15 $\pi$	3.56	2.99	2.70	2.92	3.06	3.27	3.20	3.63	3.56	3.63
0.20 $\pi$	3.56	3.06	3.20	3.13	3.13	3.35	3.56	3.42	3.35	3.63
0.25 $\pi$	3.56	2.92	3.49	3.06	3.27	3.56	3.77	3.84	3.84	3.63
0.30 $\pi$	3.63	2.92	3.13	3.13	2.85	3.35	3.84	3.63	3.56	2.99
0.35 $\pi$	3.84	3.35	3.27	3.49	2.92	3.35	3.63	3.84	3.27	3.06
0.40 $\pi$	3.84	3.70	3.42	3.56	3.84	3.63	3.49	3.35	3.70	3.77
0.45 $\pi$	3.91	3.84	3.77	3.63	3.13	3.27	3.42	3.91	4.13	3.84
0.50 $\pi$	3.99	4.06	3.49	3.49	3.35	3.27	2.92	3.27	3.20	3.20
0.55 $\pi$	3.99	3.84	3.63	4.13	3.84	4.27	4.13	3.42	4.34	4.13
0.60 $\pi$	4.06	4.13	4.63	3.99	4.70	4.34	4.06	3.99	4.70	4.63
0.70 $\pi$	4.48	4.98	4.77	4.91	4.41	4.98	4.91	5.20	6.12	6.55
0.80 $\pi$	4.63	4.48	4.77	4.98	4.34	5.62	6.62	6.55	7.40	7.69
0.90 $\pi$	4.63	4.91	4.84	4.06	5.27	5.55	6.48	7.54	8.04	9.61
$\pi$	4.63	5.34	4.84	4.13	4.70	5.98	6.69	8.04	8.40	9.11

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	13.76		12.87		12.97					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	12.48	10.59	11.19	11.19	12.38	12.18	11.09	11.58	13.56	11.98
0.05 $\pi$	12.67	10.79	11.19	11.49	12.08	12.18	12.67	12.38	12.77	12.18
0.10 $\pi$	12.57	10.99	10.99	10.79	12.48	11.09	12.48	10.79	12.48	13.27
0.15 $\pi$	12.77	11.58	11.39	10.49	11.49	11.78	10.10	12.77	11.09	11.98
0.20 $\pi$	12.77	11.39	10.49	11.68	11.29	11.09	10.30	11.88	11.58	12.97
0.25 $\pi$	12.57	11.39	10.49	11.39	11.88	11.09	11.58	11.78	11.29	10.89
0.30 $\pi$	13.07	11.29	11.68	10.20	11.58	12.48	12.77	13.37	11.29	12.18
0.35 $\pi$	13.37	12.48	11.19	12.08	11.19	11.78	13.86	14.55	14.36	14.36
0.40 $\pi$	12.97	12.57	11.58	12.57	13.07	13.37	13.76	14.16	15.25	14.75
0.45 $\pi$	13.66	13.07	13.56	13.56	13.07	12.97	13.27	13.66	13.56	14.36
0.50 $\pi$	13.76	13.86	14.26	13.47	14.16	14.26	13.17	13.27	13.76	14.26
0.55 $\pi$	13.96	14.46	14.36	14.46	14.26	14.16	14.85	15.05	15.64	14.55
0.60 $\pi$	14.55	15.94	15.45	15.15	16.04	15.25	16.34	15.94	18.12	17.52
0.70 $\pi$	15.84	15.74	14.85	15.74	16.34	17.33	16.24	16.73	16.34	18.12
0.80 $\pi$	15.94	14.95	16.14	16.83	15.84	16.34	16.53	17.43	16.14	20.89
0.90 $\pi$	15.94	15.45	17.03	14.95	15.15	16.44	16.73	19.01	20.50	20.40
$\pi$	15.64	15.74	16.24	14.65	14.75	16.73	17.33	19.11	21.39	22.77

表19. SID error rates (ML, 8-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	1.42	1.64	1.85							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	1.28	1.57	1.28	1.21	1.21	1.07	1.57	1.71	1.64	2.06
0.05 $\pi$	1.28	1.57	1.28	1.35	1.14	1.07	1.85	1.71	1.78	2.42
0.10 $\pi$	1.28	1.35	1.14	1.49	1.21	1.07	1.35	1.21	1.57	1.99
0.15 $\pi$	1.21	1.49	1.00	1.35	1.28	1.14	1.57	1.28	1.21	1.42
0.20 $\pi$	1.28	1.49	1.07	1.42	1.35	1.35	1.07	1.21	1.57	1.21
0.25 $\pi$	1.42	1.64	1.00	1.07	1.78	1.14	1.49	0.85	0.85	1.14
0.30 $\pi$	1.71	1.35	1.35	1.14	1.21	1.42	0.85	0.71	0.85	1.14
0.35 $\pi$	1.57	1.14	1.57	1.49	1.21	1.07	1.14	1.35	1.21	1.07
0.40 $\pi$	1.28	1.28	1.21	1.42	1.07	0.85	1.00	1.28	1.42	1.14
0.45 $\pi$	1.35	1.07	1.07	1.28	1.28	1.28	1.07	1.07	0.64	0.93
0.50 $\pi$	1.49	1.42	1.35	1.35	1.57	1.35	1.35	1.14	1.35	1.07
0.55 $\pi$	1.35	1.21	1.49	1.64	1.35	1.28	0.85	1.07	1.14	1.14
0.60 $\pi$	1.21	1.57	1.57	1.71	1.49	1.28	1.64	1.99	1.49	2.06
0.70 $\pi$	1.42	1.49	1.85	1.64	1.57	1.78	1.85	1.92	1.92	2.06
0.80 $\pi$	1.35	1.71	1.71	1.78	1.71	1.78	1.64	1.85	2.14	1.99
0.90 $\pi$	1.42	1.42	1.57	1.49	1.49	1.71	2.06	2.63	2.35	3.56
$\pi$	1.35	1.64	1.71	1.42	1.78	1.85	2.06	2.49	2.56	3.27

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	9.80	8.51	7.52							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.02	8.42	8.02	7.43	7.43	7.52	8.71	7.82	8.51	8.71
0.05 $\pi$	8.02	8.02	8.12	7.43	7.33	7.82	8.61	8.22	8.42	9.51
0.10 $\pi$	8.02	8.91	8.42	8.22	7.13	7.92	8.51	8.61	8.71	8.91
0.15 $\pi$	8.12	8.81	8.61	7.52	7.43	7.72	8.51	8.12	8.91	8.42
0.20 $\pi$	8.22	9.21	8.02	8.91	7.62	7.72	7.13	7.72	8.61	9.41
0.25 $\pi$	8.71	9.11	8.51	8.51	8.81	7.72	7.62	7.82	7.82	7.52
0.30 $\pi$	8.42	8.02	8.51	9.11	8.61	9.11	8.42	8.61	8.02	9.01
0.35 $\pi$	8.51	8.51	8.81	9.80	8.61	8.61	9.60	8.81	9.31	9.90
0.40 $\pi$	8.91	9.11	8.91	8.12	8.42	8.71	8.81	9.41	9.70	10.00
0.45 $\pi$	8.71	9.01	8.42	8.51	8.61	9.21	9.80	9.21	10.10	10.49
0.50 $\pi$	8.91	8.91	9.70	9.70	9.41	8.81	9.11	9.51	9.80	10.49
0.55 $\pi$	9.60	9.11	9.41	9.21	9.01	10.00	9.41	10.20	9.70	10.10
0.60 $\pi$	9.21	9.51	10.20	9.60	9.80	10.20	9.90	11.58	11.09	10.89
0.70 $\pi$	9.80	9.70	9.60	11.49	11.29	10.79	10.10	10.40	11.29	11.09
0.80 $\pi$	10.40	9.60	10.89	10.59	10.30	10.30	10.79	11.68	13.47	12.38
0.90 $\pi$	10.89	9.51	10.99	9.90	10.10	10.20	12.77	13.47	14.75	16.04
$\pi$	10.30	10.10	10.40	10.30	10.40	11.39	12.77	13.27	14.06	16.93

表20. SID error rates (ML, 16-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.14	0.21	0.21	0.21	0.21	0.21				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.21	0.28	0.21	0.14	0.28	0.36	0.07	0.14	0.14	0.43
0.05 $\pi$	0.21	0.21	0.14	0.14	0.28	0.36	0.28	0.43	0.14	0.36
0.10 $\pi$	0.21	0.21	0.14	0.28	0.14	0	0.21	0.28	0.28	0.14
0.15 $\pi$	0.14	0.36	0.21	0.28	0.14	0.14	0.28	0.28	0.36	0.36
0.20 $\pi$	0.21	0.36	0.07	0.28	0.14	0	0.14	0.14	0.14	0.21
0.25 $\pi$	0.21	0.36	0.14	0.28	0.14	0.07	0	0.14	0.07	0.14
0.30 $\pi$	0.14	0.21	0.28	0.07	0.28	0.36	0.36	0.21	0.14	0.14
0.35 $\pi$	0.14	0.21	0.28	0.28	0.36	0.21	0.21	0.28	0.21	0.14
0.40 $\pi$	0.07	0.14	0.14	0.21	0.14	0.14	0.21	0.07	0.14	0.14
0.45 $\pi$	0	0.07	0.14	0.07	0.14	0.21	0.07	0.07	0.07	0.07
0.50 $\pi$	0.14	0	0	0	0.14	0.14	0	0.14	0.21	0.21
0.55 $\pi$	0.21	0.36	0.14	0.14	0.07	0.28	0.28	0.21	0.21	0.21
0.60 $\pi$	0.21	0.14	0.14	0.14	0.14	0.14	0.21	0.28	0.14	0.14
0.70 $\pi$	0.14	0.14	0.14	0.28	0.28	0.21	0.36	0.50	0.57	0.64
0.80 $\pi$	0.07	0.14	0.28	0.07	0.36	0.50	0.50	0.36	0.50	0.71
0.90 $\pi$	0.14	0.21	0.21	0.36	0.43	0.28	0.36	0.78	0.57	1.07
$\pi$	0.14	0.28	0.14	0.36	0.43	0.36	0.36	0.85	0.64	0.71

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	6.04	5.94	5.94	5.54	5.54	5.54				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	5.74	5.25	5.94	5.05	5.45	6.14	6.63	5.94	6.24	4.95
0.05 $\pi$	5.45	5.45	5.64	5.25	5.35	5.74	7.23	6.04	6.04	5.54
0.10 $\pi$	6.04	5.35	5.74	5.15	4.95	6.04	6.44	6.34	6.63	5.35
0.15 $\pi$	5.74	5.35	6.14	5.54	5.54	5.25	6.53	6.14	5.74	5.84
0.20 $\pi$	6.24	5.45	5.84	6.04	5.94	5.05	5.64	6.53	6.53	6.34
0.25 $\pi$	5.94	5.64	6.14	5.84	5.64	6.34	6.83	5.64	6.44	5.94
0.30 $\pi$	5.84	5.84	6.34	6.04	6.24	6.24	6.04	7.23	6.93	6.83
0.35 $\pi$	5.45	5.94	6.04	6.14	6.34	6.14	7.52	7.52	9.21	8.71
0.40 $\pi$	6.14	5.74	6.14	6.34	5.45	6.53	6.04	8.42	7.72	7.52
0.45 $\pi$	6.24	6.24	5.54	6.53	6.73	6.63	6.63	7.33	6.83	8.12
0.50 $\pi$	6.53	6.14	6.73	7.13	7.13	6.73	6.53	7.52	8.02	8.71
0.55 $\pi$	6.44	7.03	6.93	6.24	7.13	7.13	6.83	7.03	7.62	8.22
0.60 $\pi$	6.53	6.93	7.23	7.72	6.83	8.12	7.52	8.12	8.32	9.90
0.70 $\pi$	6.93	7.62	7.52	7.72	8.71	7.92	8.51	8.42	8.91	9.21
0.80 $\pi$	6.53	8.32	8.12	8.71	8.22	8.71	8.02	9.51	9.70	11.49
0.90 $\pi$	6.53	7.23	8.12	8.61	8.32	8.32	10.40	10.20	12.87	13.07
$\pi$	7.43	7.52	8.51	7.92	8.81	8.22	10.89	10.69	13.56	12.77

表 21. SID error rates (ML, 32-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0	0	0.07					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0	0	0	0	0	0
0.05 $\pi$	0.07	0	0	0	0	0	0	0	0	0
0.10 $\pi$	0	0	0	0	0	0	0	0	0	0.07
0.15 $\pi$	0	0	0	0	0	0	0	0	0	0
0.20 $\pi$	0.07	0	0	0	0	0	0	0	0	0
0.25 $\pi$	0.07	0	0	0	0	0	0	0	0	0
0.30 $\pi$	0	0	0	0	0	0	0	0	0	0
0.35 $\pi$	0	0	0	0.07	0	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0
0.45 $\pi$	0	0.07	0	0.07	0	0	0	0	0	0
0.50 $\pi$	0	0	0	0	0	0	0	0	0	0
0.55 $\pi$	0.07	0	0	0	0	0	0	0	0.07	0.07
0.60 $\pi$	0	0	0	0	0	0	0	0	0	0.07
0.70 $\pi$	0	0	0	0	0	0	0.07	0	0.07	0.21
0.80 $\pi$	0	0	0	0	0	0.07	0.07	0.07	0.14	0.14
0.90 $\pi$	0	0	0	0	0.07	0.14	0.07	0.14	0.07	0.07
$\pi$	0	0	0	0.14	0.07	0.14	0	0.14	0	0.07

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	5.54	4.75	4.46			4.46				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	4.75	4.65	4.95	4.26	4.36	5.05	5.25	5.25	4.55	5.05
0.05 $\pi$	5.05	5.45	4.85	4.46	4.16	4.46	4.55	4.95	4.46	4.06
0.10 $\pi$	4.85	5.25	4.46	4.55	4.85	4.55	4.95	4.06	4.65	4.55
0.15 $\pi$	4.85	4.36	4.46	4.36	4.46	5.25	5.25	5.15	4.36	5.35
0.20 $\pi$	4.85	4.85	4.46	3.76	4.75	4.55	4.16	4.85	4.06	5.35
0.25 $\pi$	4.46	4.65	4.85	4.36	4.16	4.55	5.74	5.54	5.05	4.46
0.30 $\pi$	4.85	4.95	4.75	4.75	4.55	4.75	5.35	5.05	5.74	5.74
0.35 $\pi$	4.55	5.54	5.05	4.65	4.85	4.26	6.04	5.54	6.24	6.93
0.40 $\pi$	5.15	4.95	4.75	5.54	4.95	4.75	5.45	5.45	6.93	7.03
0.45 $\pi$	5.94	5.54	5.15	4.95	4.85	5.15	5.35	6.34	6.73	6.63
0.50 $\pi$	6.04	5.84	5.94	6.53	5.05	5.94	6.14	5.94	6.93	7.23
0.55 $\pi$	6.14	5.54	5.35	6.34	5.45	5.94	7.03	6.73	6.63	7.72
0.60 $\pi$	5.15	5.54	5.94	5.94	6.83	6.44	6.44	7.52	7.33	7.62
0.70 $\pi$	5.74	6.04	6.04	7.52	7.52	6.73	8.12	7.72	7.92	8.42
0.80 $\pi$	5.84	6.14	7.82	6.83	7.13	8.02	8.02	7.72	9.41	9.70
0.90 $\pi$	5.74	6.53	7.33	7.03	7.13	8.02	8.51	9.60	10.30	11.88
$\pi$	5.15	6.63	7.03	6.83	7.43	7.92	9.31	8.51	11.68	11.09

表 22. SID error rates (ML, 4-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	7.38		7.76		10.64					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	6.56	5.62	6.33	5.77	6.74	7.23	7.08	7.04	8.47	8.77
0.05 $\pi$	6.63	5.73	6.26	5.62	6.74	7.31	7.42	6.48	8.21	8.32
0.10 $\pi$	6.63	5.81	5.92	5.66	6.37	7.12	7.76	6.97	7.38	7.72
0.15 $\pi$	6.71	5.70	5.51	5.73	5.88	6.74	6.97	7.57	6.86	6.63
0.20 $\pi$	6.82	5.88	5.58	6.18	5.88	6.07	6.71	7.01	7.34	7.12
0.25 $\pi$	6.89	5.92	5.77	5.88	5.84	6.03	6.52	6.44	6.78	6.67
0.30 $\pi$	6.93	5.96	5.81	5.36	5.96	5.21	5.58	6.37	6.44	6.33
0.35 $\pi$	7.01	6.71	5.88	6.03	5.55	5.84	5.81	5.73	5.47	6.41
0.40 $\pi$	6.78	6.86	6.63	6.86	6.74	6.67	6.29	6.07	6.33	6.71
0.45 $\pi$	7.23	6.97	6.26	6.97	6.67	6.82	6.78	6.33	6.07	6.18
0.50 $\pi$	7.42	7.46	6.33	6.14	6.11	6.11	6.14	6.71	6.86	6.67
0.55 $\pi$	7.42	7.16	6.41	6.67	6.44	6.74	6.93	6.22	7.19	7.61
0.60 $\pi$	7.34	6.67	7.23	6.86	7.72	8.28	7.42	7.94	8.09	8.28
0.70 $\pi$	7.76	7.72	7.72	8.73	8.24	9.03	9.29	8.80	9.18	9.44
0.80 $\pi$	7.46	7.08	8.24	8.69	9.18	9.22	9.70	9.63	10.19	10.23
0.90 $\pi$	7.46	8.17	8.36	8.36	9.18	9.22	9.14	10.27	10.38	11.69
$\pi$	7.57	8.43	8.65	8.58	8.69	9.33	9.40	10.15	10.87	11.84

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	19.83		18.71		21.60					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	19.27	16.84	18.36	18.00	18.41	18.05	16.73	17.44	19.42	18.81
0.05 $\pi$	19.47	17.04	18.20	18.20	18.10	18.00	17.55	17.29	19.37	18.86
0.10 $\pi$	19.52	17.09	17.85	17.95	19.22	17.55	18.15	16.53	18.86	18.56
0.15 $\pi$	19.68	16.38	17.14	17.90	17.95	18.81	17.49	18.31	17.04	18.10
0.20 $\pi$	19.78	16.84	16.58	17.65	18.00	17.44	17.24	17.60	18.00	19.22
0.25 $\pi$	19.68	17.49	16.38	17.09	18.20	17.70	17.19	17.75	17.04	17.14
0.30 $\pi$	19.52	18.61	16.73	16.43	17.49	16.63	17.09	18.10	16.68	17.90
0.35 $\pi$	20.03	19.57	18.15	17.34	16.53	16.84	17.49	17.65	17.85	18.66
0.40 $\pi$	19.52	19.27	19.12	19.37	19.02	18.86	17.70	18.20	18.81	18.91
0.45 $\pi$	19.88	19.68	19.73	20.18	19.62	18.81	18.31	18.41	18.31	19.12
0.50 $\pi$	19.78	19.83	19.68	19.32	19.22	18.71	18.05	18.86	19.78	20.33
0.55 $\pi$	19.88	20.44	19.93	20.08	19.83	20.13	20.39	20.64	19.98	19.52
0.60 $\pi$	20.49	20.94	21.15	20.69	21.70	22.36	22.92	22.06	23.23	22.31
0.70 $\pi$	21.04	20.79	21.55	22.97	22.06	23.58	22.92	23.38	24.04	25.00
0.80 $\pi$	21.10	21.50	22.06	22.62	22.16	22.46	23.17	23.68	22.52	25.10
0.90 $\pi$	21.10	21.86	22.67	21.60	21.30	22.67	22.72	23.94	24.59	25.71
$\pi$	20.84	22.01	22.31	21.55	21.40	22.97	23.02	23.63	25.41	26.52

表 23. SID error rates (ML, 8-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	2.47		3.11		5.02					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	1.76	2.29	2.25	2.17	2.21	2.74	2.85	2.70	3.41	4.08
0.05 $\pi$	1.76	2.21	2.21	2.25	2.29	2.55	2.77	2.59	3.22	4.27
0.10 $\pi$	1.80	2.02	2.02	2.29	2.14	2.66	2.55	2.32	2.81	3.18
0.15 $\pi$	1.65	2.10	1.95	2.10	2.17	2.21	2.74	2.59	2.51	2.81
0.20 $\pi$	1.69	2.02	1.84	2.25	2.21	2.36	2.40	2.40	3.15	2.77
0.25 $\pi$	1.87	2.10	1.95	1.87	2.17	1.95	2.32	1.91	2.17	2.62
0.30 $\pi$	2.25	1.91	1.84	1.76	2.02	1.91	1.50	1.80	2.29	2.21
0.35 $\pi$	2.21	1.69	2.10	2.06	1.72	2.06	2.17	2.06	2.06	2.25
0.40 $\pi$	2.10	1.99	1.72	1.91	1.80	1.87	1.95	1.95	2.06	2.47
0.45 $\pi$	2.14	2.02	1.84	1.95	1.99	1.95	1.91	1.95	1.80	1.76
0.50 $\pi$	2.47	2.55	2.14	1.95	2.36	2.32	2.32	2.62	2.59	2.51
0.55 $\pi$	2.25	2.17	2.29	2.36	2.10	2.06	1.95	1.95	2.55	2.47
0.60 $\pi$	2.10	2.51	2.62	2.32	2.21	2.25	2.29	2.59	2.29	2.92
0.70 $\pi$	2.29	2.47	2.44	2.36	2.51	2.59	2.81	2.81	3.37	2.77
0.80 $\pi$	2.32	2.40	2.32	2.85	3.07	3.37	3.18	3.03	3.41	3.45
0.90 $\pi$	2.44	2.21	2.51	3.00	3.00	3.03	3.26	3.41	3.18	4.20
$\pi$	2.44	2.44	2.74	2.88	2.88	3.63	3.18	3.48	3.48	4.35

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	12.88		11.97		14.40					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	11.05	9.74	11.66	10.85	10.75	10.75	11.41	11.66	12.02	12.73
0.05 $\pi$	11.11	9.69	11.56	10.65	10.60	11.05	11.16	11.61	12.53	13.13
0.10 $\pi$	11.05	9.99	11.41	10.90	10.85	11.51	11.11	11.31	11.82	12.68
0.15 $\pi$	11.05	10.24	11.16	10.34	10.70	11.16	11.66	11.41	12.12	11.51
0.20 $\pi$	11.26	10.65	9.84	11.36	10.24	10.55	10.60	11.31	12.17	13.44
0.25 $\pi$	11.76	10.80	9.89	11.00	11.51	10.29	10.60	11.21	11.46	11.11
0.30 $\pi$	12.27	11.21	10.14	10.19	10.85	11.46	10.85	11.11	11.26	12.32
0.35 $\pi$	11.97	11.61	10.95	11.36	10.45	10.95	11.31	11.21	11.61	12.68
0.40 $\pi$	12.27	12.58	12.37	11.71	11.76	11.76	11.11	11.66	11.71	13.03
0.45 $\pi$	12.17	12.53	11.76	12.27	12.22	12.32	12.53	12.37	12.83	13.39
0.50 $\pi$	12.47	12.47	12.93	12.47	12.83	12.93	13.13	12.78	13.34	14.00
0.55 $\pi$	13.24	12.68	12.37	12.27	11.92	13.44	13.64	13.95	13.79	14.20
0.60 $\pi$	13.03	12.83	12.98	12.42	13.34	13.89	13.74	15.06	15.01	14.71
0.70 $\pi$	12.88	12.47	12.93	14.86	14.55	14.81	15.72	16.02	16.48	15.26
0.80 $\pi$	13.44	12.63	14.15	14.50	14.91	15.16	15.47	15.21	16.48	16.43
0.90 $\pi$	13.44	12.37	14.76	14.55	15.31	14.71	16.58	16.63	17.49	18.46
$\pi$	13.08	12.88	14.35	14.60	15.47	15.21	15.62	16.18	17.55	18.76

表 24. SID error rates (ML, 16-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.34	0.56	1.39							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.37	0.30	0.45	0.49	0.52	0.52	0.45	0.41	0.75	1.24
0.05 $\pi$	0.34	0.26	0.37	0.49	0.56	0.64	0.64	0.67	0.56	1.27
0.10 $\pi$	0.30	0.30	0.34	0.60	0.41	0.37	0.49	0.64	0.75	0.82
0.15 $\pi$	0.30	0.41	0.41	0.56	0.49	0.49	0.64	0.75	0.56	0.79
0.20 $\pi$	0.49	0.30	0.22	0.49	0.52	0.41	0.52	0.52	0.52	0.45
0.25 $\pi$	0.45	0.49	0.19	0.49	0.26	0.34	0.49	0.71	0.60	0.64
0.30 $\pi$	0.45	0.26	0.22	0.19	0.45	0.52	0.49	0.45	0.52	0.52
0.35 $\pi$	0.49	0.41	0.34	0.22	0.26	0.30	0.45	0.56	0.41	0.34
0.40 $\pi$	0.45	0.45	0.45	0.34	0.26	0.26	0.37	0.34	0.26	0.37
0.45 $\pi$	0.30	0.34	0.26	0.30	0.30	0.30	0.30	0.30	0.37	0.37
0.50 $\pi$	0.45	0.26	0.26	0.22	0.34	0.34	0.15	0.49	0.56	0.49
0.55 $\pi$	0.41	0.45	0.49	0.34	0.34	0.41	0.41	0.30	0.34	0.45
0.60 $\pi$	0.45	0.41	0.49	0.37	0.41	0.30	0.45	0.49	0.56	0.41
0.70 $\pi$	0.45	0.41	0.41	0.49	0.45	0.49	0.64	0.75	0.71	0.86
0.80 $\pi$	0.34	0.52	0.52	0.52	0.67	0.79	0.75	0.37	0.71	1.01
0.90 $\pi$	0.45	0.49	0.64	0.67	0.79	0.56	0.49	0.86	0.71	1.16
$\pi$	0.45	0.60	0.56	0.56	0.75	0.56	0.52	0.94	0.71	0.90

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	8.67	7.20	10.34							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	7.61	6.59	7.30	6.74	7.25	7.30	8.21	7.45	7.71	7.96
0.05 $\pi$	7.66	6.59	7.25	7.00	7.00	7.20	8.27	7.76	7.56	8.06
0.10 $\pi$	7.61	6.85	6.74	7.05	6.74	7.15	7.86	7.40	8.01	7.20
0.15 $\pi$	7.81	7.00	6.69	7.51	6.95	7.40	7.61	8.06	7.45	8.01
0.20 $\pi$	8.06	6.64	6.74	7.30	7.76	6.74	7.25	7.86	8.62	8.47
0.25 $\pi$	8.01	7.20	7.00	7.00	7.35	7.61	7.25	7.40	8.42	8.01
0.30 $\pi$	8.42	7.40	7.05	7.35	7.00	7.56	6.80	7.61	7.81	8.67
0.35 $\pi$	8.11	8.06	7.91	7.76	7.35	7.71	8.37	7.86	9.43	9.38
0.40 $\pi$	8.47	8.21	8.52	8.11	7.91	8.21	7.61	9.03	9.08	9.08
0.45 $\pi$	8.57	8.62	8.42	9.08	9.58	9.18	9.08	9.48	8.92	9.53
0.50 $\pi$	9.03	8.92	8.98	9.08	8.77	9.13	9.43	9.99	9.58	10.40
0.55 $\pi$	8.77	8.57	9.03	8.77	9.13	9.33	9.48	9.23	9.84	10.65
0.60 $\pi$	8.52	8.82	8.92	9.63	9.38	10.45	10.14	10.29	10.85	11.51
0.70 $\pi$	8.67	9.79	9.74	10.24	11.05	11.05	11.87	11.71	11.36	11.71
0.80 $\pi$	8.82	10.04	10.24	11.46	11.56	11.87	10.95	11.05	11.46	13.08
0.90 $\pi$	8.82	9.23	10.90	11.21	11.61	10.80	11.71	11.82	13.44	14.20
$\pi$	9.33	9.48	10.90	10.65	11.61	10.65	11.61	12.02	13.54	13.39



表 25. SID error rates (ML, 32-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.04		0.11		0.37					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.07	0	0.07	0.07	0.11	0.07	0.11	0.11	0.22	0.15
0.05 $\pi$	0.11	0.04	0.11	0.07	0.07	0.07	0.15	0.15	0.19	0.11
0.10 $\pi$	0.04	0.04	0.04	0.07	0.04	0.07	0.19	0.19	0.11	0.11
0.15 $\pi$	0.07	0.04	0	0.04	0.07	0.04	0.11	0.15	0.22	0.07
0.20 $\pi$	0.07	0	0	0.11	0.04	0.04	0.04	0.15	0.11	0.07
0.25 $\pi$	0.15	0.04	0.04	0.04	0.07	0.07	0.11	0.07	0.11	0.04
0.30 $\pi$	0.07	0.07	0.04	0.07	0.04	0.04	0.07	0.11	0.11	0
0.35 $\pi$	0.07	0	0.04	0.04	0	0.04	0.04	0.04	0.04	0.07
0.40 $\pi$	0.07	0.07	0.07	0	0.11	0.04	0.07	0	0.07	0.04
0.45 $\pi$	0.07	0.11	0.07	0.04	0	0	0.04	0.04	0	0.04
0.50 $\pi$	0.04	0.07	0.07	0.07	0.11	0.07	0.04	0	0	0
0.55 $\pi$	0.11	0.07	0.15	0.04	0.11	0.04	0.11	0.07	0.07	0.07
0.60 $\pi$	0.07	0.04	0.11	0.11	0.11	0.04	0.04	0.04	0.04	0.04
0.70 $\pi$	0.07	0.04	0.07	0.04	0.04	0	0.07	0	0.11	0.11
0.80 $\pi$	0.04	0.04	0.11	0.07	0.04	0.04	0.15	0.04	0.11	0.11
0.90 $\pi$	0.04	0.04	0.07	0.11	0.04	0.15	0.07	0.15	0.04	0.04
$\pi$	0.07	0.07	0.04	0.11	0.07	0.19	0	0.07	0.07	0.04

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	7.05		5.93		7.15					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	6.54	5.12	5.83	5.22	5.43	5.48	6.14	6.44	5.78	6.44
0.05 $\pi$	6.49	5.98	5.83	5.38	5.32	5.83	6.09	6.14	6.24	5.68
0.10 $\pi$	6.49	5.93	5.73	5.53	5.98	5.88	5.68	5.38	6.09	6.29
0.15 $\pi$	6.69	5.38	5.48	5.48	5.22	5.83	5.93	6.39	6.19	6.54
0.20 $\pi$	6.49	6.03	4.92	5.07	5.27	5.38	5.43	5.73	5.12	6.29
0.25 $\pi$	6.69	5.98	5.53	5.43	4.97	5.32	5.73	5.73	5.78	5.68
0.30 $\pi$	6.85	6.59	5.93	5.53	5.88	5.73	5.58	5.58	6.29	6.90
0.35 $\pi$	6.74	7.20	6.85	6.14	5.78	5.93	6.34	6.03	6.29	7.10
0.40 $\pi$	6.95	7.35	6.74	7.45	6.69	6.34	6.29	6.80	7.61	8.01
0.45 $\pi$	7.61	7.71	7.35	7.20	6.80	6.95	7.15	7.71	7.71	7.81
0.50 $\pi$	7.56	7.51	7.40	7.56	7.15	7.45	7.71	7.71	8.47	8.67
0.55 $\pi$	7.86	7.61	7.45	8.11	7.05	7.96	7.96	8.21	8.16	8.87
0.60 $\pi$	7.25	7.30	7.81	7.81	8.37	8.01	8.57	8.67	8.57	9.08
0.70 $\pi$	7.45	8.06	7.56	8.57	9.03	8.87	9.79	9.58	10.14	10.19
0.80 $\pi$	7.35	7.61	8.92	8.47	8.87	9.48	9.94	9.63	10.09	10.50
0.90 $\pi$	7.91	7.76	8.87	8.67	8.47	9.18	9.94	10.19	10.95	11.87
$\pi$	7.35	7.81	8.32	8.82	8.87	9.53	10.04	9.38	11.71	11.66

## 付録E GPD 学習によるシステムの識別誤り率

表 26. SID error rates (Sample-by-sample mode GPD, 4-mix., 13-females) (%)

(a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.55	0.47	0.71							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.71	0.40	0.40	0.40	0.63	0.55	0.79	0.71	0.71	0.40
0.05 $\pi$	0.71	0.47	0.40	0.40	0.63	0.63	0.71	0.32	0.40	0.32
0.10 $\pi$	0.79	0.32	0.40	0.40	0.63	0.95	0.71	0.71	0.47	0.55
0.15 $\pi$	0.71	0.32	0.16	0.47	0.40	0.87	0.71	0.71	0.32	0.55
0.20 $\pi$	0.79	0.63	0.32	0.40	0.47	0.55	0.47	0.55	0.32	0.32
0.25 $\pi$	0.87	0.55	0.24	0.24	0.32	0.32	0.40	0.32	0.40	0.40
0.30 $\pi$	0.79	0.71	0.40	0.24	0.32	0.16	0.32	0.47	0.47	0.55
0.35 $\pi$	0.87	0.87	0.71	0.55	0.47	0.24	0.40	0.40	0.32	0.47
0.40 $\pi$	0.55	0.71	0.71	0.79	0.40	0.55	0.55	0.55	0.47	0.71
0.45 $\pi$	0.47	0.63	0.47	0.55	0.79	0.40	0.47	0.40	0.71	0.32
0.50 $\pi$	0.71	0.55	0.55	0.40	0.32	0.47	0.55	0.47	0.63	0.71
0.55 $\pi$	0.55	0.47	0.63	0.63	0.71	0.71	0.79	0.79	0.71	0.71
0.60 $\pi$	0.55	0.47	0.71	0.79	0.87	1.27	0.63	0.95	0.63	0.55
0.70 $\pi$	0.55	0.79	0.95	0.63	0.87	0.47	1.03	0.79	1.27	1.11
0.80 $\pi$	0.63	0.95	0.71	0.71	0.71	0.95	0.79	1.03	0.87	1.03
0.90 $\pi$	0.71	0.95	0.63	0.63	1.03	1.19	1.27	1.03	1.03	0.87
$\pi$	0.87	0.95	0.71	0.79	0.87	1.19	1.19	1.19	1.11	0.87

(b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	16.63	13.31	15.90							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	17.78	14.55	16.11	14.24	13.93	14.76	14.45	14.03	12.79	14.14
0.05 $\pi$	17.57	14.76	15.90	13.83	13.83	14.45	15.49	14.55	13.93	13.41
0.10 $\pi$	17.67	15.07	16.22	13.51	14.66	15.59	16.42	14.45	13.51	15.59
0.15 $\pi$	17.78	14.66	14.66	14.03	14.55	13.93	13.72	15.70	15.38	17.15
0.20 $\pi$	17.88	15.28	13.83	15.70	13.62	12.89	14.66	14.45	15.59	18.61
0.25 $\pi$	17.36	15.90	13.62	14.86	14.86	14.14	12.68	12.58	13.62	15.70
0.30 $\pi$	17.67	18.19	14.97	14.45	15.28	14.35	14.76	12.79	13.62	14.03
0.35 $\pi$	17.78	17.67	16.74	15.07	15.07	13.83	15.28	14.14	13.72	15.07
0.40 $\pi$	16.94	18.09	17.67	17.67	16.84	16.42	15.38	16.11	17.98	19.54
0.45 $\pi$	17.15	17.46	17.98	16.63	16.94	15.80	16.74	17.15	16.94	18.09
0.50 $\pi$	17.15	16.84	17.46	16.74	16.74	16.63	16.84	17.67	17.46	19.13
0.55 $\pi$	17.26	17.36	17.36	17.98	17.46	17.36	17.88	19.75	18.30	19.23
0.60 $\pi$	17.57	18.09	17.57	17.57	17.36	18.92	19.65	19.02	20.17	20.89
0.70 $\pi$	17.57	17.46	17.98	19.75	21.41	20.89	19.23	20.79	23.08	24.01
0.80 $\pi$	16.42	17.57	20.06	20.58	20.27	19.75	21.93	23.28	22.77	23.60
0.90 $\pi$	16.74	17.46	19.33	20.79	18.40	20.69	21.93	22.25	21.62	23.70
$\pi$	17.26	17.57	19.85	21.00	17.88	20.48	21.62	22.35	20.79	23.18

表 27. SID error rates (Sample-by-sample mode GPD, 8-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0.16	0.16	0.16	0.16				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0.08	0.16	0.08	0.08	0.08	0.08	0.08	0.08	0.24
0.05 $\pi$	0	0.08	0.16	0.08	0.08	0.08	0.08	0.16	0.08	0.24
0.10 $\pi$	0	0.08	0.16	0.08	0.08	0.08	0.16	0.16	0.16	0.08
0.15 $\pi$	0	0.08	0.16	0.08	0.08	0.08	0.08	0.08	0.08	0.16
0.20 $\pi$	0	0	0.16	0.08	0.08	0.16	0.08	0.08	0.08	0.16
0.25 $\pi$	0	0	0.08	0.08	0.08	0.16	0.16	0.16	0.08	0
0.30 $\pi$	0	0	0	0.16	0.16	0.08	0.08	0.16	0.16	0
0.35 $\pi$	0	0	0	0	0.08	0.08	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0.16
0.45 $\pi$	0	0	0	0	0	0	0	0	0.08	0.08
0.50 $\pi$	0	0	0	0	0	0	0	0	0	0
0.55 $\pi$	0	0	0	0	0.08	0	0.08	0.08	0.16	0.08
0.60 $\pi$	0	0	0	0.08	0	0	0	0.08	0.08	0
0.70 $\pi$	0	0	0	0	0.08	0.08	0.16	0.08	0.08	0.32
0.80 $\pi$	0	0	0.08	0.08	0.08	0.08	0.08	0.16	0.08	0.24
0.90 $\pi$	0	0	0	0	0.08	0.08	0.08	0.08	0.16	0.24
$\pi$	0	0	0.08	0.08	0.08	0.08	0.16	0.08	0	0.08

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	11.85	9.67	12.58							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	10.60	8.63	10.39	9.04	9.56	10.50	10.29	11.12	11.12	11.02
0.05 $\pi$	10.60	8.73	10.50	8.63	9.56	10.91	10.08	10.39	10.19	10.50
0.10 $\pi$	10.39	8.73	11.12	8.73	10.29	11.23	9.67	10.91	9.36	10.39
0.15 $\pi$	11.12	9.15	10.29	9.77	9.36	9.36	10.50	10.29	10.81	11.02
0.20 $\pi$	10.91	9.15	9.67	9.98	8.73	8.94	9.98	9.88	11.54	12.37
0.25 $\pi$	11.64	9.36	8.21	11.23	11.02	8.52	9.88	9.04	9.77	11.33
0.30 $\pi$	11.75	10.60	9.46	8.94	10.29	9.36	10.29	9.46	10.08	10.81
0.35 $\pi$	11.85	11.12	9.98	10.19	8.73	9.04	9.36	10.60	11.02	12.79
0.40 $\pi$	11.64	11.75	12.16	11.85	11.43	11.02	10.81	10.81	11.33	12.58
0.45 $\pi$	12.16	11.75	12.58	12.58	12.27	11.95	11.54	12.68	12.37	11.43
0.50 $\pi$	11.85	12.27	12.47	12.37	11.54	12.58	13.72	13.31	13.83	14.86
0.55 $\pi$	11.85	12.37	12.79	11.23	11.75	12.58	13.20	13.41	15.59	14.86
0.60 $\pi$	12.37	12.37	11.75	12.06	12.99	13.72	13.93	14.76	13.51	13.93
0.70 $\pi$	11.95	12.06	13.20	13.62	12.79	14.76	14.86	15.80	15.70	15.28
0.80 $\pi$	12.27	12.47	13.20	13.72	13.93	16.63	15.28	14.24	14.66	16.11
0.90 $\pi$	11.43	12.47	12.68	14.45	16.22	14.55	15.07	13.83	14.55	14.24
$\pi$	11.85	12.06	13.41	14.86	15.59	14.55	14.45	14.03	15.28	14.55

表 28. SID error rates (Sample-by-sample mode GPD, 16-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0	0	0.16					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0	0	0	0	0	0.08
0.05 $\pi$	0	0	0	0	0	0	0	0	0.08	0
0.10 $\pi$	0	0	0	0	0	0	0	0	0	0.08
0.15 $\pi$	0	0	0	0	0	0	0	0	0	0
0.20 $\pi$	0	0	0	0	0	0	0	0	0	0
0.25 $\pi$	0	0	0	0	0	0	0	0	0	0
0.30 $\pi$	0	0	0	0	0	0	0	0	0.08	0
0.35 $\pi$	0	0	0	0	0	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0
0.45 $\pi$	0	0	0	0	0	0	0	0	0	0
0.50 $\pi$	0	0	0.08	0	0	0	0	0	0	0
0.55 $\pi$	0	0	0	0	0	0	0	0	0	0
0.60 $\pi$	0.08	0.08	0.08	0	0	0	0	0	0	0
0.70 $\pi$	0.08	0	0	0	0	0	0.08	0	0	0
0.80 $\pi$	0.08	0	0	0	0	0	0	0	0	0
0.90 $\pi$	0.08	0	0	0	0.08	0	0	0.08	0.08	0
$\pi$	0	0	0	0	0	0	0	0	0.08	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	9.36	6.96	11.02							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.42	6.55	6.76	7.59	7.80	7.69	7.90	7.80	7.28	8.32
0.05 $\pi$	8.63	6.03	6.86	7.38	7.80	7.80	7.90	7.59	6.76	7.69
0.10 $\pi$	8.52	7.28	6.76	7.69	7.07	7.07	8.00	7.17	8.11	6.55
0.15 $\pi$	8.52	6.44	5.61	7.38	6.24	7.48	7.28	7.59	7.38	8.11
0.20 $\pi$	8.84	6.44	6.24	6.86	7.38	7.48	8.32	7.38	7.69	8.73
0.25 $\pi$	8.94	6.76	6.34	6.34	6.55	6.44	5.51	6.55	8.00	8.00
0.30 $\pi$	8.73	7.80	6.34	7.07	6.76	6.65	7.17	5.61	7.28	8.84
0.35 $\pi$	9.36	8.94	8.63	7.80	7.48	7.90	7.59	7.07	8.11	8.52
0.40 $\pi$	10.19	9.04	9.04	9.15	9.36	8.00	7.38	8.94	9.77	10.19
0.45 $\pi$	9.25	9.67	9.67	9.46	10.08	10.19	9.56	9.56	9.77	10.08
0.50 $\pi$	9.98	9.77	9.98	9.56	9.88	10.19	10.50	10.71	11.12	10.91
0.55 $\pi$	9.25	9.04	8.84	9.36	9.56	10.71	10.91	11.12	11.12	10.50
0.60 $\pi$	9.77	9.56	9.15	9.67	10.81	10.91	10.81	11.23	11.85	11.64
0.70 $\pi$	9.56	10.29	10.08	10.39	11.54	12.27	12.79	11.75	12.37	12.16
0.80 $\pi$	9.04	10.08	10.29	10.81	10.91	12.16	11.12	11.64	11.54	12.06
0.90 $\pi$	9.15	9.36	10.81	11.54	12.99	11.64	11.23	12.47	11.23	12.37
$\pi$	9.77	9.04	10.08	10.71	12.37	12.27	10.60	10.60	11.33	11.54

表 29. SID error rates (Sample-by-sample mode GPD, 32-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0	0	0.08					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0	0	0	0	0	0
0.05 $\pi$	0	0	0	0	0	0	0	0	0	0
0.10 $\pi$	0	0	0	0	0	0	0	0	0	0
0.15 $\pi$	0	0	0	0	0	0	0	0	0	0
0.20 $\pi$	0	0	0	0	0	0	0	0	0	0
0.25 $\pi$	0	0	0	0	0	0	0	0	0	0
0.30 $\pi$	0	0	0	0	0	0	0	0	0	0
0.35 $\pi$	0	0	0	0	0	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0
0.45 $\pi$	0	0	0	0	0	0	0	0	0	0
0.50 $\pi$	0	0	0	0	0	0	0	0	0	0
0.55 $\pi$	0	0.08	0	0	0	0	0	0	0	0
0.60 $\pi$	0	0	0	0	0	0	0	0	0	0
0.70 $\pi$	0	0	0	0	0	0	0	0	0	0
0.80 $\pi$	0	0	0	0	0	0	0	0	0	0
0.90 $\pi$	0	0	0	0	0	0	0	0	0	0
$\pi$	0	0	0	0	0	0	0	0	0	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	8.32	6.34	8.21							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.21	5.30	6.24	5.93	5.82	6.03	6.13	6.44	6.03	6.44
0.05 $\pi$	7.90	6.03	6.24	6.24	5.93	6.24	7.07	6.55	7.17	5.93
0.10 $\pi$	7.90	6.24	6.03	6.03	6.24	6.44	6.03	6.96	6.55	6.34
0.15 $\pi$	7.80	6.13	6.55	6.34	6.13	6.24	5.82	7.07	6.96	7.17
0.20 $\pi$	8.11	6.24	5.41	6.24	5.41	5.93	6.13	6.03	5.30	6.44
0.25 $\pi$	8.00	6.86	6.34	6.34	5.51	5.61	5.20	5.51	5.82	6.34
0.30 $\pi$	8.63	7.38	5.93	6.13	7.28	7.17	5.41	5.41	6.65	7.90
0.35 $\pi$	8.73	8.52	7.28	6.96	6.13	6.86	6.13	6.34	6.55	7.17
0.40 $\pi$	8.21	8.73	8.11	8.32	8.11	7.07	6.96	7.07	7.69	8.52
0.45 $\pi$	8.63	8.73	9.15	9.04	8.42	8.32	8.52	8.11	8.94	8.63
0.50 $\pi$	8.63	8.73	8.52	7.80	8.42	7.90	8.32	9.04	9.56	9.98
0.55 $\pi$	9.04	8.94	9.25	8.94	7.59	8.73	8.11	9.36	9.77	9.46
0.60 $\pi$	8.94	8.21	8.84	8.94	8.84	8.63	9.77	9.88	10.08	10.39
0.70 $\pi$	8.52	9.25	8.32	8.42	9.88	10.50	10.71	10.91	10.91	11.23
0.80 $\pi$	8.21	8.32	8.52	9.04	9.67	10.29	10.50	9.88	9.88	10.29
0.90 $\pi$	9.36	7.80	8.52	9.15	8.73	9.98	10.08	9.56	10.81	10.50
$\pi$	8.73	7.69	8.42	10.08	9.67	9.98	10.08	9.46	10.71	12.37

表 30. SID error rates (Batch mode GPD, 4-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	1.74		1.50		2.06					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	2.06	1.11	1.34	1.03	0.95	1.50	1.50	1.58	1.58	1.82
0.05 $\pi$	1.98	1.03	1.19	1.19	0.95	1.90	1.42	1.19	1.90	1.98
0.10 $\pi$	2.14	1.11	1.19	1.03	1.27	1.74	1.50	1.42	1.27	1.34
0.15 $\pi$	2.14	1.34	1.11	0.95	0.87	1.34	1.42	1.27	0.95	1.27
0.20 $\pi$	2.06	1.50	0.95	1.11	1.03	0.71	1.34	1.34	1.03	1.34
0.25 $\pi$	1.82	1.74	1.34	0.95	1.11	1.11	0.71	0.79	0.79	1.11
0.30 $\pi$	1.50	1.74	1.50	0.87	1.03	0.87	0.71	1.19	1.27	1.19
0.35 $\pi$	1.34	2.22	1.74	1.50	0.87	0.87	0.71	0.87	0.79	1.27
0.40 $\pi$	1.58	1.42	1.90	1.82	1.58	1.11	1.42	1.74	1.58	2.14
0.45 $\pi$	1.50	1.50	1.42	1.42	1.27	1.19	1.50	1.66	1.34	1.66
0.50 $\pi$	1.74	1.58	1.58	1.66	1.82	1.27	1.34	1.58	1.50	1.66
0.55 $\pi$	1.66	1.82	1.42	1.82	1.74	1.82	1.74	2.22	2.22	2.53
0.60 $\pi$	1.82	1.50	1.74	1.74	1.90	2.45	2.37	2.61	2.77	2.06
0.70 $\pi$	1.58	1.42	1.74	2.37	1.90	2.37	2.37	2.69	2.53	2.45
0.80 $\pi$	1.74	1.58	1.42	1.98	2.69	2.53	3.09	3.24	2.53	2.06
0.90 $\pi$	1.74	1.90	2.06	2.61	2.45	2.77	2.85	2.53	1.90	1.90
$\pi$	1.58	1.58	2.14	2.53	2.37	2.93	2.85	2.69	2.22	2.77

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	17.78		13.93		16.32					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	17.36	14.86	15.70	14.76	14.35	14.24	14.14	13.41	13.51	14.66
0.05 $\pi$	17.57	14.86	16.01	14.03	14.35	15.49	14.45	13.62	15.07	13.51
0.10 $\pi$	17.26	15.18	16.11	13.62	15.07	15.49	16.32	14.14	14.45	14.97
0.15 $\pi$	17.57	14.35	15.18	15.07	14.45	14.66	14.55	14.14	15.18	14.76
0.20 $\pi$	17.88	15.28	14.97	15.59	14.45	13.51	14.24	14.35	15.80	18.19
0.25 $\pi$	17.36	16.01	15.28	14.97	15.07	13.83	13.41	13.10	12.89	16.11
0.30 $\pi$	17.88	18.09	14.55	14.55	15.38	14.76	15.18	12.89	14.03	15.49
0.35 $\pi$	17.46	17.67	17.26	15.18	14.76	14.86	15.38	14.66	13.83	15.80
0.40 $\pi$	17.78	17.57	17.98	18.61	17.78	15.80	15.18	15.38	17.57	17.98
0.45 $\pi$	17.98	17.67	17.78	17.88	17.46	16.74	15.80	17.15	15.90	17.26
0.50 $\pi$	17.46	17.15	17.26	16.84	16.32	17.36	17.05	17.15	17.57	19.54
0.55 $\pi$	17.46	17.67	17.98	17.78	16.94	17.57	17.57	19.75	18.82	18.40
0.60 $\pi$	17.46	17.57	17.46	18.19	18.61	19.54	18.71	20.17	20.27	20.17
0.70 $\pi$	17.78	17.36	17.78	20.06	20.17	20.27	20.17	20.06	23.49	23.91
0.80 $\pi$	17.46	17.98	20.37	20.48	19.54	17.98	20.48	22.04	21.62	24.01
0.90 $\pi$	17.46	18.19	19.75	19.23	17.88	20.06	21.83	21.41	22.25	23.80
$\pi$	17.36	17.67	20.17	18.71	18.19	19.44	21.21	22.25	21.73	21.93

表 31. SID error rates (Batch mode GPD, 8-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.08	0.32	0.32	0.16	0.24	0.16	0.32	0.40	0.40	0.47
0.05 $\pi$	0	0.32	0.32	0.16	0.24	0.24	0.24	0.32	0.32	0.47
0.10 $\pi$	0	0.32	0.24	0.16	0.16	0.32	0.24	0.24	0.32	0.47
0.15 $\pi$	0.08	0.24	0.24	0.24	0.24	0.24	0.32	0.32	0.47	0.32
0.20 $\pi$	0.08	0.08	0.24	0.32	0.32	0.16	0.40	0.24	0.32	0.32
0.25 $\pi$	0.08	0.08	0.24	0.16	0.24	0.40	0.40	0.32	0.24	0.47
0.30 $\pi$	0	0.08	0.08	0.32	0.24	0.24	0.32	0.32	0.40	0.24
0.35 $\pi$	0	0.08	0.08	0	0.16	0.16	0.16	0.24	0.16	0.16
0.40 $\pi$	0.08	0.08	0	0	0	0	0.16	0.08	0.16	0.32
0.45 $\pi$	0.08	0.08	0.08	0	0.08	0	0.08	0.08	0.16	0.16
0.50 $\pi$	0.08	0.08	0.08	0	0	0	0.08	0.24	0.24	0.32
0.55 $\pi$	0.08	0.08	0.16	0.24	0.24	0.24	0.32	0.16	0.16	0.32
0.60 $\pi$	0.16	0.24	0.08	0.16	0.24	0.32	0.24	0.32	0.32	0.08
0.70 $\pi$	0.24	0.08	0.08	0.24	0.16	0.32	0.24	0.24	0.47	0.47
0.80 $\pi$	0.08	0.08	0.16	0.08	0.16	0.16	0.32	0.47	0.40	0.32
0.90 $\pi$	0.08	0.08	0.16	0.08	0.32	0.32	0.40	0.40	0.24	0.63
$\pi$	0.08	0	0.16	0.08	0.40	0.32	0.55	0.24	0.40	0.95

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	12.27	10.59	13.51							
0.05 $\pi$	10.71	8.52	10.19	8.94	11.12	10.71	10.19	11.43	11.12	11.23
0.10 $\pi$	10.60	8.42	10.39	8.84	10.71	10.81	10.29	10.39	10.60	11.02
0.15 $\pi$	10.71	8.73	11.02	9.15	10.71	11.43	9.98	10.50	9.77	11.33
0.20 $\pi$	11.12	9.04	10.91	9.56	9.15	10.81	11.02	10.91	10.81	10.60
0.25 $\pi$	11.12	9.46	9.46	10.60	9.56	9.46	10.60	10.39	11.12	12.68
0.30 $\pi$	12.37	9.67	8.42	10.81	10.60	9.56	10.29	9.98	10.19	11.64
0.35 $\pi$	12.27	10.60	9.46	9.88	10.29	9.88	10.50	9.67	10.19	11.02
0.40 $\pi$	11.85	11.54	10.81	9.98	9.46	9.77	9.67	10.60	10.60	13.41
0.45 $\pi$	12.16	12.06	11.85	12.16	12.27	11.64	11.02	11.02	11.43	12.68
0.50 $\pi$	12.68	12.27	12.68	13.20	13.31	12.47	12.06	12.47	12.47	11.23
0.55 $\pi$	12.16	12.27	12.47	12.27	12.16	12.68	13.51	13.72	13.83	15.07
0.60 $\pi$	12.37	12.79	12.68	12.16	12.47	12.47	12.99	13.10	15.80	15.38
0.70 $\pi$	12.79	12.47	12.27	11.95	13.10	14.24	14.45	14.35	12.79	14.55
0.80 $\pi$	12.37	12.06	12.79	12.89	12.68	14.97	15.18	15.90	15.90	15.59
0.90 $\pi$	12.89	12.79	12.58	13.93	14.86	16.22	15.59	13.72	15.49	15.80
$\pi$	11.95	12.16	12.89	15.07	15.49	14.55	14.66	13.62	14.24	14.76
$\pi$	11.85	12.27	13.41	14.14	16.01	14.76	14.14	13.10	15.80	14.76

表32. SID error rates (Batch mode GPD, 16-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0.08	0.08	0.16	0.16				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0.08	0.16	0.08	0.16	0.08	0.16
0.05 $\pi$	0	0	0	0	0	0	0.08	0.16	0.08	0.08
0.10 $\pi$	0	0	0	0.08	0	0	0.16	0	0.24	0.16
0.15 $\pi$	0	0	0	0	0	0	0.16	0	0.08	0.16
0.20 $\pi$	0	0	0	0	0.08	0	0	0.08	0	0
0.25 $\pi$	0.08	0	0	0	0.16	0.16	0	0	0	0.08
0.30 $\pi$	0.16	0	0	0	0	0.16	0.16	0.16	0.16	0
0.35 $\pi$	0.16	0	0	0	0	0	0	0	0.08	0.08
0.40 $\pi$	0.08	0.08	0	0	0	0	0	0	0	0
0.45 $\pi$	0.08	0.16	0	0	0	0	0	0	0	0
0.50 $\pi$	0	0	0.08	0.08	0	0	0.16	0	0	0
0.55 $\pi$	0.08	0.08	0.16	0.16	0.08	0.24	0	0	0	0
0.60 $\pi$	0.08	0.08	0.08	0.08	0	0	0	0	0.16	0
0.70 $\pi$	0.08	0.08	0	0	0	0.16	0.16	0	0	0.16
0.80 $\pi$	0.08	0	0	0	0.24	0.16	0.08	0.08	0	0
0.90 $\pi$	0.08	0	0	0.08	0.08	0.16	0	0.08	0.08	0
$\pi$	0.08	0	0	0.16	0.08	0.08	0	0.16	0.08	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	9.88	7.48	11.64							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.94	6.96	6.55	7.69	7.69	7.90	8.00	8.21	7.28	8.63
0.05 $\pi$	8.94	6.65	7.28	8.00	8.63	7.69	8.11	8.32	7.28	7.90
0.10 $\pi$	8.73	6.86	6.65	8.42	7.17	7.38	8.63	8.32	8.11	7.17
0.15 $\pi$	8.73	6.76	5.93	7.80	6.34	7.90	7.07	8.84	7.90	8.11
0.20 $\pi$	8.84	6.55	7.07	7.28	7.69	7.69	7.90	7.48	8.73	8.84
0.25 $\pi$	8.63	6.55	6.86	6.96	7.48	7.69	6.34	7.17	8.42	8.32
0.30 $\pi$	9.04	7.80	6.86	7.59	7.17	7.07	7.38	6.24	7.59	9.46
0.35 $\pi$	9.46	9.15	8.42	8.42	7.38	7.59	8.00	7.80	8.11	9.15
0.40 $\pi$	10.29	9.15	8.94	9.67	9.15	8.32	7.90	8.94	9.98	10.39
0.45 $\pi$	9.77	10.08	10.08	10.19	10.91	10.60	9.77	10.39	9.67	9.88
0.50 $\pi$	9.98	10.39	10.29	10.50	10.19	10.29	10.91	10.81	10.91	11.23
0.55 $\pi$	10.08	9.36	9.36	9.88	9.98	10.50	11.12	11.12	10.81	11.64
0.60 $\pi$	9.98	9.77	9.46	10.50	10.60	11.02	11.54	11.33	12.58	11.85
0.70 $\pi$	9.36	10.39	10.81	10.71	11.43	13.20	12.89	12.89	12.47	12.37
0.80 $\pi$	9.67	9.98	10.91	11.33	12.06	12.58	12.06	11.54	11.12	10.91
0.90 $\pi$	9.77	9.88	10.50	11.02	12.58	12.58	11.75	12.06	11.54	12.16
$\pi$	10.39	10.39	10.50	10.71	12.99	12.27	10.91	10.81	12.47	12.37



表 33. SID error rates (Batch mode GPD, 32-mix., 13-females) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0	0	0.08	0				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0	0	0	0	0.08	0
$0.05\pi$	0	0	0	0	0	0	0	0	0	0
$0.10\pi$	0	0	0	0	0	0	0.16	0	0	0.08
$0.15\pi$	0	0	0	0	0	0	0	0	0	0
$0.20\pi$	0	0	0	0	0	0	0	0.08	0	0
$0.25\pi$	0	0	0	0	0	0	0	0	0	0
$0.30\pi$	0	0	0	0	0	0	0	0.08	0.08	0
$0.35\pi$	0	0	0	0	0	0	0	0	0	0
$0.40\pi$	0	0	0	0	0	0	0	0	0	0
$0.45\pi$	0.08	0	0.08	0	0	0	0	0	0	0
$0.50\pi$	0	0.08	0.08	0	0	0	0	0	0	0
$0.55\pi$	0.08	0.08	0	0	0	0	0	0	0	0
$0.60\pi$	0.08	0.08	0	0.08	0	0	0	0	0	0
$0.70\pi$	0.08	0	0	0	0	0	0	0	0	0
$0.80\pi$	0	0	0	0	0	0	0.08	0	0	0
$0.90\pi$	0	0	0	0.08	0	0	0	0	0	0
$\pi$	0	0	0	0	0	0.08	0	0	0	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	8.32	8.32	6.55	6.55	8.42	8.42				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.32	5.30	6.76	5.93	6.13	6.03	7.07	6.44	6.24	6.76
$0.05\pi$	8.00	6.13	6.76	6.24	5.93	6.86	7.48	6.76	7.59	6.24
$0.10\pi$	7.90	5.93	6.44	6.13	6.96	6.55	6.65	6.96	6.96	6.96
$0.15\pi$	7.69	6.24	6.24	6.44	5.93	6.24	6.24	7.69	7.59	7.69
$0.20\pi$	8.00	6.65	5.30	6.24	5.61	6.13	6.44	6.34	5.72	7.07
$0.25\pi$	8.11	6.76	6.24	6.44	5.72	5.93	5.30	5.41	6.44	6.44
$0.30\pi$	8.73	7.90	6.34	6.13	6.96	6.96	5.61	5.93	6.86	8.11
$0.35\pi$	8.63	8.63	8.00	7.38	6.44	7.28	6.34	6.34	6.55	7.17
$0.40\pi$	8.21	9.15	8.63	8.94	8.73	7.48	7.07	7.59	8.00	8.63
$0.45\pi$	8.73	9.25	9.88	9.25	8.32	8.84	8.84	8.94	8.73	9.04
$0.50\pi$	8.94	8.84	8.52	8.00	8.63	8.21	8.52	9.25	9.46	10.29
$0.55\pi$	8.94	9.15	9.56	9.25	7.80	8.84	8.32	9.56	9.77	9.36
$0.60\pi$	8.84	8.73	9.46	8.94	9.36	8.94	10.39	10.19	10.08	10.71
$0.70\pi$	8.94	9.46	8.42	8.84	10.29	10.50	11.02	11.12	11.33	11.23
$0.80\pi$	8.21	8.73	9.15	8.94	9.98	10.29	10.81	10.71	9.88	10.50
$0.90\pi$	9.67	8.11	8.94	9.36	8.94	10.08	10.50	9.98	10.81	11.23
$\pi$	9.15	8.32	8.84	10.29	9.46	10.08	9.98	9.56	11.33	12.27

表 34. SID error rates (Sample-by-sample mode GPD, 4-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.43		0.07		0.14					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0.14	0.07	0.07	0.07	0	0	0	0.14	0.14
0.05 $\pi$	0.07	0.21	0.07	0.07	0.07	0	0	0	0.14	0.07
0.10 $\pi$	0.07	0.14	0	0.07	0	0	0	0.07	0.14	0.07
0.15 $\pi$	0.07	0.21	0.14	0.14	0	0.07	0	0	0.07	0.21
0.20 $\pi$	0.14	0.14	0.28	0.14	0.07	0.07	0	0	0.14	0.14
0.25 $\pi$	0.07	0.07	0.21	0.07	0.21	0.14	0.07	0	0	0.14
0.30 $\pi$	0.07	0	0.21	0.14	0.07	0.21	0.07	0.07	0	0.21
0.35 $\pi$	0.07	0.14	0.14	0.28	0.14	0	0.14	0.21	0.14	0.07
0.40 $\pi$	0.07	0.07	0	0.21	0.21	0.21	0.07	0.07	0.14	0.14
0.45 $\pi$	0.21	0.07	0.07	0	0.07	0.14	0.14	0.07	0.14	0.21
0.50 $\pi$	0.43	0.28	0.21	0	0	0.07	0.07	0.14	0.07	0.14
0.55 $\pi$	0.36	0.28	0.07	0.07	0	0	0	0	0.14	0.14
0.60 $\pi$	0.28	0.21	0.14	0	0	0	0	0.21	0.14	0.14
0.70 $\pi$	0.21	0.14	0	0.07	0.14	0.21	0.36	0.36	0.50	0.50
0.80 $\pi$	0.14	0	0.07	0.14	0.28	0.43	0.64	0.36	0.50	0.57
0.90 $\pi$	0.14	0	0.14	0.28	0.28	0.36	0.36	0.43	0.28	0.43
$\pi$	0.14	0	0.14	0.28	0.36	0.28	0.50	0.43	0.57	0.43

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	9.41		7.72		9.41					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	9.70	8.51	9.41	8.61	8.51	7.92	8.71	7.82	8.61	8.71
0.05 $\pi$	9.70	8.22	9.51	8.81	8.32	8.22	8.81	7.82	7.82	8.71
0.10 $\pi$	9.60	8.12	8.22	8.22	8.71	7.62	9.31	8.12	7.82	9.21
0.15 $\pi$	9.01	8.42	8.91	8.02	8.32	8.32	8.02	9.41	9.31	8.91
0.20 $\pi$	9.11	9.01	8.22	9.21	7.72	9.21	7.72	9.01	9.80	8.32
0.25 $\pi$	9.01	9.51	8.51	8.91	9.21	9.80	9.51	9.01	8.51	8.71
0.30 $\pi$	9.51	9.51	8.81	9.01	9.60	9.51	9.51	10.00	9.41	10.30
0.35 $\pi$	9.70	9.80	10.10	9.90	8.71	9.51	11.19	10.79	11.29	11.78
0.40 $\pi$	10.20	9.90	9.21	9.31	10.49	10.20	10.99	10.89	11.88	11.29
0.45 $\pi$	9.70	10.10	10.30	9.70	9.51	9.90	10.20	11.29	11.58	12.57
0.50 $\pi$	9.90	9.90	10.40	10.89	11.29	11.29	11.78	11.29	12.67	11.68
0.55 $\pi$	10.59	10.49	10.79	11.29	10.89	10.89	11.98	10.89	11.68	13.37
0.60 $\pi$	10.59	11.29	11.49	11.39	12.57	12.67	12.87	15.05	15.25	15.25
0.70 $\pi$	11.58	10.49	10.99	13.37	13.66	13.07	13.66	14.55	16.44	17.72
0.80 $\pi$	11.49	10.69	13.37	13.27	13.76	14.36	16.93	16.93	18.02	18.91
0.90 $\pi$	10.59	11.19	13.56	12.08	14.55	14.65	17.23	18.22	17.72	21.19
$\pi$	10.99	11.68	13.76	13.07	14.36	16.04	16.63	17.03	20.20	21.29

表35. SID error rates (Sample-by-sample mode GPD, 8-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)			0.42 (Mel)			0.55 (Bark)			
0	0			0.07			0.14			
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.07	0	0	0	0	0	0	0	0.07	0.14
0.05 $\pi$	0.07	0	0	0	0	0	0	0	0	0.07
0.10 $\pi$	0.07	0	0	0	0	0	0	0	0.07	0.07
0.15 $\pi$	0.07	0	0	0	0	0	0	0	0	0
0.20 $\pi$	0	0	0	0	0	0	0	0	0.07	0
0.25 $\pi$	0	0	0	0	0	0	0	0	0	0.07
0.30 $\pi$	0	0	0	0	0	0	0	0	0	0
0.35 $\pi$	0	0	0	0.07	0.07	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0.07	0	0	0	0	0
0.45 $\pi$	0	0	0	0	0	0	0	0	0	0
0.50 $\pi$	0	0	0	0	0	0	0	0.07	0.07	0
0.55 $\pi$	0	0	0	0	0	0	0	0.07	0	0.07
0.60 $\pi$	0	0	0	0	0	0.07	0	0	0	0
0.70 $\pi$	0	0	0	0.07	0.07	0.07	0.07	0.07	0.14	0.21
0.80 $\pi$	0	0	0	0.07	0.07	0	0.07	0	0.07	0.21
0.90 $\pi$	0	0	0	0.07	0	0.07	0.07	0.07	0.07	0.14
$\pi$	0	0	0	0.14	0.14	0.07	0.14	0	0.07	0.07

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)			0.42 (Mel)			0.55 (Bark)			
0	8.02			6.73			6.14			
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	7.52	6.24	6.44	5.64	6.04	6.34	6.93	6.34	6.73	7.03
0.05 $\pi$	7.43	6.34	6.14	6.24	6.24	6.63	6.44	6.63	6.93	7.13
0.10 $\pi$	7.23	6.83	6.73	6.44	5.25	5.25	6.34	6.73	6.93	6.93
0.15 $\pi$	6.73	7.33	6.53	6.34	4.95	6.83	6.73	6.34	6.73	7.23
0.20 $\pi$	7.13	7.62	6.24	6.93	6.24	6.24	6.24	5.84	7.03	7.52
0.25 $\pi$	7.43	8.02	6.73	6.44	7.13	6.14	5.84	6.53	6.44	7.52
0.30 $\pi$	7.52	7.72	7.52	6.83	6.53	7.62	6.93	6.83	6.93	7.82
0.35 $\pi$	7.72	7.43	7.52	7.23	6.93	6.44	7.92	8.51	8.22	8.91
0.40 $\pi$	7.92	7.43	7.43	6.83	7.23	7.33	7.52	7.82	8.02	8.61
0.45 $\pi$	7.62	8.02	6.93	6.93	7.33	7.72	8.61	7.72	8.02	8.51
0.50 $\pi$	8.12	8.02	8.71	8.81	8.22	7.82	7.82	7.82	8.51	9.51
0.55 $\pi$	8.61	8.42	8.22	8.02	8.71	8.42	7.43	8.32	8.51	9.11
0.60 $\pi$	7.92	8.42	9.11	9.41	9.11	8.91	8.61	9.60	9.60	9.90
0.70 $\pi$	8.71	8.91	9.41	10.10	9.51	8.91	8.91	10.30	11.39	11.58
0.80 $\pi$	9.01	8.91	9.21	9.01	9.90	9.11	10.59	11.09	10.99	11.19
0.90 $\pi$	8.71	9.70	8.91	9.01	9.70	9.70	11.98	12.08	13.96	15.05
$\pi$	8.71	9.60	9.41	9.41	9.51	10.59	12.38	12.57	13.96	16.24

表 36. SID error rates (Sample-by-sample mode GPD, 16-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0	0	0	0				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0	0	0	0	0	0
0.05 $\pi$	0	0	0	0	0	0	0	0	0	0
0.10 $\pi$	0	0	0	0	0	0	0	0	0	0
0.15 $\pi$	0	0	0	0	0	0	0	0	0	0
0.20 $\pi$	0	0	0	0	0	0	0	0	0	0
0.25 $\pi$	0	0	0	0	0	0	0	0	0	0
0.30 $\pi$	0	0	0	0	0	0	0	0	0	0
0.35 $\pi$	0	0	0	0	0	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0
0.45 $\pi$	0	0	0	0	0	0	0	0	0	0
0.50 $\pi$	0	0	0	0	0	0	0	0	0	0
0.55 $\pi$	0	0	0	0	0	0	0.07	0	0	0
0.60 $\pi$	0	0	0	0	0	0	0	0	0	0.07
0.70 $\pi$	0	0	0	0	0	0	0	0	0.07	0.07
0.80 $\pi$	0	0	0	0	0	0	0	0	0	0
0.90 $\pi$	0	0	0	0	0	0	0	0	0	0
$\pi$	0	0	0	0	0	0	0	0	0	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	5.25	5.45	5.45	5.05	5.05	5.05				
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	5.54	4.55	5.15	4.55	4.95	5.25	6.34	5.15	4.55	4.65
0.05 $\pi$	5.35	4.85	5.15	4.75	4.75	4.65	6.63	5.64	5.15	4.55
0.10 $\pi$	5.54	5.15	5.05	4.55	4.26	5.25	6.04	5.25	5.64	4.26
0.15 $\pi$	5.15	5.15	5.05	4.46	5.15	4.75	5.74	6.14	4.46	4.85
0.20 $\pi$	5.74	4.85	4.85	5.15	4.75	4.75	5.05	5.84	5.54	5.45
0.25 $\pi$	5.54	5.05	5.35	4.65	5.35	5.45	6.14	5.64	6.04	5.84
0.30 $\pi$	5.64	5.45	5.45	4.95	4.95	5.45	5.35	6.14	5.84	5.94
0.35 $\pi$	5.74	5.35	5.35	5.74	5.64	5.84	7.03	7.23	8.42	8.42
0.40 $\pi$	5.84	5.64	5.84	6.14	5.05	6.04	6.24	7.72	7.33	7.03
0.45 $\pi$	6.04	6.14	5.15	6.24	6.34	6.14	7.03	6.44	6.73	7.72
0.50 $\pi$	5.54	5.45	6.04	6.63	6.73	6.63	5.54	6.93	6.63	8.22
0.55 $\pi$	5.64	5.94	6.34	6.14	6.34	6.63	6.53	7.33	7.03	7.92
0.60 $\pi$	6.24	6.24	6.63	6.83	6.93	7.62	6.83	7.43	7.82	8.12
0.70 $\pi$	6.24	6.63	7.13	7.52	8.22	7.52	8.51	8.02	8.61	9.41
0.80 $\pi$	6.14	7.72	7.23	7.92	8.32	8.42	7.92	8.81	9.41	10.89
0.90 $\pi$	6.44	7.33	8.02	7.52	8.12	7.92	9.41	9.90	11.39	11.88
$\pi$	6.14	7.13	7.82	7.33	8.42	7.52	10.69	10.20	12.87	12.18

表 37. SID error rates (Sample-by-sample mode GPD, 32-mix., 22-males) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)	
0	0	0	0	0	0	0
$0.05\pi$	0	0	0	0	0	0
$0.10\pi$	0	0	0	0	0	0
$0.15\pi$	0	0	0	0	0	0
$0.20\pi$	0	0	0	0	0	0
$0.25\pi$	0	0	0	0	0	0
$0.30\pi$	0	0	0	0	0	0
$0.35\pi$	0	0	0	0	0	0
$0.40\pi$	0	0	0	0	0	0
$0.45\pi$	0	0	0	0	0	0
$0.50\pi$	0	0	0	0	0	0
$0.55\pi$	0	0	0	0	0	0
$0.60\pi$	0	0	0	0	0	0
$0.70\pi$	0	0	0	0	0	0
$0.80\pi$	0	0	0	0	0	0
$0.90\pi$	0	0	0	0	0	0
$\pi$	0	0	0	0	0	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)	
0	5.55	4.46	4.36			
$0.05\pi$	4.95	5.25	4.75	4.26	4.16	4.06
$0.10\pi$	4.65	5.05	4.36	4.46	4.75	4.55
$0.15\pi$	4.75	4.46	4.16	4.55	4.26	4.95
$0.20\pi$	4.36	4.65	4.36	3.86	4.46	4.06
$0.25\pi$	4.26	4.46	4.65	4.26	3.96	4.26
$0.30\pi$	4.55	4.85	4.75	4.65	4.46	4.65
$0.35\pi$	4.36	5.35	4.95	4.55	4.75	4.06
$0.40\pi$	4.85	4.95	4.55	5.45	4.95	4.46
$0.45\pi$	5.74	5.25	4.85	4.75	4.85	4.85
$0.50\pi$	5.84	5.84	5.64	6.14	5.05	6.04
$0.55\pi$	5.64	5.35	5.15	6.04	5.35	5.64
$0.60\pi$	5.05	5.64	5.74	5.64	6.83	6.34
$0.70\pi$	5.45	5.84	6.44	7.33	7.33	6.83
$0.80\pi$	5.84	5.94	7.62	6.73	7.23	7.62
$0.90\pi$	5.35	6.24	6.63	6.73	7.03	8.22
$\pi$	5.25	6.63	6.73	6.53	7.13	7.62

表 38. SID error rates (Sample-by-sample mode GPD, 4-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0.19		0.22		0.34					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0.22	0.19	0.15	0.19	0.22	0.26	0.30	0.26	0.30	0.19
0.05 $\pi$	0.19	0.15	0.19	0.19	0.22	0.30	0.34	0.15	0.22	0.15
0.10 $\pi$	0.26	0.19	0.15	0.26	0.22	0.37	0.34	0.34	0.22	0.26
0.15 $\pi$	0.30	0.26	0.19	0.26	0.19	0.22	0.26	0.22	0.22	0.30
0.20 $\pi$	0.37	0.22	0.07	0.22	0.19	0.26	0.15	0.19	0.22	0.22
0.25 $\pi$	0.30	0.30	0.11	0.11	0.19	0.19	0.19	0.19	0.15	0.22
0.30 $\pi$	0.26	0.30	0.19	0.15	0.15	0.15	0.19	0.19	0.15	0.22
0.35 $\pi$	0.30	0.26	0.34	0.15	0.22	0.11	0.19	0.19	0.15	0.19
0.40 $\pi$	0.26	0.26	0.26	0.34	0.22	0.30	0.22	0.19	0.30	0.19
0.45 $\pi$	0.26	0.22	0.15	0.19	0.19	0.11	0.19	0.22	0.26	0.26
0.50 $\pi$	0.22	0.22	0.22	0.07	0.11	0.19	0.22	0.19	0.26	0.26
0.55 $\pi$	0.26	0.19	0.26	0.26	0.34	0.30	0.30	0.22	0.22	0.19
0.60 $\pi$	0.30	0.19	0.34	0.26	0.30	0.41	0.19	0.41	0.30	0.30
0.70 $\pi$	0.22	0.41	0.30	0.26	0.34	0.26	0.41	0.30	0.71	0.64
0.80 $\pi$	0.34	0.41	0.22	0.34	0.37	0.52	0.60	0.52	0.60	0.56
0.90 $\pi$	0.34	0.41	0.19	0.37	0.56	0.49	0.52	0.67	0.52	0.60
$\pi$	0.37	0.45	0.37	0.45	0.34	0.64	0.71	0.52	0.64	0.67

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	14.00		11.46		13.29					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	13.54	11.82	12.68	11.66	11.41	11.51	12.27	11.56	11.56	12.53
0.05 $\pi$	13.64	12.22	12.93	11.56	11.00	11.92	12.78	11.61	11.71	11.87
0.10 $\pi$	13.59	12.53	12.47	11.41	11.92	11.76	13.34	12.22	11.36	12.68
0.15 $\pi$	13.95	12.17	12.53	11.76	12.22	11.66	11.05	13.54	12.93	13.64
0.20 $\pi$	14.00	12.78	11.66	12.53	11.11	11.61	11.76	12.68	13.13	13.64
0.25 $\pi$	13.84	13.34	11.71	12.53	12.02	12.02	12.07	11.21	12.22	13.39
0.30 $\pi$	14.15	14.00	12.58	12.47	12.98	12.73	12.88	12.12	12.12	12.78
0.35 $\pi$	13.89	14.25	13.59	12.37	11.92	12.27	13.69	13.13	12.88	13.69
0.40 $\pi$	14.05	14.71	13.64	13.84	12.98	13.39	13.24	14.20	15.06	15.62
0.45 $\pi$	13.34	13.89	14.20	14.20	14.30	13.69	13.64	13.89	14.50	15.67
0.50 $\pi$	13.74	13.79	13.89	14.00	14.10	14.00	14.45	15.21	15.47	15.72
0.55 $\pi$	14.20	14.05	13.89	14.35	13.95	14.45	15.16	15.82	15.92	16.02
0.60 $\pi$	14.10	14.81	14.81	14.91	15.57	16.08	16.48	17.19	18.00	18.15
0.70 $\pi$	14.81	14.15	14.91	17.19	17.44	18.86	17.60	19.37	21.10	21.86
0.80 $\pi$	14.71	14.50	16.99	17.55	17.34	17.70	20.54	21.60	21.50	22.77
0.90 $\pi$	14.30	14.76	17.39	16.89	17.70	18.81	21.65	22.01	21.25	24.39
$\pi$	13.95	15.57	17.65	17.55	17.04	19.73	21.60	20.69	21.86	24.24

表 39. SID error rates (Sample-by-sample mode GPD, 8-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0		0.11		0.11					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0.04	0.04	0.04	0.07	0.07	0.07	0.04	0.07	0.19
0.05 $\pi$	0	0.04	0.04	0.04	0.07	0.07	0.07	0.04	0.04	0.07
0.10 $\pi$	0	0.04	0.07	0.04	0.04	0.07	0.07	0.07	0.07	0.07
0.15 $\pi$	0	0	0.07	0.04	0.04	0.07	0.04	0.07	0.04	0.04
0.20 $\pi$	0	0	0.07	0.04	0.04	0.07	0.07	0.07	0.07	0.04
0.25 $\pi$	0	0	0.04	0.04	0.04	0.04	0.04	0.07	0.04	0.07
0.30 $\pi$	0	0	0	0.04	0.07	0.04	0.04	0.04	0.04	0.04
0.35 $\pi$	0	0	0	0	0.07	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0.04
0.45 $\pi$	0	0	0	0	0	0	0	0	0.04	0.04
0.50 $\pi$	0	0	0	0	0	0	0	0.04	0.04	0
0.55 $\pi$	0	0	0	0	0	0	0	0	0	0.07
0.60 $\pi$	0	0	0	0	0	0.04	0	0	0.04	0
0.70 $\pi$	0	0	0	0.04	0.04	0.04	0.07	0.04	0.11	0.19
0.80 $\pi$	0	0	0	0	0.07	0.04	0.07	0	0.04	0.19
0.90 $\pi$	0	0	0.04	0.04	0.04	0.04	0.04	0.07	0.07	0.11
$\pi$	0	0	0	0.07	0.07	0.04	0.07	0.04	0	0.07

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	9.94		7.86		9.28					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	8.92	7.56	8.32	7.40	7.76	8.72	8.77	8.57	9.08	9.18
0.05 $\pi$	9.13	7.61	8.37	7.61	8.06	9.18	8.77	8.47	8.47	9.13
0.10 $\pi$	8.92	7.76	8.67	7.81	7.51	8.57	8.21	9.13	8.62	8.52
0.15 $\pi$	9.08	7.96	8.57	8.27	6.90	8.01	8.77	8.47	8.92	9.38
0.20 $\pi$	8.98	8.42	8.06	8.82	7.30	7.86	7.96	8.01	9.53	10.29
0.25 $\pi$	9.28	8.77	7.56	8.72	8.92	7.61	7.71	7.86	8.52	9.43
0.30 $\pi$	9.99	9.58	8.47	7.81	8.82	8.77	8.27	8.67	8.87	9.58
0.35 $\pi$	9.63	9.28	8.87	8.62	7.96	7.96	8.47	9.08	9.58	10.70
0.40 $\pi$	9.69	9.53	9.58	9.53	9.63	9.28	9.13	9.33	10.09	10.40
0.45 $\pi$	9.89	9.94	9.89	9.79	9.74	10.04	10.04	10.24	10.29	10.19
0.50 $\pi$	10.09	9.84	10.40	10.29	10.04	10.65	11.00	10.70	11.16	12.07
0.55 $\pi$	10.34	10.50	10.50	9.48	10.24	10.29	10.40	11.05	12.47	12.32
0.60 $\pi$	10.24	10.60	10.60	10.85	11.41	11.56	11.66	11.92	11.92	12.27
0.70 $\pi$	10.45	10.45	11.76	11.82	11.56	12.37	12.27	12.73	14.45	14.20
0.80 $\pi$	10.80	10.60	11.76	11.76	12.78	13.74	13.03	13.39	14.15	14.91
0.90 $\pi$	10.14	11.26	11.36	12.02	13.49	12.68	14.25	13.79	15.26	15.72
$\pi$	10.04	11.05	11.56	12.58	12.88	13.54	14.40	14.05	15.26	16.73

表 40. SID error rates (Sample-by-sample mode GPD, 16-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	0	0	0	0	0.04					
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	0	0	0	0	0	0	0	0	0	0.04
0.05 $\pi$	0	0	0	0	0	0	0	0	0	0
0.10 $\pi$	0	0	0	0	0	0	0	0	0	0.04
0.15 $\pi$	0	0	0	0	0	0	0	0	0	0
0.20 $\pi$	0	0	0	0	0	0	0	0	0	0
0.25 $\pi$	0	0	0	0	0	0	0	0	0	0
0.30 $\pi$	0	0	0	0	0	0	0	0	0	0
0.35 $\pi$	0	0	0	0	0	0	0	0	0	0
0.40 $\pi$	0	0	0	0	0	0	0	0	0	0
0.45 $\pi$	0	0	0	0	0	0	0	0	0	0
0.50 $\pi$	0	0	0	0	0	0	0	0	0	0
0.55 $\pi$	0	0	0	0	0	0	0	0	0	0
0.60 $\pi$	0	0	0	0	0	0	0	0	0	0
0.70 $\pi$	0	0	0	0	0	0	0	0	0	0.04
0.80 $\pi$	0.04	0	0	0	0	0	0	0	0	0
0.90 $\pi$	0	0	0	0	0.04	0	0	0	0.04	0
$\pi$	0	0	0	0	0	0	0	0	0	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)					
0	7.30	6.03	7.86							
$\theta \downarrow, \alpha \rightarrow$	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
0	7.20	5.43	6.03	5.88	6.19	6.39	7.10	6.14	5.83	6.29
0.05 $\pi$	7.10	5.32	6.03	5.78	6.24	6.09	7.15	6.44	5.73	6.59
0.10 $\pi$	7.10	5.88	5.83	6.03	5.63	6.19	6.95	6.29	6.54	5.48
0.15 $\pi$	7.15	5.83	5.22	5.83	5.83	6.24	6.19	6.64	5.83	6.64
0.20 $\pi$	7.20	5.58	5.53	6.34	5.88	6.14	6.85	6.49	6.54	7.30
0.25 $\pi$	7.30	5.83	6.03	5.63	5.93	5.78	5.73	5.88	7.30	6.74
0.30 $\pi$	7.40	6.59	5.78	5.98	5.98	6.09	6.19	5.68	6.44	7.35
0.35 $\pi$	7.76	7.30	6.80	6.69	6.59	6.95	7.15	7.00	8.21	8.21
0.40 $\pi$	8.27	7.40	7.51	7.76	6.95	6.90	6.85	8.01	8.37	8.57
0.45 $\pi$	7.56	7.81	7.40	7.91	8.11	8.11	8.16	7.86	8.32	8.82
0.50 $\pi$	7.51	7.35	8.21	8.16	8.37	8.37	8.32	8.98	8.87	9.74
0.55 $\pi$	7.61	7.51	7.51	7.71	7.91	8.92	8.72	8.87	9.23	9.28
0.60 $\pi$	7.81	7.76	7.91	8.52	9.03	9.43	9.03	9.38	10.14	9.94
0.70 $\pi$	7.76	8.42	8.77	8.98	9.74	10.19	10.95	10.09	10.40	10.95
0.80 $\pi$	7.61	8.87	8.72	9.48	9.33	10.24	9.79	10.14	10.70	11.56
0.90 $\pi$	7.76	8.21	9.38	9.53	10.29	9.63	10.29	11.11	12.07	13.03
$\pi$	7.86	8.06	8.87	9.03	10.60	9.99	10.70	11.16	12.37	12.53



表 41. SID error rates (Sample-by-sample mode GPD, 32-mix., 35-speakers) (%)

## (a) Training data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)	
0	0	0	0	0	0	0
$0.05\pi$	0	0	0	0	0	0
$0.10\pi$	0	0	0	0	0	0
$0.15\pi$	0	0	0	0	0	0
$0.20\pi$	0	0	0	0	0	0
$0.25\pi$	0	0	0	0	0	0
$0.30\pi$	0	0	0	0	0	0
$0.35\pi$	0	0	0	0	0	0
$0.40\pi$	0	0	0	0	0	0
$0.45\pi$	0	0	0	0	0	0
$0.50\pi$	0	0	0	0	0	0
$0.55\pi$	0	0	0	0	0	0
$0.60\pi$	0	0	0	0	0	0
$0.70\pi$	0	0	0	0	0	0
$0.80\pi$	0	0	0	0	0	0
$0.90\pi$	0	0	0	0	0	0
$\pi$	0	0	0	0	0	0

## (b) Test data

$\theta \downarrow, \alpha \rightarrow$	0 (Linear)		0.42 (Mel)		0.55 (Bark)	
0	6.90	5.17	5.98			
$0.05\pi$	6.34	4.97	5.38	5.27	4.92	5.63
$0.10\pi$	6.24	5.58	5.27	5.07	5.48	5.48
$0.15\pi$	6.24	5.27	5.22	5.32	5.17	5.43
$0.20\pi$	6.14	5.32	4.87	5.12	4.97	5.02
$0.25\pi$	6.03	5.63	5.58	5.32	4.67	4.92
$0.30\pi$	6.54	5.83	5.27	5.32	5.73	5.83
$0.35\pi$	6.54	6.90	5.98	5.63	5.53	5.63
$0.40\pi$	6.54	6.85	6.34	6.74	6.49	5.78
$0.45\pi$	7.15	6.85	7.00	6.90	6.74	6.64
$0.50\pi$	7.10	7.30	7.00	6.95	6.80	7.00
$0.55\pi$	7.10	7.05	7.20	7.51	6.64	7.30
$0.60\pi$	6.85	6.90	7.30	7.15	8.06	7.51
$0.70\pi$	6.95	7.61	7.45	7.81	8.62	8.47
$0.80\pi$	7.00	7.05	8.01	7.81	8.62	8.92
$0.90\pi$	7.10	7.10	7.61	7.86	8.06	9.08
$\pi$	6.69	7.30	7.71	8.27	8.21	9.13