

隠れマルコフモデルを用いた最ゆう推定に基づく誤り訂正

米崎 正<sup>†</sup>      吉田 幸司<sup>†</sup>      八木 敏男<sup>†</sup>      鹿野 清宏<sup>††</sup>

An Error Correction Approach Based on the Maximum Likelihood Estimation Combined with Hidden Markov Models

Tadashi YONEZAKI<sup>†</sup>, Koji YOSHIDA<sup>†</sup>, Toshio YAGI<sup>†</sup>, and Kiyohiro SHIKANO<sup>††</sup>

あらまし 隠れマルコフモデル (HMM) を利用した通信路誤り訂正法を提案する。提案法は受信符号列による制約の下、情報源モデル (HMM) と通信路モデルから算出される二つの確率を用いて、送信符号列を最ゆう推定するものである。したがって、誤り訂正符号を付加する必要はなく、既存の符号化方式に対して提案法を復号器の前処理に用いることで誤り耐性を向上させることができる。LSP パラメータを量子化した符号列を用いて、提案法に基づく誤り訂正によるケプストラムひずみの低減効果を調べた。その結果、3% ランダム誤りを付加した符号列に対して、提案法は約 39% ひずみを低減 (16 状態 HMM) することが確かめられた。

キーワード 誤り訂正, 隠れマルコフモデル, 情報源モデル, 通信路モデル, 最ゆう推定

1. ま え が き

音声符号化方式の研究では情報源符号化と通信路符号化を分離し、それぞれの最適化を図る手法が一般的である。一方、情報源-通信路統合符号化についても様々な検討がされている。これは、通信路での誤りを考慮した系で冗長性を除去し最適な情報源符号化を実現する手法と、情報源符号化で除去することができない冗長性を利用して誤り訂正能力を向上させる手法に大別できる。

前者は、更に、二つの手法に分類することができる。

一つめが、ベクトル量子化器を設計するときに通信路誤りを考慮する手法である。これに関して、熊沢らは通信路誤りを考慮したひずみ尺度を用いて符号帳を生成することを提案した [1]。一方、相澤らはベクトル量子化の符号語に対して誤りの影響が小さくなるような通信路符号を割り当てることを提案した [2]。更に Farvardin は多レベルブロック符号によって一つの符号内のビット間で誤り率が異なる場合の符号割り当てについて論じている [3]。

二つめが、通信路誤りに強い符号化系を情報源符号化に適用する手法である。誤り訂正で用いられるトレリス符号を情報源符号化へ用いた情報源-通信路統合符号化はその一つであり、Dunham らはこの符号化方式がひずみ-伝送速度関数に十分近い性能を有することを示した [4]。これを受けて、トレリス符号化を利用した様々な音声符号化方式が検討された。波形符号化への適用 [5], [6], CELP 符号化方式の音源への適用 [7], [8], LSP の量子化への適用 [9], [10] 等である。

これらに対して情報源符号化で除去することができない冗長性を積極的に利用する手法として、埴らは音声のスペクトルパラメータを符号語間の遷移を制約した符号帳に基づいて符号化することで誤り耐性を向上させた情報源符号化方式を提案した [11]。この符号語間遷移の制約は、符号語の遷移に偏りがあること、つまり、符号化における冗長性を利用することで、符号化利得を損なうことなく実現している。これは符号化系の構築時に情報源の動的な統計的性質を用いる手法であるが、近年、符復号化処理において情報源の動的な統計的性質を利用する手法が提案されつつある。

Fingscheidt らは通信路から得られる受信符号の確からしさを入力として量子化値の統計モデル (マルコフモデル) を駆動、復号することで誤りの影響を低減することができることを示した [12]。これは通信路の誤りに対して、誤りを訂正するのではなく隠ぺいする

<sup>†</sup> 松下通信工業株式会社テレコム研究所, 横須賀市  
Telecom Research Laboratory, Matsushita Communication Industrial CO., Ltd., Yokosuka-shi, 239-0847 Japan

<sup>††</sup> 奈良先端科学技術大学院大学情報科学研究科, 生駒市  
Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma-shi, 630-0101 Japan

ことを目的とした復号法である。一方、誤り訂正を目的とした手法として、AlajajiらはマルコフモデルとFECを組み合せ手法を提案した[13]。この中で、スカラ量子化したLSPパラメータをモデル化した2次マルコフモデルを軟判定復号と組み合わせることで最高5dBの符号化利得があることを報告している。更に、Garcia-Friasらは、隠れマルコフモデル(HMM)とターボ符号[14]を組み合せた手法を提案した[15]。ターボ符号は畳込み符号をベースにしており、畳込み符号に対応する状態とHMM状態との積空間に広がるトレリスに対して最ゆう復号することで誤り訂正能力を向上させることができる。Garcia-Friasらは、この手法を既知の隠れマルコフ情報源が出力する2値符号列に対して用いることで最適な復号化系が実現できることを[15]に示した。

本論文は、このように情報源の動的な統計的性質を利用して誤り訂正する一手法を提案する。本手法は[15]と同様にHMMを情報源モデルとして送信符号列を最ゆう推定するものである。しかしながら、音声符号化で用いられるLSPパラメータを符号化した多値符号列を情報源とする場合、情報源の出力を2値とした[15]のように積空間を探索することは処理量の点から不可能である。そこで、本論文ではHMMが最ゆう推定を効率良く実現するために必要となる“フォワード-バックワード特性”を有することを利用して、実現可能な誤り訂正アルゴリズムを提案する。

情報源の動的な統計的性質を利用して誤り訂正する手法として、量子化に用いる符号帳に対し全符号語間の遷移確率をあらかじめ求めておき誤り訂正に利用する手法も考えることができる。これは符号語数と同じ状態数のマルコフモデルを情報源モデルとして誤り訂正する手法といえる。この手法では符号語間の遷移確率を蓄積するために符号語数の2乗ワードのメモリが必要となり、符号語数の増加に伴い必要となるメモリ量は指数的に増大する。一方、HMMを情報源モデルとすることで必要となるメモリ量は状態数に対応して増加することとなりマルコフモデルを用いる場合と比較して少なくなる。更に、符号列に対応して一意に状態遷移するマルコフモデルと比較し状態間の遷移が確率的な重み付けの遷移として表現されるHMMは話者によって異なる送信符号列の時間変動に対してロバストなモデルといえるので提案法ではHMMを用いることとした。

提案法は符号間の冗長性を利用して誤り訂正する手

法であり、誤り訂正符号を付加することなく誤り耐性を向上させることができる。また、提案法は符号化系を再構築する必要がなく、復号化の前処理として用いることで既存の符号化方式の通信路誤り耐性を向上させることができる。

## 2. 最ゆう推定に基づく誤り訂正

符号列の統計的確率モデル  $\mathbf{M}$  が与えられたとき、モデル  $\mathbf{M}$  が符号列  $\hat{\mathbf{x}} = \{\hat{x}_n | n = 1, \dots, N\}$  を出力する確率  $P(\hat{\mathbf{x}}|\mathbf{M})$  を算出することができる。一方、符号列  $\mathbf{y} = \{y_n | n = 1, \dots, N\}$  を受信したとき、送信された符号列が  $\hat{\mathbf{x}}$  である確率  $P(\hat{\mathbf{x}}|\mathbf{y})$  は通信路モデルを定義することで算出することができる。本論文ではこの二つの確率を  $\gamma$  の重み付けでかけ合せた式(1)を符号列  $\hat{\mathbf{x}}$  のゆう度とし、送信符号列を最ゆう推定する手法を提案する。

$$P(\hat{\mathbf{x}}) = P(\hat{\mathbf{x}}|\mathbf{M})^\gamma \cdot P(\hat{\mathbf{x}}|\mathbf{y})^{(1-\gamma)} \quad (1)$$

図1に提案する誤り訂正法のブロック図を示す。

### 2.1 符号列モデルに基づく確率

符号列モデルとしてHMMを用いることで、 $P(\hat{\mathbf{x}}|\mathbf{M})$  はフォワードパス確率分布、及び、バックワードパス確率分布を用いて効率良く算出することができる。 $n$ におけるフォワードパス確率分布、バックワードパス確率分布を、それぞれ、 $\{\alpha(q, n) | q = 1, \dots, N_s\}$ 、 $\{\beta(q, n) | q = 1, \dots, N_s\}$ と記すこととする。ここで、 $N_s$ はHMMの状態数である。このとき、状態  $s$  から  $q$ へ遷移する確率を  $a_{sq}$ 、その遷移において符号  $x_n$  を出力する確率を  $b_{sq}(x_n)$  とすると、 $\alpha(q, n)$ 、 $\beta(q, n)$  は次の再帰式で定義される。

$$\alpha(q, n) = \sum_{s=1}^{N_s} \alpha(s, n-1) a_{sq} b_{sq}(x_n) \quad (2)$$

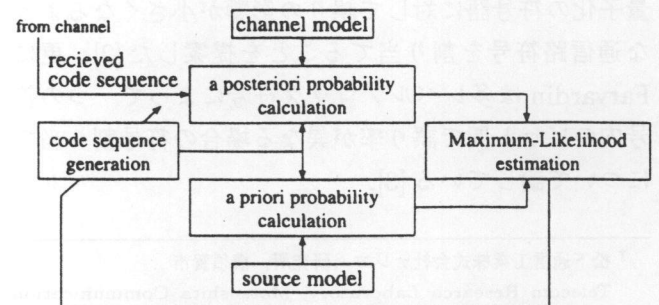


図1 最ゆう推定に基づく誤り訂正法ブロック図  
Fig. 1 Block diagram of the proposed error correction algorithm based on Maximum-Likelihood estimation.



$$\beta(s, n) = \sum_{q=1}^{N_s} \beta(q, n+1) a_{sq} b_{sq}(x_{n+1}) \quad (3)$$

初期値は、それぞれ、 $\{\alpha(q, 0) = \pi_q \mid q = 1, \dots, N_s\}$ ,  $\{\beta(q, N) = 1 \mid q \text{は最終状態}\}$  とする。ただし、 $\sum_1^{N_s} \pi_q = 1$  である。

図 2 に 4 状態のエルゴディック HMM によるフォワードパス、バックワードパス確率分布の遷移例を示す。(a) に HMM 状態遷移をトレリス線図により示した。○ は初期状態、● は最終状態を表している。(b) は状態間の遷移により HMM が出力する符号列である。(c) はフォワードパス確率分布、及び、バックワードパス確率分布の遷移を示しており、出力される符号に対応して式 (2), (3) より算出される。

符号列  $\hat{x}$  がモデル  $M$  より出力される確率  $P(\hat{x}|M)$  は任意の  $n$  におけるフォワードパス確率  $\alpha(q, n)$  とバックワードパス確率  $\beta(q, n)$  を用いて次式で表される。

$$\begin{aligned} P(\hat{x}|M) &= \sum_{q=1}^{N_s} \alpha(q, n) \cdot \beta(q, n) \\ &= \sum_{q=1}^{N_s} \sum_{s=1}^{N_s} \alpha(s, n-1) a_{sq} b_{sq}(\hat{x}_n) \cdot \beta(q, n) \end{aligned} \quad (4)$$

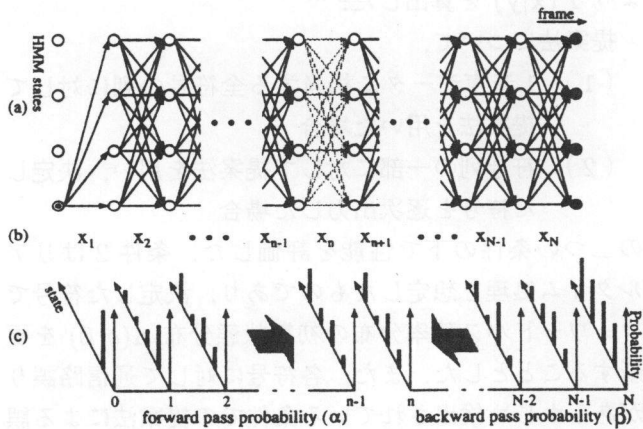


図 2 フォワード-バックワードパス確率分布の遷移例：(a) 4 状態エルゴディック HMM のトレリス線図、○ 初期状態、● 最終状態、(b) 出力符号列、(c) 確率分布の遷移

Fig. 2 Example of a transition of forward-backward pass probability distribution : (a) the trellis diagram of 4 states ergodic HMM, ○ indicates an initial state, ● indicates final states, (b) an output code sequence, (c) the transition of the probability distribution.

このように HMM では状態  $s$  から  $q$  へ遷移する確率  $a_{sq}$  と状態  $s$  から  $q$  への遷移で符号  $x$  を出力する確率  $b_{sq}(x)$  をパラメータとして情報源をモデル化し符号列  $\hat{x}$  が出力される確率を算出する。したがって、モデルの学習は学習データに対して式 (4) が最大となるようにパラメータを再帰的に更新していく [16]。本提案で用いたエルゴディック HMM [17] はすべての状態間を遷移することができるモデルであるため、全状態間の遷移確率と出力確率を学習により求め符号列をモデル化する。

次に式 (4) における符号列  $\hat{x}$  のうち符号  $\hat{x}_n$  のみを他の符号  $\bar{x}_n$  へ置換した符号列を  $\bar{x}$  とする。このとき、モデルが符号列  $\bar{x}$  を出力する確率  $P(\bar{x}|M)$  は、符号列  $\hat{x}$  に対して  $\{\alpha(q, n), \beta(q, n) \mid q = 1, \dots, N_s, n = 1, \dots, N\}$  があらかじめ算出されていれば、式 (4) の  $b_{sq}(\hat{x}_n)$  を  $b_{sq}(\bar{x}_n)$  に置換することで容易に算出することがわかる。

ところで、 $N$  が大きくなると式 (2), (3) の演算でけた落ちを防ぐために必要とされる精度が増大する。そこで、次式で示す正規化した  $\bar{\alpha}(q, n)$ ,  $\bar{\beta}(q, n)$  を  $P(\hat{x}|M)$  の算出に用いる。

$$\begin{aligned} \bar{\alpha}(q, n) &= \frac{1}{S_\alpha^{(n)}} \alpha(q, n) \\ \bar{\beta}(q, n) &= \frac{1}{S_\beta^{(n)}} \beta(q, n) \end{aligned} \quad (5)$$

ただし、

$$\begin{aligned} S_\alpha^{(n)} &= \sum_{q=1}^{N_s} \alpha(q, n) \\ S_\beta^{(n)} &= \sum_{q=1}^{N_s} \beta(q, n) \end{aligned}$$

このとき式 (4) は次式となる。

$$\begin{aligned} P(\hat{x}|M) &= \prod_{i=1}^n S_\alpha^{(i)} \cdot \left\{ \sum_{q=1}^{N_s} \bar{\alpha}(q, n) \cdot \bar{\beta}(q, n) \right\} \cdot \prod_{j=n}^N S_\beta^{(j)} \\ &= \prod_{i=1}^{n-1} S_\alpha^{(i)} \cdot \prod_{j=n}^N S_\beta^{(j)} \\ &\quad \times \sum_{q=1}^{N_s} \sum_{s=1}^{N_s} \bar{\alpha}(s, n-1) a_{sq} b_{sq}(\hat{x}_n) \cdot \bar{\beta}(q, n) \end{aligned} \quad (6)$$

## 2.2 通信路モデルに基づく確率

符号列  $\mathbf{y}$  を受信したときに送信符号列が  $\hat{\mathbf{x}}$  である確率  $P(\hat{\mathbf{x}}|\mathbf{y})$  は変復調方式や通信路のモデルを定義することで算出することができる。通信路モデルとして、記憶のない2元対称通信路を仮定すると  $P(\hat{\mathbf{x}}|\mathbf{y})$  は次式となる。

$$\begin{aligned} P(\hat{\mathbf{x}}|\mathbf{y}) &= \prod_{n=1}^N P(\hat{x}_n|y_n) \\ &= \prod_{n=1}^N r^{D_h^{(n)}} (1-r)^{(b-D_h^{(n)})} \end{aligned} \quad (7)$$

ただし、 $D_h^{(n)}$  は  $\hat{x}_n$  と  $y_n$  のハミング距離、 $b$  は1符号を表すビット数である。

$D_h^{(n)}$  は  $n$  について独立なので、 $P(\hat{\mathbf{x}}|\mathbf{M})$  と同様に、任意の  $\hat{x}_n$  の置換に対して  $P(\hat{\mathbf{x}}|\mathbf{y})$  は容易に算出することができる。

## 2.3 最ゆう推定

式(1)のゆう度を最大にする  $\hat{\mathbf{x}}$  は再帰的アルゴリズムにより求める。ここで、 $\hat{\mathbf{x}}^{(k)}$  を  $k$  回目の繰返しにより推定される符号列とし、符号列  $\hat{\mathbf{x}}^{(k-1)}$  と  $\hat{\mathbf{x}}^{(k)}$  はあるひとつの符号  $\hat{x}_n^{(k)}$  のみが異なることとする。したがって、1回の繰返しでは元の符号列に対して1符号のみを置換したときに最大のゆう度を与える符号列を求めることとなる。更に、本論文では式(1)のゆう度を直接求めるのではなく、 $\frac{P(\hat{\mathbf{x}}^{(k)})}{P(\hat{\mathbf{x}}^{(k-1)})}$  を求め、 $\frac{P(\hat{\mathbf{x}}^{(k)})}{P(\hat{\mathbf{x}}^{(k-1)})} \leq 1$  をアルゴリズムの終了条件とした。これは  $k-1$  回目の置換によりゆう度が極大値を取ることを意味しており、 $\hat{\mathbf{x}}^{(k-1)}$  を本アルゴリズムにより誤り訂正された符号列として出力することとなる。このように、提案法では、非現実的な処理量を要する全探索ではなく、再帰的アルゴリズムに基づいて符号列を推定するため極大値の検出をアルゴリズムの終了条件とした。 $\frac{P(\hat{\mathbf{x}}^{(k)})}{P(\hat{\mathbf{x}}^{(k-1)})}$  は式(8)により任意の  $\hat{\mathbf{x}}^{(k)}$  に対して算出することができる。式(8)において、 $\bar{\alpha}^{(k)}$ 、 $\bar{\beta}^{(k)}$  は  $\hat{\mathbf{x}}^{(k)}$  に対するフォワードパス確率、及び、バックワードパス確率である。また、 $\hat{\mathbf{x}}^{(0)} = \mathbf{y}$  とする。

$$\begin{aligned} &\frac{P(\hat{\mathbf{x}}^{(k)})}{P(\hat{\mathbf{x}}^{(k-1)})} \\ &= \frac{P(\hat{\mathbf{x}}^{(k)}|\mathbf{M})^\gamma P(\hat{\mathbf{x}}^{(k)}|\mathbf{y})^{1-\gamma}}{P(\hat{\mathbf{x}}^{(k-1)}|\mathbf{M})^\gamma P(\hat{\mathbf{x}}^{(k-1)}|\mathbf{y})^{1-\gamma}} \end{aligned}$$

$$\begin{aligned} &= \left( \frac{\sum_{s=1}^{N_s} \sum_{q=1}^{N_s} \bar{\alpha}^{(k-1)}(s, n-1) a_{sq} b_{sq}(\hat{x}_n^{(k)})}{\bar{\beta}^{(k-1)}(q, n)} \right)^\gamma \\ &\quad \cdot \left( \frac{P(\hat{x}_n^{(k)}|y_n)}{P(\hat{x}_n^{(k-1)}|y_n)} \right)^{(1-\gamma)} \end{aligned} \quad (8)$$

本アルゴリズムでは  $\hat{\mathbf{x}}^{(0)}$  が最大のゆう度を与えるとき通信路誤りを検出しなかったといえる。

## 3. 評価実験

### 3.1 実験条件

評価には10次のLSPパラメータを8bit符号帳で符号化した符号列を用いた。分析にはハミング窓を用い分析長を25ms、分析周期を20msとした。音声データは音響学会音声データベースを8kHzへダウンサンプリングしたものを使用し、HMMの学習用として1200発話(男性6名、女性6名)、評価用として学習内100発話(男性6名、女性6名)、学習外100発話(男性14名、女性18名)を用いた。HMMはエルゴディックHMMを用い、得られた符号列に対してビット誤り率0.1~3%の誤りをランダムに付加した符号列を入力として誤り訂正を行った。通信路は2元対称通信路を仮定し付加した誤り率  $r$  を既知とし式(7)より  $P(\hat{\mathbf{x}}|\mathbf{y})$  を算出した。

提案法について、

- (1) 1音声データに相当する全符号の列に対して提案法を用いた場合
- (2) 符号列の一部に対して提案法を用い、決定した符号を逐次出力した場合

の二つの条件の下で性能を評価した。条件2はリアルタイム処理を想定したものであり、決定した符号でフォワードパス確率分布の初期状態分布  $\alpha(q, 0)$  を更新することとした。また、各符号に対して通信路誤りがあらかじめ検出されている条件での提案法による誤り訂正の性能についても調べた。この場合、アルゴリズムの終了条件に  $k > 0$  を加え、更に、誤りが検出されていない符号は置換せず、検出されている符号に対してのみ符号を置換し誤り訂正することとした。

評価尺度として、誤りなしの符号列より得られる量子化パラメータからのケプストラム距離を用い、平均ひずみと2~4dB及び4dBを越えるひずみが全符号

数に占める割合 (Outlier) を測定した。

### 3.2 符号列/通信路モデルに対する重み付け

本節では 3.4, 3.5 の評価で用いる式 (1) の重み付け  $\gamma$  の値を決定する。図 3 に符号列モデルとして 8 状態 HMM を用いたときの  $\gamma$  の値と平均ひずみの関係を示す。通信路モデルを 2 元対称通信路としたので  $\gamma = 0.0$  が誤り訂正を行わず受信符号を用いて復号したときの通信路誤りによるひずみを示している。また、 $\gamma = 1.0$  は符号列モデルのみを用いて誤り訂正した結果であり、過度に符号列モデルに依存することでひずみが大きくなってしまふことを示している。これは誤っていない符号を間違えて '誤り' と検出したために生じたもので、統計モデルに基づく誤り検出を行う上で符号列を完全にモデル化できないことに起因する。図 3 より  $P(\hat{x}|M)$  に対する重み付け  $\gamma$  として 0.5 を用いることとする。

### 3.3 HMM 状態数と誤り訂正能力の関係

本節では、HMM の状態数と誤り訂正能力の関係を調べる。図 4 に 3% 誤りに対して提案法の HMM 状態数を 2~64 と変化させたときの平均ひずみを示す。図より、状態数を増加させるに従いひずみが減少し誤り訂正能力が向上していることがわかる。しかしながら、状態数を増加させるとともに、状態数に対するひずみ低減効率は劣化している。

ところで、本論文ではエルゴディック HMM を用いているため全状態間に対して遷移確率と出力確率を蓄積する必要がある。つまり、状態数の 2 乗に比例して要するメモリ量が増大する。また、ゆう度算出におい

ても全状態間の遷移確率を利用して符号の出力確率を算出する必要があるため、状態数が増加するにつれ処理量も増大する。

以上より、実現性を考慮して以下の実験では 2~16 状態 HMM を用いて提案法を評価する。

### 3.4 誤り訂正能力の評価

本節では、提案法の誤り訂正能力を評価する。まず、3.1 の条件 (1) について HMM の状態数を 2, 4, 8, 16 としたときの提案法の性能を図 5 に示す。図より、HMM の状態数を増すに従って誤り訂正能力が向上することがわかる。3% 誤りでは、状態数に対して、それぞれ、13%, 23%, 33%, 39% 程度のひずみ低減効果が確認された。次に、3% 誤りに対する Outlier を表 1 に示す。表 1 では、2~4 dB の割合はほとんど変化していないが 4 dB を越える符号の割合が大きく減っている。したがって、性能が大きく劣化する符号誤りに

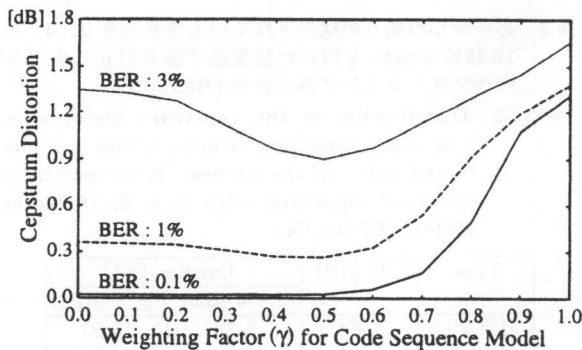


図 3 8 状態 HMM を用いて提案法により誤り訂正したときの符号列モデルに対する重み付け ( $\gamma$ ) とケプストラムひずみの関係

Fig. 3 The relation between the value of a weighting factor ( $\gamma$ ) for the code sequence model and cepstrum distortion. The code sequence is corrected by the proposed error correction algorithm with 8 states HMM.

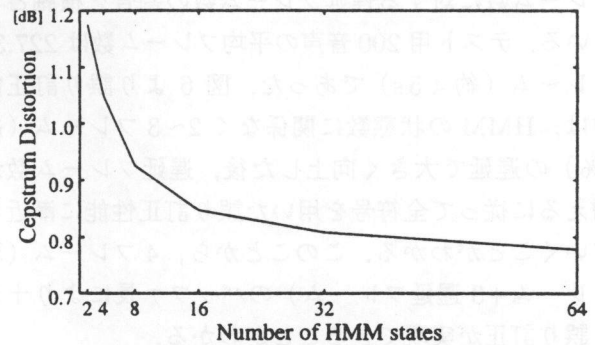


図 4 2~64 状態 HMM を用いて誤り訂正したときのケプストラムひずみ (BER=3%)

Fig. 4 The cepstrum distortion of the corrected code sequence using 2 to 64 states HMM. (BER=3%)

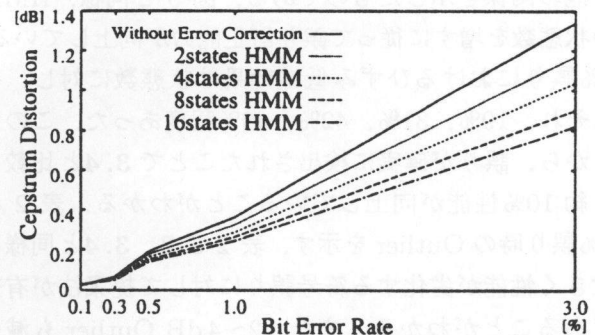


図 5 2, 4, 8, 16 状態 HMM を用いた提案法で誤り訂正したときの通信路上のビット誤り率に対するケプストラムひずみ

Fig. 5 The cepstrum distortion vs. bit error rate on channel. The code sequence is corrected by proposed algorithm with 2, 4, 8, 16 states HMM.



表 1 2, 4, 8, 16 状態 HMM を用いた提案法による誤り訂正時のケプストラムひずみの分布 (BER=3%)

Table 1 Distribution of the cepstrum distortions. The code sequence is corrected by proposed algorithm with 2, 4, 8, 16 states HMM. (BER=3%)

Type	CD [dB]	Outlier [%]	
		2 - 4 [dB]	> 4 [dB]
訂正なし	1.34	5.78	15.81
2states	1.17	5.77	14.01
4states	1.04	5.73	12.49
8states	0.90	5.78	10.91
16states	0.82	5.79	9.90

対して提案法が有効であることがわかる。

次に、3.1の条件(2)に対応して、実システムでの利用を考慮して演算量及びアルゴリズム遅延に影響するバッファ長と誤り訂正能力の関係について評価した。誤り率3%におけるアルゴリズム遅延フレーム数と性能の関係を図6に示す。図6は音声データの全フレーム数に対する遅延フレーム数の割合を横軸としている。テスト用200音声の平均フレーム数は227.38フレーム(約4.5s)であった。図6より誤り訂正能力は、HMMの状態数に関係なく2~3フレーム(約1%)の遅延で大きく向上した後、遅延フレーム数が増えるに従って全符号を用いた誤り訂正性能に漸近していくことがわかる。このことから、4フレーム(現フレーム+3遅延フレーム)のバッファ長により十分な誤り訂正が実現できることがわかる。

### 3.5 誤り検出との組合せ

本節では、誤りがあらかじめ検出されているとしたときの誤り訂正能力について調べる。まず、3.1の条件(1)について実験した。図7はHMMの状態数と性能の関係を示したものである。図5と同様、HMMの状態数を増すに従って誤り訂正能力が向上している。3%誤りにおけるひずみ低減効果は状態数に対し、それぞれ、19%、31%、42%、49%程度あった。このことから、誤りが確実に検出されたことで3.4と比較して約10%性能が向上していることがわかる。表2に3%誤り時のOutlierを示す。表2より、3.4と同様に大きく性能が劣化する符号誤りに対して提案法が有効であることがわかる。また、2~4dB Outlierも減少していることから、3.4で訂正することができなかった劣化の少ない符号誤りも訂正していることがわかる。

次に、3.1の条件(2)に対応して図8に誤り率3%での遅延量と性能の関係を示す。3.4の結果を踏まえ、3フレーム遅延までを評価した。本実験ではあらかじめ

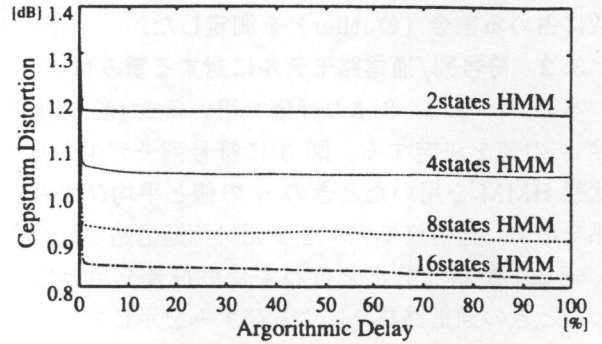


図 6 2, 4, 8, 16 状態 HMM を用いて誤り訂正したときのアルゴリズム遅延とケプストラムひずみの関係 (BER=3%)

Fig. 6 The relation between the algorithmic delay and the cepstrum distortion of the corrected code sequence using 2, 4, 8, 16 states HMM. (BER=3%)

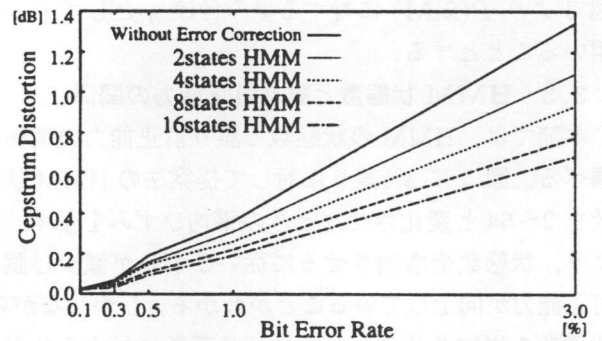


図 7 あらかじめ誤りが検出されている符号列を2, 4, 8, 16 状態 HMM を用いた提案法で誤り訂正したときの通信路上のビット誤り率に対するケプストラムひずみ

Fig. 7 The cepstrum distortion vs. bit error rate on channel. The code sequence, whose errors are detected prior to the process, is corrected by proposed algorithm with 2, 4, 8, 16 states HMM.

表 2 あらかじめ誤りが検出されている符号列を2, 4, 8, 16 状態 HMM を用いた提案法で誤り訂正したときのケプストラムひずみの分布 (BER=3%)

Table 2 Distribution of the cepstrum distortions. The code sequence, whose errors are detected prior to the process, is corrected by proposed algorithm with 2, 4, 8, 16 states HMM. (BER=3%)

Type	CD [dB]	Outlier [%]	
		2 - 4 [dB]	> 4 [dB]
訂正なし	1.34	5.78	15.81
2states	1.09	5.77	13.14
4states	0.93	5.69	11.09
8states	0.78	5.69	9.21
16states	0.68	5.57	8.03

誤りが検出されているため、16 状態の場合、遅延なしでも全符号を用いて訂正した場合に対して70%程度の

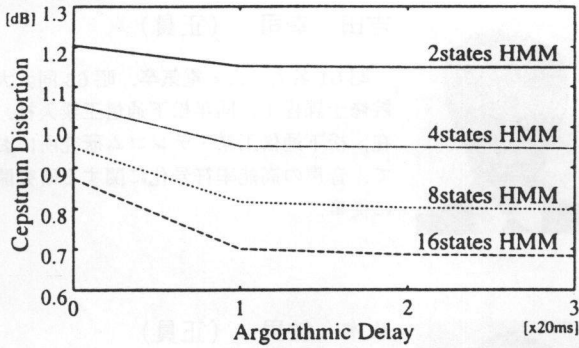


図8 あらかじめ誤りが検出されている符号列を2, 4, 8, 16状態HMMを用いて誤り訂正したときのアルゴリズム遅延とケプストラムひずみの関係(BER=3%)  
 Fig.8 The relation between the algorithmic delay and the cepstrum distortion of the code sequence, whose errors are detected prior to the process, is corrected by proposed algorithm using 2, 4, 8, 16 states HMM. (BER=3%)

性能を実現しており、また1フレーム遅延では97%もの性能を実現している。このことから、誤り訂正のみを行う場合バッファ長は2フレーム(現フレーム+1遅延フレーム)あれば十分であるといえる。

### 3.6 処理量

本節では1フレーム遅延で8bit符号を訂正することを仮定して、提案法の演算量を見積もった。以下に1符号置換の処理量を示す。対象としている符号数は二つ(現フレーム+1遅延フレーム)なので、複数の符号の誤り訂正では、この値を2倍した処理量で十分に収束すると考えられる。

まず、16状態HMMを用いてすべてのビット誤りの組合せに対して最ゆう推定した場合、1フレームあたり約457kステップが必要であった。つまり、分析周期20msでは約23MOPSとなる。この値は1bit誤りのみを対象に訂正を行うことで約5MOPSとなる。次に、8状態HMMを用いた場合、1フレームあたり約181kステップ(約9MOPS)が必要であるが、1bit誤りのみを対象とすると約1.7MOPSで実現できることがわかった。

## 4. むすび

本論文では情報源符号化後に残る符号間遷移の冗長性を利用して誤り訂正を行う手法を提案した。提案法は符号列モデルと通信路モデルから算出される符号列の生起確率に基づいて送信された符号列を最ゆう推定するものである。符号列モデルとしてHMMを利用した効率良い最ゆう推定アルゴリズムを示し、シミュ

レーションにより性能を評価した。

LSPパラメータを符号化した符号列に対して評価した結果、3%ランダム誤りに対して、16状態HMMを用いることで約39%ケプストラムひずみを低減できることを示した。特に、劣化の大きな誤りに対して提案法が有効であることを確認した。更に、あらかじめ誤りが検出されている場合、約49%のケプストラムひずみ低減効果があることを示した。

提案法は符号間遷移の冗長性を利用した誤り訂正法であり、誤り訂正符号を付加することなく通信路誤り耐性を向上させることができる。また、符号復号化のアルゴリズムを変更する必要がないため、復号化の前処理として適用することにより、既存の符号化方式の通信路誤り耐性を向上させることができる。

本論文ではLSPパラメータのみを対象として提案法を評価した。つまり、LSPパラメータのフレーム間相関を利用した誤り訂正法の訂正能力について検討した。一方、HMMは出力パラメータとして音声符号化で用いられる他のパラメータをも含めて情報源をモデル化することができる。今後の課題として、このようなパラメータ間の相関を利用したときの誤り訂正能力の評価、既存の符号化方式へ適用したときの性能評価が挙げられる。

## 文 献

- [1] 熊沢宏之, 笠原正雄, 滑川敏彦, “通信路誤りを考慮したベクトル量子化器の構成,” 信学論(B), vol.J67-B, no.1, pp.1-8, Jan. 1984.
- [2] 相澤清晴, 原島 博, 宮川 洋, “通信路歪を軽減するベクトル量子化の伝送符号の割りあてに関する一検討,” 信学技報, no.CS84-32, 1984.
- [3] N. Farvardin, “A study of vector quantization for noisy channels,” IEEE Trans. Inf. Theory, vol.IT-36, no.4, pp.799-809, July 1990.
- [4] J.G. Dunham and R.M. Gray, “Joint source and noisy channel trellis encoding,” IEEE Trans. Inf. Theory, vol.IT-27, no.4, pp.516-519, July 1981.
- [5] M.W. Marcellin, T.R. Fischer, and J.D. Gibson, “Predictive trellis coded quantization of speech,” IEEE Trans. ASSP., vol.38, no.1, pp.46-55, Jan. 1990.
- [6] P. Sriram and M.W. Marcellin, “Performance of adaptive prediction algorithms for trellis coded quantization of speech,” IEEE Trans. Commun., vol.42, no.2/3/4, pp.1512-1517, Feb./March/April 1994.
- [7] S. Kang and T.R. Fischer, “Trellis excitation speech coding at low bit rates,” IEEE Trans. Commun., vol.42, no.2/3/4, pp.1902-1910, Feb./March/April 1994.
- [8] A. Popescu, N. Moreau, and C. Lamblin, “Celp cod-



ing using trellis-coded vector quantization of the excitation," IEEE Trans. Speech and Audio Process., vol.3, no.6, pp.464-472, Nov. 1995.

- [9] D. Rowe and P. Secker, "A robust 2400 bit/s mbe-lpc speech coder incorporating joint source and channel coding," Proc. ICASSP, pp.II141-II144, San Francisco, U.S.A., March 1992.
- [10] K.T. Malone and T.R. Fischer, "Trellis-searched adaptive predictive coding of speech," IEEE Trans. Speech and Audio Process., vol.1, no.2, pp.196-206, April 1993.
- [11] 楠 雅典, 長谷川孝明, "誤りのある通信路における音声符号化のための遷移を制約したベクトル量子化法について," 信学論 (A), vol.J77-A, no.6, pp.905-914, June 1994.
- [12] T. Fingscheidt and P. Vary, "Robust speech decoding: A universal approach to bit error concealment," Proc. ICASSP, pp.1667-1670, Munich, Germany, April 1997.
- [13] F.I. Alajaji, N.C. Phamdo, and T.E. Fuja, "Channel codes that exploit the residual redundancy in CELP-encoded speech," IEEE Trans. Speech and Audio Process., vol.4, no.5, pp.325-336, Sept. 1996.
- [14] C. Berrow and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," IEEE Trans. Commun., vol.44, no.10, pp.1261-1271, Oct. 1996.
- [15] J. Garcia-Frias and J.D. Villasenor, "Combining hidden markov source models and parallel concatenated codes," IEEE Commun. Lett., vol.1, no.4, pp.111-113, July 1997.
- [16] 鹿野清宏, 中村 哲, 伊勢史郎, 音声・音情報のデジタル信号処理, 昭晃堂, 東京, 1997.
- [17] J.R. Deller, Jr., J.G. Proakis, and J.H.L. Hansen, Discrete-Time Processing of Speech Signals, Macmillan Publishing Company, New York, 1993.

(平成 11 年 3 月 24 日受付, 7 月 16 日再受付)



吉田 幸司 (正員)

平 2 阪大・基礎工・制御卒。平 3 松下通信工業入社。平 6~8 奈良先端大学院大在籍, 平 8 同大修士課程了。現在, 松下通信工業・テレコム研究所において, 音声符号化を中心とした音声信号処理の研究開発に従事。平 10 より奈良先端大学院大博士後期課程に在籍, 音響学会会員。



吉田 幸司 (正員)

昭 61 名大・工・電気卒。昭 63 同大大学院修士課程了。同年松下通信工業入社。現在, 松下通信工業・テレコム研究所において, 音声の高効率符号化に関する研究開発に従事。



八木 敏男 (正員)

昭 53 明大・工・電気卒・昭 55 同大大学院修士課程了。同年松下通信工業入社。現在, 松下通信工業・テレコム研究所においてアプリケーションプロトコル等, デジタル携帯端末を用いた情報通信に関する研究開発に従事。音響学会会員。



鹿野 清宏 (正員)

昭 45 名大・工・電気卒。昭 47 同大大学院修士課程了。同年電電公社武蔵野電気通信研究所入所。昭 59~61 カーネギーメロン大客員研究員。昭 61~平 2ATR 自動翻訳電話研究所音声情報処理研究室長。平 4NTT ヒューマンインタフェース研究所主席研究員。平 6 より奈良先端科学技術大学院大学情報科学研究科教授。音情報処理学講座を担当。工博。主として音声・音情報処理の研究及び研究指導に従事。昭 50 本会米沢賞, 平 3 IEEE SP 1990 Senior Award, 平 6 日本音響学会技術開発賞受賞。IEEE, 音響学会, 情報処理学会各会員。