

遷移を制約したベクトル量子化による  
音声の情報源 - 通信路統合符号化方式に関する研究

Studies on Joint Source-Channel Coding Methods of Speech  
using Constrained Transition Vector Quantization

1995年3月

埼玉大学大学院理工学研究科（博士後期課程）

生産情報科学専攻（主指導教官 小林 禧夫）

埴 雅典

## 要旨

近年、デジタルセルラフォンやパーソナルハンディフォンなどの移動音声通信システムが急速に普及しつつある。これらのシステムのユーザ数のさらなる増加によって周波数帯域の不足が問題となり、この問題を解決するために高能率な情報源符号化システムが必要となることが予想される。一方、移動音声通信システムでは無線通信路を用いるが、無線通信路ではフェージングなどの影響で通信環境が大きく変動し多くの通信路誤りが重畳するため、劣悪な通信環境下での通信を可能とする通信路符号化を行なわなくてはならない。従来、通信路誤り対策には、情報源符号化と通信路符号化を縦列接続することでそれぞれを独立して行なうタンデム符号化を用いることが一般的であった。しかし近年、より高能率な符号化システムの実現を目的として、情報源符号化と通信路符号化を同時に行なう情報源 - 通信路統合符号化方式の検討が行なわれている。本論文では、高能率な符号化を可能とするベクトル量子化において通信路誤りを訂正可能にする新しい情報源 - 通信路統合符号化方式について述べ、次にその手法を実際の音声符号化システムへ応用した例について述べる。

第1章では、まず、音声の情報源 - 通信路統合符号化へのベクトル量子化の応用の歴史について述べ、本研究の位置付けを行なう。

第2章では、本研究を理解する上で必要な基礎知識として、従来の音声符号化システムおよび従来の情報源 - 通信路統合符号化法について総括し、その問題点を整理する。

第3章では、新しい情報源 - 通信路統合符号化法として、遷移を制約したベクトル量子化 (CTVQ) を提案する。CTVQでは、ベクトル量子化のコードベクトル間の遷移を制約して符号化に用いることで、符号器が出力可能な符号系列を制約する。通信路で符号系列に誤りが重畳した場合には、受信符号系列はこの制約を満たさなくなるため、復号器において通信路で符号に重畳した誤りが検出可能となり、さらに畳み込み符号の



復号と同様にビタビアルゴリズムを適用することによって誤りの訂正が可能となる。第3章で提案する手法は、ベクトル量子化のコードベクトルとトレリスの状態を対応させる手法であり、状態ラベル型CTVQ(LS-CTVQ)と呼ぶことにする。LS-CTVQを用いて符号化することで、従来のタンデム符号化により符号化する場合と比較して6dB以上劣悪な通信環境において同じ復号品質が得られ、優れた誤り耐性を示す。

CTVQの誤り訂正はトレリス上でビタビアルゴリズムを実行して行なわれるが、この時に必要な計算量と記憶量はトレリスの状態数に比例する。LS-CTVQではコードベクトルとトレリスの状態が対応するため、ベクトル量子化の精度を向上するためにコードベクトルを増やした場合、同時に誤り訂正に必要な記憶量と演算量も増加してしまう。第4章では、この問題に対処する改良法として、遷移ラベル型CTVQ(LT-CTVQ)を提案する。LT-CTVQではLS-CTVQとは異なり、トレリスの状態間の遷移にコードベクトルを対応させたいうで、全コードベクトル数よりも状態数を少なく設定する。これによってLS-CTVQに比べて計算量と記憶量を低減することが可能になる。LT-CTVQにおいて状態数を少なく設定して計算量を抑えた場合、LS-CTVQに比べてわずかに量子化ひずみが増大するものの、ほぼ同程度の通信路誤り耐性を有し、計算量と記憶量の低減に有効であることを示す。

続く第5章では、CTVQの実際の音声符号化システムへの応用例として、残差励振線形予測(RELP)音声符号化システムにLT-CTVQを適用して構成した、2.8kbps音声符号化システムについて述べる。このシステムでは、RELPシステムの出力をLT-CTVQで符号化することによって、高能率化と誤り耐性の強化を同時に行なう。LT-CTVQを用いて構成したシステムでは、情報源符号化に有限状態ベクトル量子化、通信路符号化に畳み込み符号を用いたタンデムシステムに比べて、7dB劣悪な通信路においても同じ品質の音声通信が可能であることを示す。

第6章は結論であり、本論文を総括する。

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 ベクトル量子化の音声の情報源 - 通信路統合符号化への 応用の歴史 . . . . .	1
1.2 本論文の位置づけと目的 . . . . .	2
1.3 本論文の構成 . . . . .	5
<b>第2章 音声の情報源符号化と通信路符号化に関する基礎理論</b>	<b>7</b>
2.1 まえがき . . . . .	7
2.2 音声の情報源符号化方式 . . . . .	7
2.2.1 音声の性質 . . . . .	8
2.2.2 波形符号化方式 . . . . .	11
2.2.3 ボコーダ方式 . . . . .	13
2.2.4 ベクトル量子化 . . . . .	16
2.2.5 ハイブリッド方式 . . . . .	18
2.3 音声符号化システムにおける通信路符号化方式 . . . . .	20
2.3.1 誤り検出と訂正の原理 . . . . .	21
2.3.2 誤り検出符号の音声符号化システムへの応用 . . . . .	23
2.3.3 誤り訂正符号の音声符号化システムへの応用 . . . . .	23
2.3.4 音声符号化システムにおけるその他の通信路符号 化方式 . . . . .	24
2.4 本章のまとめ . . . . .	25
<b>第3章 遷移を制約したベクトル量子化法</b>	<b>28</b>
3.1 まえがき . . . . .	28
3.2 トレリスの状態にラベル付けを行なう遷移を制約したベ クトル量子化 (LS-CTVQ) . . . . .	29
3.2.1 符号化アルゴリズム . . . . .	30
3.2.2 遷移を制約したコードブックの作成 . . . . .	31
3.2.3 LS-CTVQ 符号のビット誤り訂正法 . . . . .	32

3.3	計算機シミュレーション . . . . .	38
3.3.1	コードベクトル間の遷移を制約する影響 . . . . .	38
3.3.2	ランダム誤り下での基礎的検討 . . . . .	40
3.3.3	フェージング環境下における LS-CTVQ 符号の軟 判定復号の効果 . . . . .	44
3.3.4	従来のタンデム符号化との比較 . . . . .	46
3.4	本章のまとめ . . . . .	49
<b>第 4 章</b>	<b>遷移を制約したベクトル量子化の改良法</b>	<b>51</b>
4.1	まえがき . . . . .	51
4.2	トレリスの遷移にラベル付けを行なう遷移を制約したベ クトル量子化 (LT-CTVQ) . . . . .	52
4.2.1	符号化アルゴリズム . . . . .	52
4.2.2	遷移ラベル型 CTVQ 符号の誤り訂正 . . . . .	54
4.2.3	コードブック設計アルゴリズム . . . . .	56
4.3	計算機シミュレーション . . . . .	58
4.3.1	ランダム誤り下での基礎的検討 . . . . .	58
4.3.2	フェージング環境下における LT-CTVQ 符号の軟 判定復号の効果 . . . . .	62
4.3.3	従来のタンデム符号化方式との比較 . . . . .	64
4.3.4	LS-CTVQ と LT-CTVQ の比較 . . . . .	66
4.4	本章のまとめ . . . . .	68
<b>第 5 章</b>	<b>遷移を制約したベクトル量子化の 2.8kbps 残差励振線形予 測音声符号化システムへの応用</b>	<b>69</b>
5.1	はじめに . . . . .	69
5.2	CTVQ を用いた 2.8kbps 残差励振線形予測音声符号化シ ステム . . . . .	70
5.3	計算機シミュレーション . . . . .	73
5.3.1	線スペクトル対パラメータの符号化 . . . . .	73
5.3.2	残差信号の符号化 . . . . .	76

5.3.3	セグメンタル SNR による復号音声の品質評価 . . .	77
5.4	本章のまとめ . . . . .	81
第 6 章	結論	83
	謝辞	86
	参考文献	87

## 第1章 序論

### 1.1 ベクトル量子化の音声の情報源 - 通信路統合符号化への応用の歴史

ベクトル量子化 (VQ)<sup>(1-3)</sup> は 1950 年代の Dudley による Phonetic Pattern Recognition Vocoder<sup>(4)</sup> や 1960 年代の Smith による研究<sup>(5)</sup> など、古くから音声の符号化に用いられてきた。しかし、本格的に広まったのは 1980 年に Buzo らによりコードブック学習アルゴリズムとして LBG アルゴリズム<sup>(6)</sup> が提案されてからである。LBG アルゴリズムの登場後、量子化時の計算量やコードブックの保持に必要な記憶量、さらに符号化のビットレートを低減するための様々な改良型 VQ が提案され<sup>(7-11)</sup>、数百 bps から十数 kbps までの幅広い範囲のビットレートの様々な音声符号化システムに応用されてきた<sup>(12-26)</sup>。最近では、VQ を用いて音声の励振波形を高品質に伝送する符号励振線形予測 (CELP)<sup>(51)</sup> 音声符号化システムが広く実用化されており、高能率な音声符号化システムを構成する場合に VQ は不可欠な技術となっている。

一方、音声に限らず符号化システムを通信システムに適用する場合には、通信路で生じるビット誤りに対処するためになんらかの形で通信路符号化を行なう必要がある。特に VQ を用いた符号化システムのように一つの符号が多く情報を担っている場合には、わずかな通信路誤りが重大な品質劣化を起す可能性がある。Shannon によって情報源符号化と通信路符号化がその最適性を失うことなく分離することが可能であると示唆されて以来、一般的に情報源符号化と通信路符号化は独立して行なわれてきたが、他方でこれらを総合的に扱う情報源 - 通信路統合符号化の検討も行なわれていた<sup>(27-34)</sup>。

VQ の符号化への適用において通信路誤りの影響を考慮し、情報源符号化と通信路符号化を総合的に検討した例としては、1984 年に熊沢らによって発表された通信路誤りを考慮したベクトル量子化法<sup>(35)</sup> がある。これは通信路において重畳する誤りをあらかじめ想定してコードブックを



デザインし、さらに符号化時に通信路誤りが重畳することを想定してベクトルの選定を行なうことで、ベクトル量子化器を特定の通信路に整合させる手法である。その後、1987年に熊沢らのアルゴリズムと同様の手法によって通信路誤りを考慮してコードベクトルにバイナリコードを割り当てる手法がいくつか検討されていることを除けば<sup>(36-39)</sup>、VQにおける通信路誤りの影響を考慮した検討はなく、数年間はこの分野における大きな進展はなかったようである。しかし90年代になると、Farvardinによるシミュレーテッドアニーリングを用いて通信路誤りを考慮したバイナリコード割り当てを行なうアルゴリズム<sup>(41)</sup>や、Moriyaによる二通信路共役ベクトル量子化法<sup>(44)</sup>など新たな手法が相次いで報告されている。さらにFarvardinらは、熊沢らとは異なる手法でベクトル量子化器を通信路に最適化するアルゴリズム<sup>(39)</sup>を提案し、さらに木探索型VQや多段VQに拡張<sup>(42)</sup>している。上述した手法以外にも、1990年代に入ってからVQの情報源 - 通信路統合符号化への応用例がいくつも報告されており<sup>(40,43,45-48)</sup>、さらには日本の次世代の携帯電話用のハーフレート符号化システム<sup>(52)</sup>の一部にもこの種のベクトル量子化法<sup>(47)</sup>が採用されるなど、今後さらなる発展が期待できる分野である。

## 1.2 本論文の位置づけと目的

近年、携帯電話などの移動通信システムが急速に普及しつつあるが、今後のユーザー数の急増ともなう同時通信局数の逼迫が予想され、周波数を効率的に利用できる音声符号化方式が求められている。しかし、このような高能率な音声符号化方式は概して通信路で重畳する符号誤りに対して脆弱で、通信路符号化として誤り訂正符号化などを行なうものの、わずかな通信路誤りが重大な品質劣化を引き起こす可能性がある。この問題に対処するために、高能率でしかも通信路誤りに強い符号化方式の開発が急務である。

高能率な符号化のためには符号から冗長を取り除く必要があり、通信路誤り耐性の強化のためには符号に冗長を含ませなくてはならない。こ

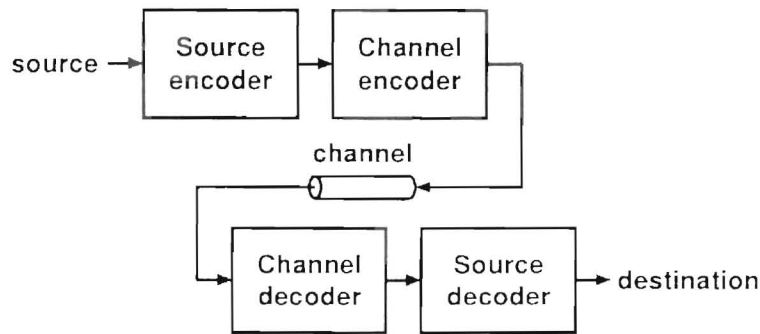


図 1.1 情報源 - 通信路タンデム符号化

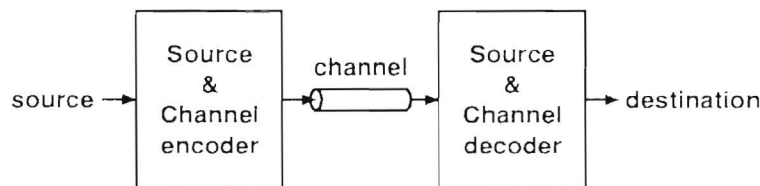


図 1.2 情報源 - 通信路統合符号化

のように、高能率化を目的とする情報源符号化と通信路誤りに対する耐性の強化を目的とする通信路符号化は本来相反するものであり、それぞれを独立して行なったりえでタンデム接続する情報源 - 通信路タンデム符号化方式が一般的に用いられてきた(図 1.1)。しかし近年、符号化システム全体での符号化特性の向上を目的として、情報源符号化と通信路符号化を統合した情報源 - 通信路統合符号化方式(図 1.2)の検討がいくつかなされている(27, 28, 30, 32, 40-42, 44)。

北米および日本における携帯電話用のフルレート音声符号化システム<sup>(28)</sup>(誤り制御込みで 11.2kbps 程度)や、汎ヨーロッパ移動無線(GSM)用のハーフレート音声符号化システム<sup>(30, 32)</sup>(誤り制御込みで 5.6kbps)用には、音声符号化器の出力する符号の各ビット位置における誤りによって復号音声に生じる品質劣化(誤り感度)を調べ、誤り感度が高いビットほど強力な誤り保護を行なう符号化システムが採用されている。この種の手法は従来からのタンデム符号化方式の延長と考えられ、符号化のビット

レートが比較的高い場合 (数 kbps から 10 数 kbps 程度) には有効である。

一方、ベクトル量子化は高能率な符号化を実現するうえで必須の技術である。ベクトル量子化は符号化のビットレートが比較的低い場合に良く用いられるが、この場合にはわずかなビット誤りが重大な品質劣化を引き起こす。この問題に対処するために、ベクトル量子化における情報源 - 通信路統合符号化法として、前節で述べた Moriya による二通信路共役ベクトル量子化法<sup>(44)</sup>や Farvardin による通信路誤りに強いバイナリコード割り当て法<sup>(41)</sup>など、符号の誤り感度を下げることによって誤り耐性の強化を行なう手法の検討がいくつかなされている。日本における次世代の携帯電話用の音声符号化システムにおいてもこの種の手法がその一部に用いられている。これらの手法は、通信路における誤りがあまり多くない場合には有効だが、誤りの多い劣悪な通信路では誤り訂正符号化などと組み合わせる必要がある。

このように、現在、高能率な符号化システムを検討する場合に、通信路誤りに対する耐性を考慮することは重要な課題の一つとなっている。このような背景の下で、本論文では新しい音声の情報源 - 通信路統合符号化方式として、遷移を制約したベクトル量子化 (CTVQ) の検討を行なう。CTVQ では、情報源符号に残される冗長を積極的に利用することで、通信路で重畳した符号の誤りを訂正可能である。さらに、CTVQ の誤り訂正時の記憶量と演算量を低減する改良法の検討、および実際の音声符号化システムへの応用として、2.8kbps の残差励振線形予測音声符号化システムに適用した場合についての検討を行なう。提案する手法を用いることで、情報源符号化と通信路符号化を独立して行なうタンデム符号化方式より優れた符号化特性を持つシステムが構成可能で、符号化のビットレートが低いシステムを劣悪な通信路に適用する場合に有効であることを示す。



### 1.3 本論文の構成

本論文は図 1.3 のように 6 章構成であり、その内容は以下の通りである。第 2 章では、音声の情報源 - 通信路統合符号化に関する基礎理論と題して、音声の線形予測、ベクトル量子化、誤り制御技術など、本論文を理解する上で必要な基礎技術の解説を行なう。第 3 章では、トレリスの状態にラベルづけを行なう遷移を制約したベクトル量子化法を提案し、音声の線形予測ケプストラムの符号化／伝送に適用した場合の性能について述べる。第 4 章では、CTVQ の復号時の計算量と記憶量を低減する改良法として、トレリスの遷移にラベルづけを行なう CTVQ を提案する。第 5 章では、CTVQ の実システムへの応用として、2.8kbps の残差励振線形予測符号化システムへ CTVQ を応用した場合について述べる。第 6 章では、本論文を総括する。

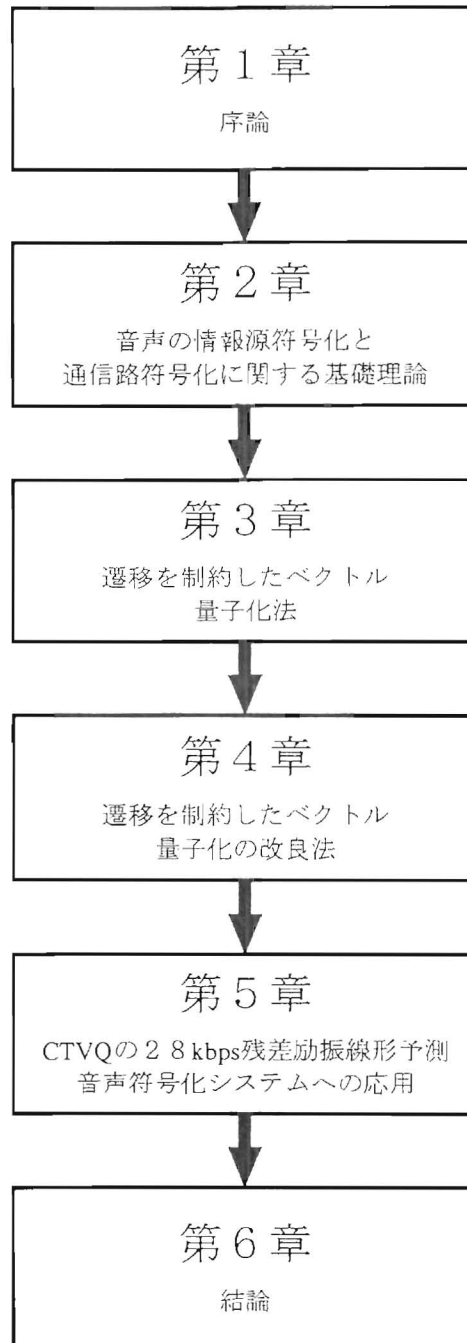


図 1.3 本論文の構成

## 第2章 音声の情報源符号化と通信路符号化に関する基礎理論

### 2.1 まえがき

現在、デジタル通信網が急速に発達しつつあり、近い将来、音声を含む様々な情報がこのデジタル通信網を通じて伝送されることが予想される。また、セルラ電話などの移動音声通信が急速に普及しつつあり、同時に通話が可能なユーザ数を増やすために、効率的に音声を伝送可能な方法が必要とされている。このような背景の下で、音声信号の効率的なデジタル伝送または蓄積を可能にする高能率な情報源符号化方式の必要性がますます高まってきている。さらに、音声をデジタル的に伝送／蓄積する場合、通信路における伝送誤りや蓄積情報の読みとり誤りなどによる再生音声の品質劣化が起こることが予想され、これに対処するために誤り訂正符号化等の通信路符号化を行なう必要がある。本章では既存の音声の情報源符号化方式とその通信路符号化方式を総括し、その問題点を整理する。

本章の構成は以下の通りである。まず2.2では、音声の情報源符号化方式について概観する。ここでは、まず音声の性質について述べた後に、もっとも基本的な符号化方式であるパルス符号化変調(PCM)に代表される波形符号化方式、符号化対象を音声に特定することで高能率化を実現するボコーダ方式、ベクトル量子化方式、波形符号化とボコーダ方式のハイブリッド方式について述べる。さらに2.3では、移動体音声通信において通信路で生じる問題について述べ、従来の音声符号化システムにおける通信路符号化法について問題点を整理する。

### 2.2 音声の情報源符号化方式

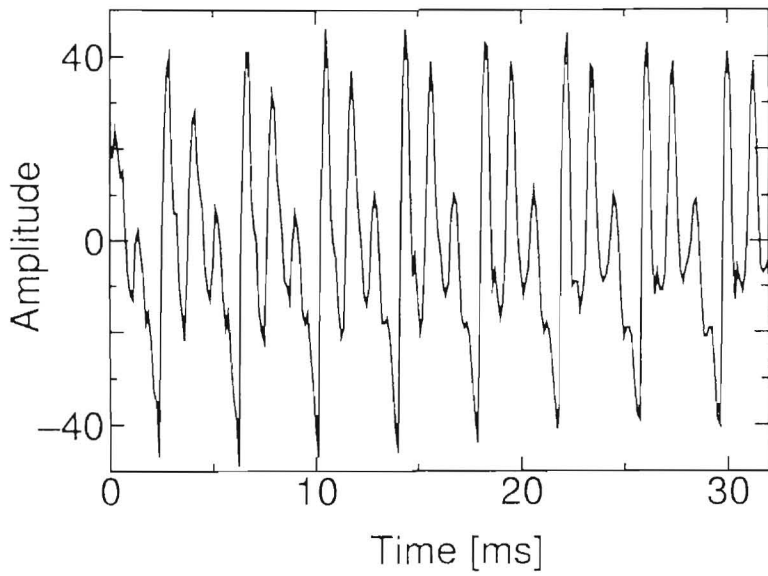
音声の情報源符号化方式の多くは、音声波形の標本化と振幅値の量子化を行なうことから始まる。標本化と量子化を行なっただけの音声信号は高いビットレートを必要とするため(標本化周波数8kHzで8ビット量

子化して64キロビット毎秒[kbps]),品質をできるだけ保ちつつ低いビットレートで高能率に表現することが音声の情報源符号化方式の目的である。この情報源符号化方式は、音声波形を直接符号化するノンパラメトリック法と、音声の生成モデルを仮定してモデルのパラメータを符号化するパラメトリック方式、さらにそれらのハイブリッド方式に大きく分類することができる。ノンパラメトリックな符号化を行なう場合、符号化に必要なビットレートは高くなるが(16kbps以上)、復号した音声は非常に高い品質のものが得られる。パラメトリック符号化方式ではノンパラメトリック方式に比べて高能率な符号化が可能(2.4kbps以下)だが、復号音声の品質は劣化する。これらの両者の特徴を併せて数kbps程度で高品質な復号音声を得ることを目的とした方式がハイブリッド符号化方式である。以下では、まず音声の性質を簡単に述べた後にこれらの音声符号化方式を概観する。

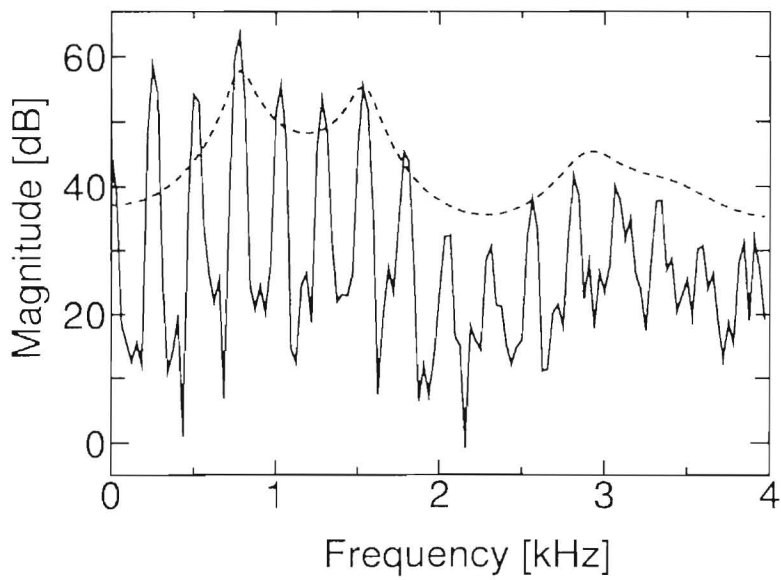
### 2.2.1 音声の性質

音声信号は非定常信号であるが、5～20ミリ秒[ms]程度の短時間には準定常的な信号とみなすことができ、音声信号を数ms程度に分割した短時間セグメントの特徴によって、/a/や/e/などの声帯の振動を伴う有声音と、/s/や/p/などの声帯の振動を伴わない無声音に分類することができる。しかし実際には、これらの中間の“わたり”も存在する。有声音/a/と無声音/s/の短時間セグメントの、時間波形および短時間スペクトルの例を図2.1と図2.2に示す。有声音/a/の時間波形はほぼ周期的で、短時間スペクトルには調波構造が見られる。これに対して、無声音/s/の時間波形はランダムで/a/に比べて振幅が小さい。これは/s/に限ったことではなく、一般的に無声音は有声音に比べてエネルギーが小さい。また無声音の短時間スペクトルには調波構造がなく、有声音に比べて高周波成分が強いのが特徴である。

図2.1の/a/の短時間スペクトルには、点線で線形予測分析によるスペクトル包絡も併せて示してある。有声音の短時間スペクトルは、スペ

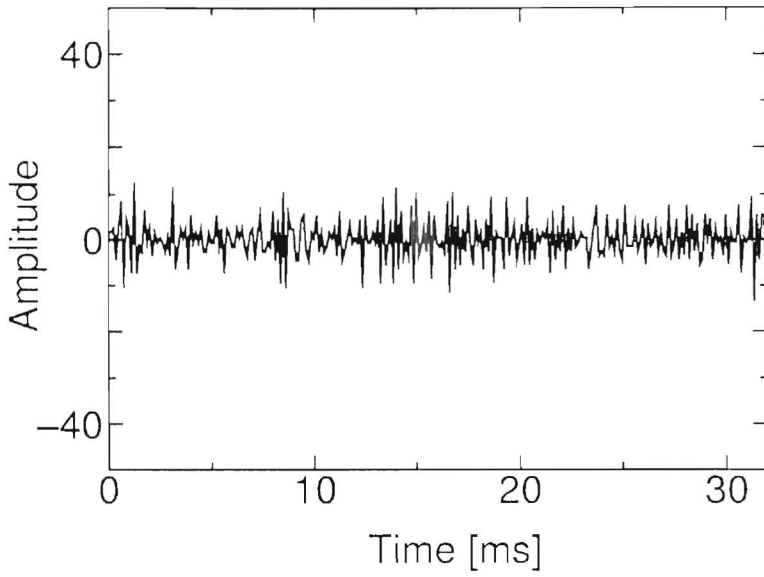


(a) /a/ の時間波形

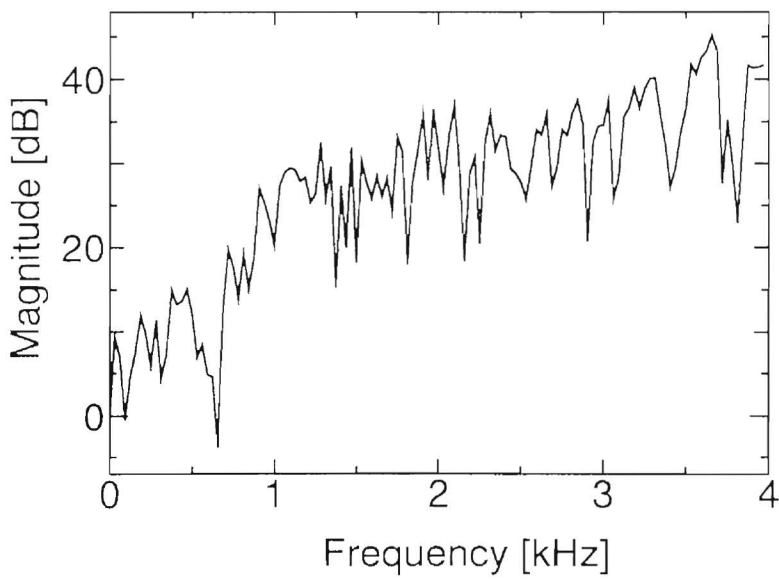


(b) /a/ の短時間スペクトル

図 2.1 有声音セグメント (/a/) とその短時間スペクトル



(a) /s/ の時間波形



(b) /s/ の短時間スペクトル

図 2.2 無声音セグメント (/s/) とその短時間スペクトル



クトル包絡のピークとして現れるホルマントと、スペクトルの微細構造を表す調波構造で特徴付けられる。図 2.1 では 3kHz 以下に 3 つのホルマントが確認できるが、この 3kHz 以下に存在するホルマントが発声内容の知覚に重要であるといわれている。このホルマントは発声時の声道の形によって決定され、無声音においても存在する。一般に、無声音ではホルマントが存在する周波数は有声音に比べ高くなり、無声音を高品質で表現するためには、高い周波数のホルマントをいかに良く保存できるかが重要になる。一方、調波構造は声帯の振動の周期によって決定され、有声音にのみ見られる。この声帯の振動の周期が調波構造の基本周期であり、これをピッチ周期または単にピッチと呼ぶ。ピッチ周期の逆数をピッチ周波数または基本周波数と呼ぶ。このピッチは、後述するボコーダ方式の音声符号化において重要な符号化パラメータの一つとなる。

### 2.2.2 波形符号化方式

音声に限らず、アナログ波形を離散時間、離散振幅近似して符号化を行なう際のもっとも基本的な手法は線形パルス符号化変調 (PCM) である。この手法では波形情報の高能率表現は全く行なわれず、他の情報源符号化方式の多くはこの線形 PCM で符号化された音声を高能率に符号化する手法であるともいえる。図 2.3 に線形 PCM の原理を示す。アナログ波形を一定の時間間隔  $T$  毎に標本化して得られる標本値は、一定の量子化幅  $\Delta$  毎に定められた値に量子化される。図 2.3 において、網かけの帯の内部に位置する標本値は量子化値  $y_j$  に量子化されるものとする、標本値  $x_i$  は  $y_j$  に量子化されることになる。この時、量子化の幅  $\Delta$  は量子化値を二進符号に符号化する際のビット数  $B$  で決まり、信号の振幅の存在範囲が  $\pm L$  とすると、 $\Delta = L/2^{(B-1)}$  となる。信号の振幅が一様に分布するものと仮定すると、この時の信号対雑音電力比 (SNR) は、

$$\text{SNR} = 10 \log_{10} 2^B \simeq 6B [\text{dB}] \quad (2.1)$$

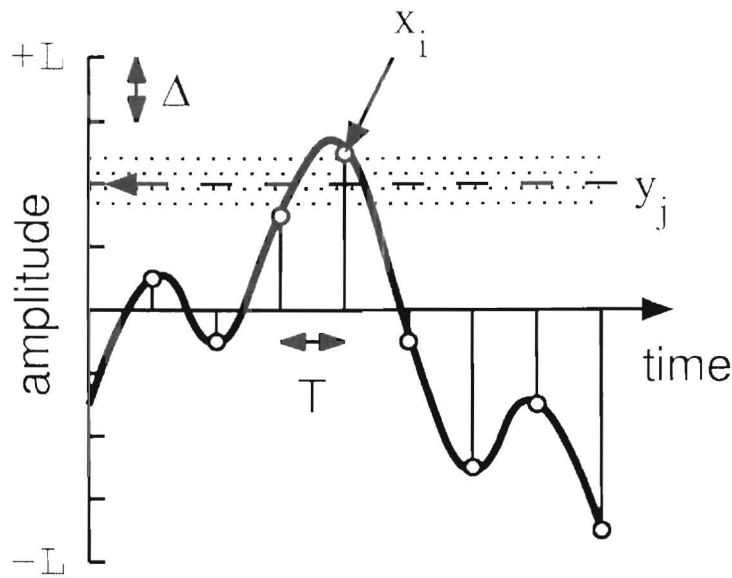


図 2.3 パルス符号化変調 (PCM) の原理

と表され、符号化のビット数を1ビット増やす毎にSNRが約6dB改善するため、所望のSNRが得られるように符号化のビット数を決めれば良いことがわかる。この線形PCMを発展させた手法として、信号の分布を考慮して最適な量子化幅を決定する最適量子化器や、信号の振幅分布を指数分布と仮定して対数圧伸を行なった上で量子化する $\mu$ -lawやA-lawなど、数多くの改良法が存在する。これらの改良法では線形PCMに比べて少ないビット数で同程度のSNRを実現できる。

PCMをさらに発展させた手法としては、時間的に連続する標本間の差分を量子化する差分PCM(DPCM)、量子化幅を適応的に変化させる適応PCM(APCM)、DPCMとAPCMのハイブリッド方式である適応差分PCM(ADPCM)などがあげられる。ここで32kbpsのADPCMはCCITTの標準方式として採用されており、さらに今後パーソナルハンディフォンシステム(PHS)の標準音声符号化方式として用いられるなど、多くの高ビットレート音声符号化方式のなかでも最も重要な方式のひとつである。



他の波形符号化方式としては、信号の増減のみを1ビットで符号化するデルタ変調(DM)、DMの量子化幅を適応的に変化させる適応デルタ変調(ADM)、信号を離散フーリエ変換や離散コサイン変換によって周波数領域で表現した後に符号化を行なう、適応変換符号化(ATC)に代表される各種変換符号化方式などがある。これらの手法は、基本的に標本値毎(変換符号化では周波数領域における標本値)に符号化を行なうスカラー符号化方式だが、これに対して複数の標本値をまとめて符号化するベクトル量子化方式も波形符号化に良く用いられる。このベクトル量子化については後述する。ここで述べた波形符号化方式は、音声波形をできる限り忠実に伝送することを目的としているためにビットレートが高く、低ビットレートの符号化には適さない。

### 2.2.3 ボコーダ方式

前節で述べた各種波形符号化方式は、信号を生成する情報源について特別なモデルを仮定しているわけではないので、その量子化対象は音声信号に限定されず高品質な復号結果が得られるが、符号化のビットレートは16kbps以上と高くなる。ビットレートを大幅に低下させるためには、情報源にモデルを仮定してモデルのパラメータを符号化するパラメトリック符号化が有効である。音声符号化ではこのような手法はボコーダ方式と総称される。ボコーダのモデルを図2.4に示す。ボコーダで



図 2.4 ボコーダのモデル

は、音声は声帯から発せられる励振信号によって声道が励振されることで発生されるものと仮定する。ここで、声道フィルタや励振信号のモデル化の相違によって様々なボコーダシステムが存在する。声道のモデル化については、声道フィルタを全極型の線形フィルタと仮定してフィルタ係数を線形予測分析で決定する線形予測符号化(LPC)システムが、数多くの応用において成功を収めている(例えば文献(51)など)。また励振モデルは、有声音をパルス列、無声音はランダムノイズ列で励振する単純な2状態励振モデルが基本である。

図2.5は、励振モデルに2状態励振モデル、声道のモデルに線形フィルタを仮定した音声の線形生成モデルを表している。線形予測符号化方式

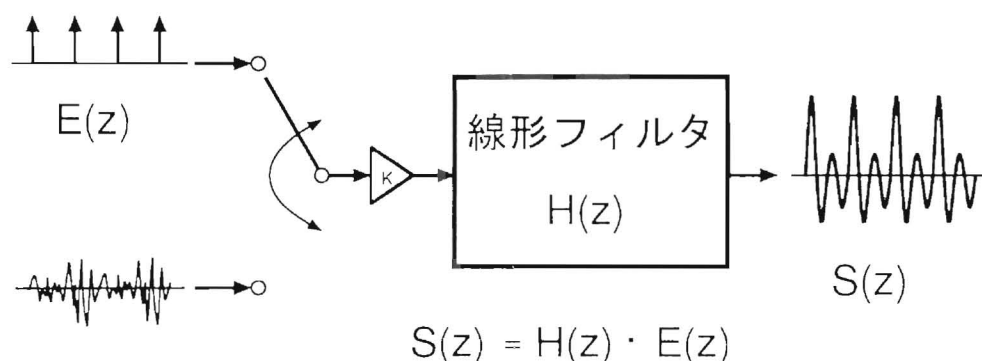


図 2.5 音声の線形生成モデル

では、この生成モデルにおいて、音源情報としては、ゲイン、有声/無声の情報、さらに有声音の場合のピッチを符号化し、声道情報として合成用の全極型フィルタの係数を符号化する。ここで、 $p$  次の全極型フィルタは  $z$  領域において次のように表現できる。

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2} - \dots - a_p z^{-p}} \quad (2.2)$$

ここで、入力音声信号を  $s(n)$ 、合成フィルタの出力を  $\hat{s}(n)$  とすると、推定誤差  $e(n) = s(n) - \hat{s}(n)$  の期待値  $E[e(n)^2]$  が最小になるように、(2.2) 式の係数  $a_i, i = 1, 2, \dots, p$  の組を線形予測分析によって求める。線形予測の推定誤差  $e(n)$  は残差信号と呼ばれ、励振信号に対応する。求まっ

たフィルタ係数(線形予測係数)は音声のスペクトル包絡を表現するパラメータであり、これを低ビットレートで符号化することで高能率な音声符号化が可能となる。しかし、線形予測係数  $a_n$  は量子化特性が悪く、低ビットレートで符号化した場合、合成フィルタの特性が符号化対象の音声のものとずれたり、フィルタが不安定になるなどの欠点を有する。これらの問題点を解決するために、偏自己相関(PARCOR)方式、線形予測ケプストラム方式、線スペクトル対(LSP)方式などの改良法が提案されている。線形予測ケプストラムは、その自乗距離として定義されるケプストラム距離が、人間の聴覚特性に良く一致することから、符号化音声の品質評価や音声認識の特徴量として良く用いられる。またLSP係数は、声道内部において  $Q$  が無限大の共振が起こる時の共振周波数を表すパラメータで、音声のホルマント周波数に対応する。これは周波数領域のパラメータであるため良好な量子化特性と補間特性を有し、音声の符号化システムで幅広く用いられている。図2.1の有声音 /a/ のLSPパラメータの例を図2.6に示す。この例では、LSPパラメータがホルマント周波数

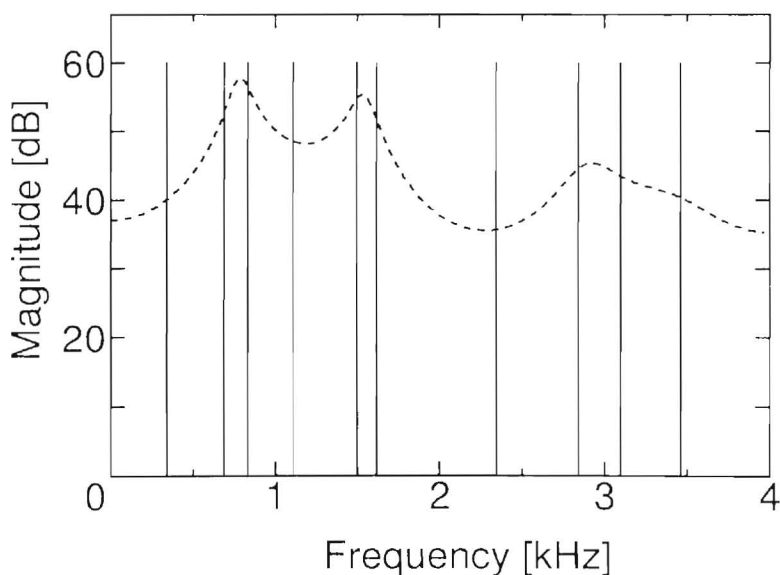


図 2.6 LSP パラメータの例(有声音 /a/)

と良く一致している。

ここで述べたように、ボコーダの励振信号にパルス列とランダムノイズ系列を用いることで、2.4kbps以下の低ビットレートでの音声符号化が可能となる。さらに、ボコーダのスペクトル包絡パラメータをベクトル量子化することで、1kbps以下の極低ビットレートの符号化も実現できる。しかし、この場合の復号音声は明らかに合成音とわかる低品質なものとなってしまったため、低ビットレートで高品質な復号音声を得るために、残差波形を波形符号化により伝送するボコーダと波形符号化のハイブリッド方式の検討が数多くなされている。これらの方式については、2.2.5で述べる。

#### 2.2.4 ベクトル量子化

2.2.2で述べた波形符号化方式が標本値毎に量子化を行なうのに対して、複数の標本値をまとめて一つのベクトルとして符号化する手法をベクトル量子化 (VQ) と呼ぶ。図 2.7 に VQ を用いた符号化システムの構成を示す。符号器と復号器は共通のコードブックを有する。このコードブック

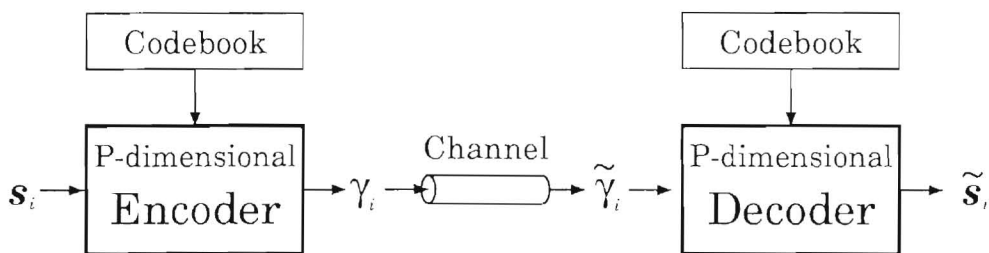


図 2.7 VQ を用いた符号化システムのブロック図

クには、 $N$  個の  $P$  次元コードベクトル  $\{\hat{s}_1, \hat{s}_2, \dots, \hat{s}_N\}$  と各コードベクトルに対応する  $\log_2 N$  ビットの通信路符号  $\{u_1, u_2, \dots, u_N\}$  が含まれる。連続する標本値系列や線形予測モデルのパラメータなどの  $P$  次元ベクトル  $s_i = [s_i(0), s_i(1), \dots, s_i(P-1)]^T$  は、符号  $\gamma_i = u_k$  に写像される。通信路に誤りがない場合 ( $\tilde{\gamma}_i = \gamma_i$ ) には、復号器の出力ベクトル  $\tilde{s}_i$  は  $\hat{s}_k$  となる。

最も基本的なベクトル量子化法は、入力ベクトルを全てのコードベクトルと比較する全探索型のメモリレス VQ である。メモリレス VQ では、入力ベクトル  $s_i$  と全てのコードベクトルの距離尺度を計算し、 $s_i$  に最も距離が近いコードベクトルと対応する符号を量子化結果として出力する。この時の距離尺度としては次の自乗距離がよく用いられる。

$$\delta(s_i, \hat{s}_k) = \sum_{j=1}^P \{s_i(j) - \hat{s}_k(j)\}^2 \quad (2.3)$$

メモリレス VQ では、ベクトルの次数  $P$  とコードベクトル数  $N$  を大きくすることで量子化ひずみを低減することが可能だが、同時に符号化に必要な計算量と記憶量も増加する。ここで 1 サンプルあたりの符号化ビット数を  $B = \log_2 N/P$  と表すと、演算量は  $O(2^{BP})$ 、記憶量は  $O(P \cdot 2^{BP})$  となり、ベクトルの次数と符号化ビット数にたいして指数的に増加する。

効率的な探索を可能にする構造化コードブックを用いることで、量子化時の計算量や記憶量を低減することができる。木探索型 VQ (TSVQ) と多段 VQ (MSVQ) では、量子化ひずみの増加を犠牲にして量子化器の複雑さを抑えることができる。また、コードベクトルを形とゲインに分割して記憶する Gain/Shape VQ (GSVQ) でもメモリレス VQ に比べてベクトル量子化器の複雑さを低減することができる。GSVQ は、近年数多く用いられている符号励振線形予測符号化 (CELP) においても励振信号の符号化に用いられている。さらに、少数のベクトルの線形結合でコードベクトルを生成するベクトル和量子化法でも量子化器の構成を簡易化でき、日本および北米の携帯電話用フルレート音声符号化方式のベクトル和励振線形予測符号化方式 (VSELP) において励振信号の符号化に用いられている。

これまでに述べた手法ではあらかじめ設計したコードブックを固定的に用いるのに対して、符号化を行ないながら適応的にコードブックを更新していく手法も存在する。これらは適応ベクトル量子化と総称され、ベクトル予測量子化 (VPQ) と有限状態ベクトル量子化 (FSVQ) に大別



できる。スカラー量子化における DPCM に相当する VPQ は、CELP や VSELP において励振信号の符号化に用いられており、より高品質な符号化を目的としたベクトル量子化法といえる。一方の FSVQ は、符号器にいくつかの状態が存在するものと仮定し、それぞれの状態においてコードベクトルの部分集合のなかから探索を行なう。さらに、探索後、一定の規則にしたがって符号器の状態を遷移する。これは各状態で小さなコードブックを用いることになり、符号器と復号器の状態遷移が同期していれば、少ないビット数で符号化することが可能である。他に、トレリス VQ などこの種の手法である。ベクトル量子化はボコーダのスペクトルパラメータや励振信号の符号化に応用され、低ビットレート音声符号化システムの構成には不可欠な技術となっている。しかし情報を高能率に圧縮するため、通信路誤りに脆弱となる。よって、ベクトル量子化の通信路誤り耐性の強化は、低ビットレート音声符号化システムにとって重要な課題である。

### 2.2.5 ハイブリッド方式

線形予測の残差信号  $e(n)$  が声道の励振信号に対応することは、2.4 ですでに述べた。この  $e(n)$  で合成フィルタを励振した場合、元の音声信号が完全に復元できる。よって  $e(n)$  を符号化して伝送することで、ボコーダの合成音の品質を改善できる。しかし、 $e(n)$  の符号化には音声信号と同じ情報量を必要とするため、単に波形符号化を行なったのでは高能率な符号化の実現は望めない。そこで様々な手法が検討されている。

残差励振線形予測 (RELP) 符号化システムは、残差信号の低域成分のみを符号化して伝送する手法である。図 2.8 に RELP システムの概要を示す。残差信号  $e(n)$  は声帯から発せられる励振信号に対応するため、有声音の場合にはインパルス列となり調波構造をもつ。しかし、RELP では残差信号の低域成分しか伝送しないため、合成時に何らかの処理によって調波構造を復元する必要がある。このため、復号した低域残差信号を全波整流回路などの非線形回路に通したうえでスペクトルの平滑化を行

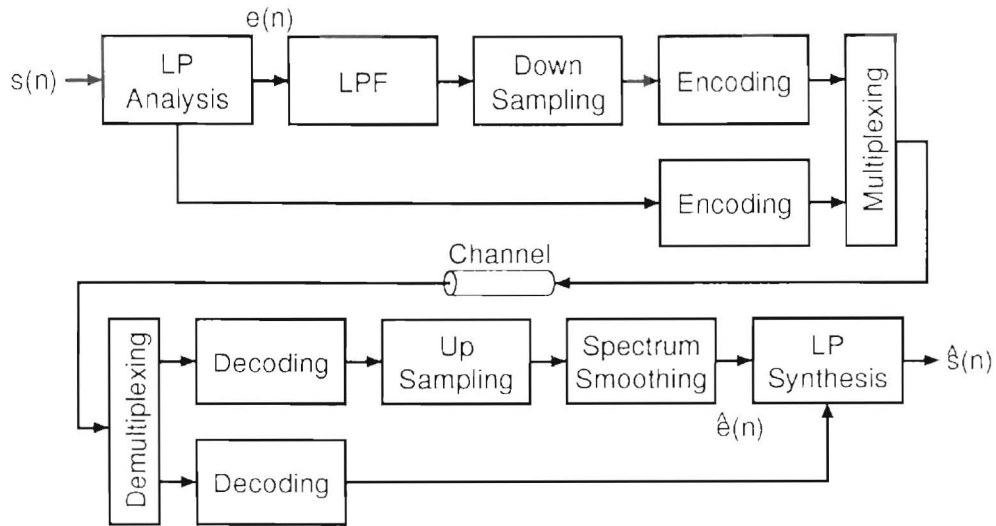


図 2.8 RELP ボコーダの概要

なって合成フィルタの励振信号  $\hat{e}(n)$  とする。RELP システムでは、残差信号  $e(n)$  からピッチ抽出を行なう必要がないため、システム構成が簡易で周囲の雑音に強く、さらにピッチの伝送誤りによる極端な品質劣化は生じないため通信路誤りに対する耐性が強いという利点を持つ。しかし残差信号の低域成分のみから励振信号  $\hat{e}(n)$  を生成するために復号音声の音質が荒々しいものとなり、後述する CELP に比べ復号音声の品質は劣る。

RELP の符号化システムは開ループシステムだが、符号器の中に復号器を持ち閉ループ探索によって残差信号の符号化を行なう手法も存在する。この種のシステムの代表的なものは符号励振線形予測 (CELP) 符号化システムである。図 2.9 に、CELP 符号化器の概要を示す。図中の実線は波形の流れを表し、太い破線は分析したパラメータの流れを示す。CELP では、入力音声と合成音声の間の聴覚的な誤差を最小にするように符号化のパラメータを決定する。このような手法は、Analysis by Synthesis (AbS) と呼ばれる。また、励振信号の符号化には適応型の Gain Shape VQ が用いられている。CELP では聴覚的な誤差を最小化するため非常に高品質な復号音声を得られるが、図からもわかるようにシス

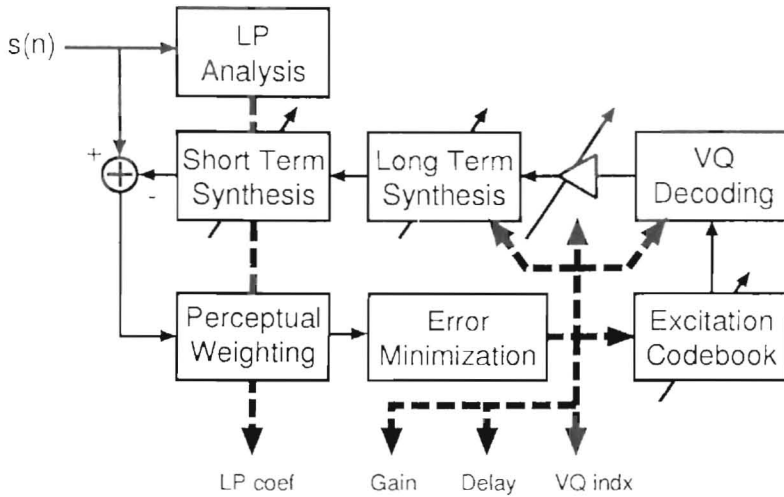


図 2.9 CELP 符号化器の概要

テム構成が複雑である。また伝送する情報が多いため、従来 CELP で音声通信に必要な品質を達成する場合のビットレートの下限は 9kbps といわれていたが、現在では 4kbps 弱での符号化も可能となり<sup>(52,60)</sup>、4～8kbps 程度の音声符号化方式において広く実用化されている。CELP の変形方式である VSELP も、携帯電話用の音声符号化システムとして実用化されている。

ここで述べた RELP と CELP 以外にも、マルチパルス励振線形予測 (MPLP) 符号化方式<sup>(49)</sup>や正規化パルス励振 (RPE) 符号化方式<sup>(49)</sup>など、ボコーダと他の手法の様々なハイブリッド方式が数多く提案されている。低いビットレートで音声通信に必要な品質の復号音声を得るためにはこれらのハイブリッド方式が不可欠であり、今後ますます重要な技術となっていくものと思われる。

### 2.3 音声符号化システムにおける通信路符号化方式

前節で述べた様々な音声符号化方式の主要な応用先として、移動体音声通信システムがあげられる。特に重要な応用先である携帯電話などの陸上移動体通信では、電波がビル壁などによって反射され多くの遅延波



が生じ、また局の移動によってドップラー効果が発生する (図 2.10)。こ

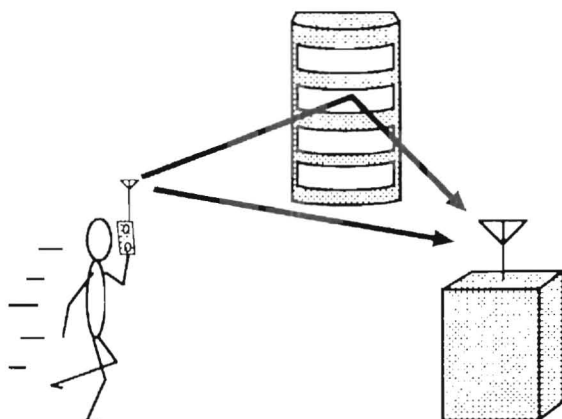


図 2.10 陸上移動体音声通信

れらの影響により、通信路の特性が時間とともに大きく変動するフェージングが発生するため、陸上移動体通信では劣悪な通信環境の下で通信を行なわなくてはならない。この劣悪な通信環境に対処するため、音声符号化システムにおいても通信路誤り対策を施す必要がある。以下では、まず誤り検出と訂正の原理を簡単に述べた後に、音声符号化システムにおける従来の通信路誤り対策について概観し、それらの問題点を述べる。

### 2.3.1 誤り検出と訂正の原理

いま、A と B の2つのシンボルを出力する情報源からの出力を符号化し、誤りのある通信路を通して伝送する場合を考える。もっとも単純な符号化法は、シンボル A に“1”，シンボル B に“0”を割り当て、1ビットで符号化する方法である。しかしこの場合には、通信路で1ビットの誤りが重畳すると、シンボル A は B に、B は A に誤ってしまい、受信側では誤りの検出さえ不可能である。次に2ビットで符号化し、シンボル A に“11”，シンボル B に“00”を割り当てる場合を考える。この場合には1ビットの誤りが重畳すると受信符号は“01”または“10”となり、これらの符号に対応するシンボルは存在しないために誤りが重畳したことを検出できる。しかし、“11”と“00”のどちらの符号が誤ったかはわ

からず、さらに2ビットの誤りが重畳した場合の誤り検出は不可能である。そこでシンボルAに“111”，シンボルBを“000”として3ビットで符号化すると、シンボルAを1ビット誤った場合に受信される“011”，“101”，“110”とシンボルBを1ビット誤った場合に受信される“100”，“010”，“001”は互いに重複していないため、1ビットの誤りは訂正可能となる。またそれぞれのシンボルを2ビット誤った場合にも送信符号の“111”と“000”にはならないため、2ビットの誤りは検出可能である。これをまとめると図2.11のように表すことができる。すなわち、通信路誤

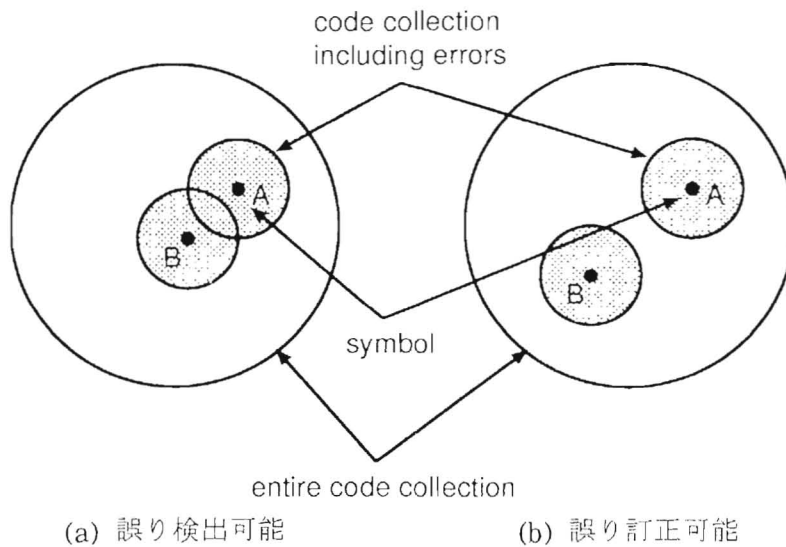


図 2.11 誤り検出と訂正の原理

りが加わった受信符号の集合が互いのシンボルに割り当てられた符号を含まなければ誤りの検出が可能で、さらにそれらの集合が互いに重複しなければ誤り訂正が可能となる。一般に、誤りの検出を目的とした符号化法を誤り検出符号化、訂正を目的としたものを誤り訂正符号と呼ぶ。次に、これらの符号化法が音声符号化でどのように用いられるかについて述べる。

### 2.3.2 誤り検出符号の音声符号化システムへの応用

音声は人間が声帯波で声道を励振して得られるものであり、人間の声道の時間的な変化は緩やかであるため、2.2.1で述べた音声のスペクトルの時間的な変化は緩やかなものとなる。このため、ボコーダにおけるスペクトルパラメータの符号化などでは、通信路で誤りが重畳した符号から音声を復号するよりも、直前の復号結果を現在の出力としたほうが良好な品質の音声を得られると考えられる。これを利用して、情報源符号を誤り検出符号化し、復号時に誤りが検出された場合には一時刻前の復号結果を出力することで、通信路誤りに対処する手法が検討されている<sup>(29,32)</sup>。この手法を CELP に適用した場合、通信路に誤り率が  $10^{-2}$  程度のビット誤りがある場合に、復号音声の主観的な品質が改善したと報告されており<sup>(29)</sup>、通信路の誤り率が比較的低い場合には少ない冗長ビットの付加で通信路誤り耐性が強化できる有効な手法である。

しかし通信路のビット誤り率が高く、いくつもの符号に連続して誤りが重畳する場合には、誤り検出符号化による手法では数十 ms にわたってほぼ同じスペクトルを持つ波形が復号されてしまい、復号音声が不明瞭になってしまう。このような場合には、次節で述べるように誤り訂正符号化を用いて通信路で符号に重畳した誤りを訂正する必要がある。

### 2.3.3 誤り訂正符号の音声符号化システムへの応用

誤り訂正符号化の最も簡単な利用法は、全ての符号ビットを対等に扱い、一括して誤り訂正符号化を行なうことである。この場合、全ての符号ビットが一様に誤りから保護されることになる。しかし実際には、誤りが復号音声に与える影響は符号ビット列のビット位置によって異なるため、一様な誤り訂正符号化は非効率的である。そこでまず、情報源符号器の出力ビット列中の各ビットを誤った場合に復号音声が受ける品質劣化の度合をセグメンタル SNR やスペクトルひずみなどによって調べ、これを通信路誤りに対する誤り感度とする。次に、この誤り感度にしたがって



符号ビット列をいくつかのグループに分類し、誤り感度の高い符号ほど多くの冗長ビットを追加して強力な誤り訂正符号化を行ない、誤り感度の低い符号は全く通信路符号化を行なわないか誤り検出符号化のみを行なうなど、非一様な誤り訂正符号化を用いることで、全てのビットを一様に誤り訂正符号化する場合に比べて少ない冗長ビットの追加で効率良く誤り耐性が強化できる。このような手法は Suda らによって導入され<sup>(27)</sup>、その後数多くの検討がなされ<sup>(28,30,32,34)</sup>、実際の音声符号化システムにおいても用いられている<sup>(52)</sup>。

この他には、通信路の誤り率に応じて情報源符号と通信路符号のビット配分を変えた符号化システムを適応的に切替える手法が、CELP をベースシステムとして検討されている<sup>(31)</sup>。この方式では、通信路符号化を行なわない同じビットレートの CELP システムと比べて、通信路誤りがある場合の復号音声のセグメンタル SNR が 2dB から 3dB 改善したと報告されている<sup>(31)</sup>。しかし、複数の符号化システムを切替えて用いるためシステムが複雑になるうえに、通信路の状態(誤り率)を推定しなくてはならないなどの欠点がある。

ここで述べた手法は、ビットレートが 5 kbps 程度までは有効だが、それ以下のビットレートでの符号化を行なう場合には、付加する冗長ビットが多いことが問題となる。冗長ビットの数を少なく抑えるためには、情報源符号自体の誤り耐性を強化する必要がある。次節でこれらの手法について述べる。

#### 2.3.4 音声符号化システムにおけるその他の通信路符号化方式

これまでに述べた方式とは異なり、誤り検出符号化や訂正符号化を行なうことで誤り耐性を強化するのではなく、通信路で重畳するビット誤りの影響をあらかじめ考慮して情報源符号化を行なうことで、情報源符号の誤り耐性の強化を行う手法が数多く検討されている<sup>(35,41-43,46)</sup>。これらの手法では、通信路誤りを考慮したひずみを定義し、情報源符号化によるひずみと通信路誤りによるひずみを同時に最小化するような符号化

を行なう。このため、通信路に誤りがないうちにも情報源符号化によるひずみを最小にすることはできないが、通信路に誤りがある場合には同じビットレートの通信路誤りを行なわないシステムに比べて大幅にひずみを低減できる<sup>(35,39)</sup>ことが報告されている。

符号化時に通信路誤りを考慮するだけでなく、符号器の設計段階から通信路誤りを考慮して最適な符号化システムを構成することを意図した検討も行なわれている<sup>(35,40,41,44,48)</sup>。Farvardin は、ベクトル量子化した符号に1ビットの誤りが生じると仮定してひずみを定義し、このひずみが最小になるようにコードベクトルと符号の対応をシミュレーテッドアニーリングによって決定する手法を提案し、さらに通信路誤りを考慮した符号化を行なう手法と組み合わせた場合の検討も行なっている。この手法を多段ベクトル量子化やトレリスベクトル量子化などの構造化したベクトル量子化法に適用する拡張が、Phamdoらによって提案されている<sup>(42,43)</sup>。またMoriya は、ベクトル量子化のコードブックを2つに分割してそれぞれのコードブックからのコードベクトルの和でベクトルを表わし、さらに通信路誤りを考慮したひずみを最小化するように符号化することで通信路誤りの影響を低減する手法を提案している<sup>(44)</sup>。この手法は、日本の次世代携帯電話用音声符号化システムの一部に用いられている。

ここで述べた手法は、符号に誤りが重畳した場合の影響を最小限に抑えることが目的であり、単独で用いるには誤り耐性が不十分であり、低ビットレートの音声符号化システムに用いるには、より誤り耐性の強い情報源 - 通信路統合符号化法が必要になると考えられる。

## 2.4 本章のまとめ

本章では、音声の情報源符号化方式とその従来の通信路符号化方式について述べた。まず、音声の情報源符号化方式を波形符号化、ポコード方式、ベクトル量子化、ハイブリッド応用方式に分類して概観した。表2.1にこれらの音声符号化方式についてまとめた。次に音声符号化シス

テムにおける従来の通信路符号化方式として、誤り検出符号化や誤り訂正符号化による誤り制御法や、誤りの影響を低く抑えるための情報源-通信路統合符号化法について概説し、低ビットレート音声符号化システムへの適用を考慮した場合のそれらの手法の問題点について述べた。これらの音声符号化における従来の誤り対策法とそれらの問題点については、表 2.2 にまとめる。ここで述べたように、4kbps 以下の低ビットレート音声符号化システムを移動通信で用いるためには、従来よりも高能率で誤り耐性が強い符号化方式が必要となる。次章では、この問題に対する一対処法として、高能率で強力な誤り訂正能力を持つ新しい情報源-通信路統合符号化方式について述べる。

表 2.2 音声符号化システムの従来の通信路誤り対策

	手法	問題点
誤り検出符号化の利用	前のフレームの情報を利用する。	符号に連続して誤りが重畳した場合、復号品質が極端に劣化する。
誤り訂正符号化の利用	誤り感度に応じて非一様な誤り訂正符号化を行なう。	付加する冗長ビットが多過ぎる。
情報源-通信路統合符号化の利用	通信路誤りによるひずみを考慮して送信符号を決定。	符号化時の計算量が多くなる。誤り耐性が不十分。
	誤りの影響が少なくなるように符号を設計する。	誤り耐性が不十分。

表 2.1 音声符号化方式の分類

	符号化方式	ビットレート	伝送する情報	標準化/応用先, 特徴 など
波形符号化方式	PCM	32 ~ 64 kbps	標本値	CCITT G.711, デジタル基幹回線網などで使用.
	ADPCM	16 ~ 32 kbps	標本値の差分	CCITT G.721, 32kbps システムが PHS 用符号化方式に採用.
	ATC	16 ~ 32 kbps	標本値の DFT 係数や DCT 係数	16kbps で PCM にくらべ SNR が 16 ~ 25dB 改善.
	VQ		標本値の VQ インデックス	CELP などと長期予測後の残差信号の符号化に利用
分析合成方式	LPC-10	2.4 kbps	線形予測係数, ピッチ, ゲイン, 有 声 / 無声	アメリカ合衆国政府標準 FS1015(2.4kbps)
	LSP-VQ	0.8 kbps	線形予測係数の VQ インデックス, ピッチ, ゲイン, 有聲 / 無声	線形予測係数をベクトル量子化することで極低ビットレートを實現.
ハイブリッド方式	RELTP	4.8 ~ 9.6 kbps	線形予測係数, 低域残差波形, ゲイ ン	残差波形の低域成分を符号化して伝送. 2 状態励振の LPC 方式より 高品質.
	CELP	3.6 ~ 16 kbps	線形予測係数, 長期予測係数, ビッ チ, 長期予測残差の VQ インデック ス, ゲイン $\times 2$	FS1016(4.8kbps), 次世代日本・北米デジタルセルラ用ハーフレー ト符号化システム.
	VSELP	6 7/8 kbps	線形予測係数, 長期予測係数, ビッ チ, 長期予測残差の VQ インデック ス $\times 2$ , ゲイン $\times 3$	日本・北米デジタルセルラ用フルレート符号化システム, 2 つのベ クトルの和で長期予測残差を表現.
	MPLP	4.8 ~ 16 kbps	線形予測係数, マルチパルス音源の パルス振幅と位置, 最大振幅値	Skyphone (British Telecom International, UK) で採用. 残差波形 を複数のパルス列で表現.
	RPE	6.4 kbps	線形予測係数, 正規化パルス音源の パルス振幅	Group Special Mobile 汎ヨーロッパパルパデジタル移動通信. 一様な間 隔でパルスが並ぶ複数のパルス列の和で残差波形を表現.



## 第3章 遷移を制約したベクトル量子化法

### 3.1 まえがき

本章では、VQを用いた新しい音声の情報源 - 通信路統合符号化方式として遷移を制約したベクトル量子化を提案し、その性能評価を行なう。

最も基本的なメモリレスVQを符号化に適用する場合には、各時点においてコードブックに含まれる全てのコードベクトルの中から探索を行なうことになる。これは、コードベクトル間の任意の遷移が許されていることを意味する。この時、符号化対象のベクトルとして記憶のない情報源からの出力を仮定している場合には、任意のコードベクトル間の遷移がほぼ等確率で発生する。

一方、音声信号は人の声道を声帯の振動で励振することで発せられる。この時、声帯や声道は急激に変化することはできないため、発生される音声信号波形の変化は比較的緩やかで過去の信号の影響を受けたものとなる。このため、音声波形やそのスペクトルパラメータを短時間毎にメモリレスVQを適用して得られる符号系列には時間的な冗長(連鎖性)が残存し、符号ベクトル間の遷移確率にはバラツキが存在する。この時間的な冗長を低減することでメモリレスVQに比べて高能率な符号化を実現する手法に2.2.4で述べたFSVQなどがあるが、これらの手法とは逆に、この冗長を利用することで通信路誤りに対する耐性を強化することも可能と考えられる。本章では、この冗長を利用して通信路誤りの耐性を強化することを目的とした新しいベクトル量子化法、遷移を制約したベクトル量子化を提案する。この手法を用いることで、通信路符号化によって明示的な冗長ビットを追加して誤り保護を行なわなくとも、通信路で重畳した誤りを復号器において訂正可能となる。

本章ではまず、遷移を制約したベクトル量子化法の符号化/コードブックデザイン/誤り訂正復号の各アルゴリズムについて述べる。次に、音声の線形予測ケプストラムを誤りのある通信路を通じて符号化/伝送するシミュレーションによって提案する手法の性能評価を行なう。



### 3.2 トレリスの状態にラベル付けを行なう遷移を制約したベクトル量子化 (LS-CTVQ)

図 3.1 に、コードベクトル数が 4 の場合の通常のメモリレス VQ のトレリス表現を示す。図中の  $\circ$  はトレリスの状態を表し、それぞれがコードベクトルに対応する。メモリレス VQ では任意のコードベクトル間の遷移が可能なので、トレリス上においても任意の状態間の遷移が可能である。メモリレス VQ を符号化に用いる場合、送信符号系列 (図中の太い実線) とそれに通信路で誤りが重畳した受信符号系列 (太い破線) のどちらも、トレリス上の一本の連続したパスとして表現可能である。このため、このトレリスを用いても復号器において通信路誤りの検出や誤り訂正を行なうことは不可能である。

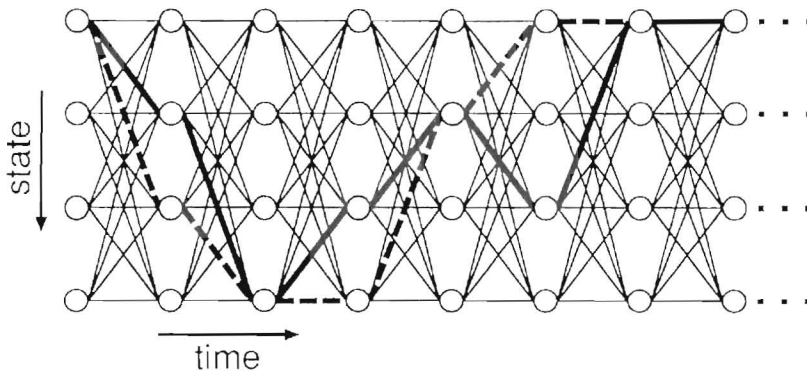


図 3.1 通常のメモリレス VQ のトレリス表現

しかし、記憶のある情報源の出力を符号化対象としている場合には、コードベクトル間の遷移確率にはバラツキが生じる。そこで、ある 2 つのコードベクトル間の遷移確率が小さい場合には、その遷移は起こらないものとして遷移を禁止する。この場合のトレリス表現を図 3.2 に示す。コードベクトル間の遷移を制限したために、図 3.1 とは異なり、トレリスの状態間の遷移も制限される。この場合、符号器が出力可能なコードベクトル系列が制限されるために、メモリレス VQ と比べて量子化ひずみみがかくらか増大する。しかし、通信路で誤りが重畳した符号系列が表現するコー

ドベクトル系列はこのトレリス上では連続した一本のパスとして表現できない。このため、復号時にトレリス上で不連続になるコードベクトル系列を受信した場合には、通信路で重畳した誤りを検出することが可能であり、さらに遅延させた後に復号判定を行なうことで通信路誤りを訂正することもできる。これは、確率の小さい遷移を制約することで生じる量子化ひずみみの増加を犠牲にして確率の大きい遷移に重畳する通信路誤りを低減し、全体的な復号品質の向上を目指した手法ともいえる。このように、コードベクトル間の遷移を制約することで通信路誤りの訂正を可能にするベクトル量子化法を、遷移を制約したベクトル量子化 (Constrained Transition Vector Quantization : 以下 CTVQ) と呼ぶことにする。本章で扱う手法は、図 3.2 のようにコードベクトルとトレリスの状態が対応しているため、これを特に状態ラベル型 CTVQ (Labeled-State CTVQ : 以下 LS-CTVQ) と呼ぶことにする。

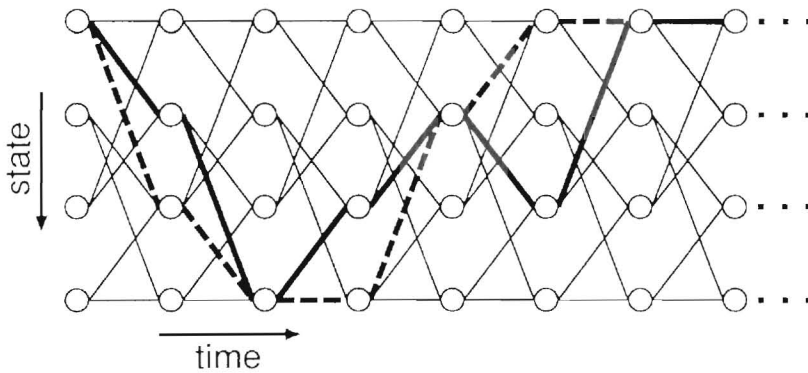


図 3.2 状態ラベル型 CTVQ のトレリス表現

### 3.2.1 符号化アルゴリズム

$x_t$  と  $y_t$  を時点  $t$  における情報源出力ベクトルとその量子化ベクトル、 $\gamma_t$  を符号器が出力する符号とする。この時、符号化は次の手順で行なう。

1.  $y_0 = c_1, \gamma_0 = b_1, t = 1$  とする。

2. 時点  $t-1$  における量子化出力  $\mathbf{y}_{t-1}$  から遷移可能なコードベクトルの集合を  $\Gamma_{t-1}$  として、この中から量子化ベクトル  $\mathbf{y}_t = q_c(\mathbf{x}_t)$  の探索を行なう。ここで  $q_c(\cdot)$  は LS-CTVQ を表し、次式で定義される。

$$q_c(\mathbf{x}_t) = \underset{\mathbf{c}_j \in \Gamma_{t-1}}{\operatorname{argmin}} \{ \delta(\mathbf{x}_t, \mathbf{c}_j) \} \quad (3.1)$$

3. 量子化出力  $\mathbf{y}_t = \mathbf{c}_i$  に対応する符号  $b_i$  を時点  $t$  における出力符号  $\gamma_t$  とする。この符号を LS-CTVQ 符号と呼ぶ。
4.  $t = t+1$  として、2. へもどる。

通常のリモレス VQ との違いは、コードブックに含まれるすべてのコードベクトルの中から探索を行なうのではなく、各時点で遷移可能なコードベクトルの部分集合の中から探索を行なう点である。このため探索に必要な計算コストはリモレス VQ に比べて減少するが、各コードベクトルから他のコードベクトルへの遷移の可否を記憶しておく必要があるため記憶コストは増加する。

LS-CTVQ で用いるコードブックは次のように表される。

$$\mathcal{C}_c = \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 & \dots & \mathbf{c}_N \\ b_1 & b_2 & \dots & b_N \\ \mathbf{t}_1 & \mathbf{t}_2 & \dots & \mathbf{t}_N \end{bmatrix} \quad (3.2)$$

$\mathbf{t}_i = [t_{i1}, t_{i2}, \dots, t_{iN}]$  はコードベクトル  $\mathbf{c}_i$  からの遷移の可否を 1, 0 で表現した行ベクトルで、 $t_{ij} = 1$  はコードベクトル  $\mathbf{c}_i$  から  $\mathbf{c}_j$  へ遷移可能、 $t_{ij} = 0$  は遷移不可能なことを表す。この  $\mathcal{C}_c$  を遷移を制約したコードブックと呼ぶことにする。

### 3.2.2 遷移を制約したコードブックの作成

遷移を制約したコードブック  $\mathcal{C}_c$  の作成にはコードベクトル間の遷移確率が必要となる。いまコードベクトル間の遷移確率をコードベクトル  $\mathbf{c}_j$  に引き続いて  $\mathbf{c}_k$  が出力される条件付き確率として定義し、 $p_{kj} = p(\mathbf{c}_k | \mathbf{c}_j)$  で表す。コードベクトルと符号が 1 対 1 に対応するので、 $p_{kj}$  は符号  $b_j$  か

ら  $b_k$  への遷移確率  $p(b_k|b_j)$  でもあることに注意する. 遷移確率  $p_{kj}$  の推定には隠れマルコフモデル (HMM)<sup>(65,66)</sup> の推定法である Baum-Welch アルゴリズム<sup>(66)</sup> を用いた. 推定した遷移確率  $p_{kj}$  に従ってコードベクトル間の遷移に制約を与える. 制約の与え方としては様々なものが考えられるが, 今回は各コードベクトルからの遷移確率が大きいものから順に  $M$  個のコードベクトルにのみ遷移可能とした. 以後, この  $M$  を遷移数と呼ぶことにする.

遷移を制約したコードブック  $C_c$  の作成は次の手順で行なう.

1.  $N$  個のコードベクトルを持つメモリレス VQ 用のコードブック  $C$  を作成する.
2. コードブック  $C$  のコードベクトルを状態に対応させた. 無相関単一正規分布エルゴディック HMM<sup>(65)</sup> を作成する. モデルの出力確率は, 各コードベクトルを平均ベクトルとする無相関正規分布で表現し. 正規分布の分散および状態間の遷移確率は乱数で初期化する.
3. Baum-Welch アルゴリズムにより, モデルの出力確率の正規分布の分散と状態遷移確率の推定を行なう. この時, 正規分布の平均 (コードベクトル) は変化させない.
4. 推定された状態遷移確率  $p_{kj}$  をコードベクトル  $c_j$  からの遷移確率とし, 各状態毎に確率の大きい方から  $M$  個の遷移を残し ( $t_{jk} = 1$ ), 残りの遷移を禁止する ( $t_{jk} = 0$ ).

### 3.2.3 LS-CTVQ 符号のビット誤り訂正法

LS-CTVQ を用いて符号化を行なった場合, コードベクトル間の遷移が制約されているために, 符号間の遷移も制約される. ここで, 通信路において符号に誤りが重畳した場合には, 復号器はトレリス上では存在しない符号系列を受信する可能性がある. そこで符号の遷移を, 図 3.2 のコードベクトルの場合と同様にトレリスで表し (図 3.3), このトレリス上



で表現可能な符号系列のうち、受信符号系列との距離が最小になるものを復号系列とすることで、通信路で符号に重畳したビット誤りが訂正可能になる。

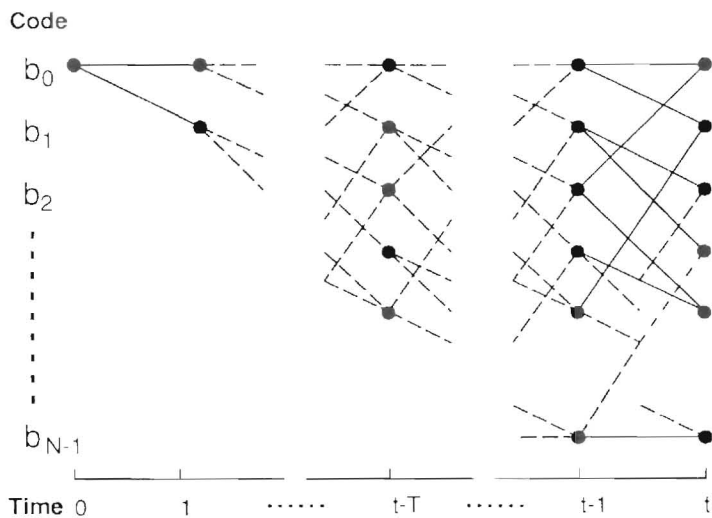


図 3.3 符号間の遷移のトレリス表現

図 3.4 のように時点  $t - T$  で符号  $b_j$  を出力し、時点  $t$  で符号  $b_k$  を出力する符号系列を考える。実線は時点  $t$  までの受信系列との距離が最小の符号系列である。この時、この最小距離符号系列は時点  $t - T$  までの距離も最小でなくてはならない。そこで、各時点でそれぞれの符号を出力する符号系列の中で受信系列との距離が最小のもの(生き残り系列)だけを残し、すべての符号を受信した時点で距離が最小の系列を受信符号系列の判定結果とする。この操作は、符号がトレリスの枝ではなく節点に割り当てられていることを除けば畳み込み符号のビタビ復号と同じであり、コードベクトル間の遷移を制約することによって、ビタビアルゴリズムを用いた誤り訂正が可能になることを意味する。しかし、すべての符号を受信した後に受信符号系列の判定を行なうと、生き残り系列の記憶に必要なメモリ量と復号判定遅延が膨大なものになってしまう。このため畳み込み符号のビタビ復号と同様に、適当な判定遅延  $l_d$  の後に受信符号の判定をしなくてはならない。



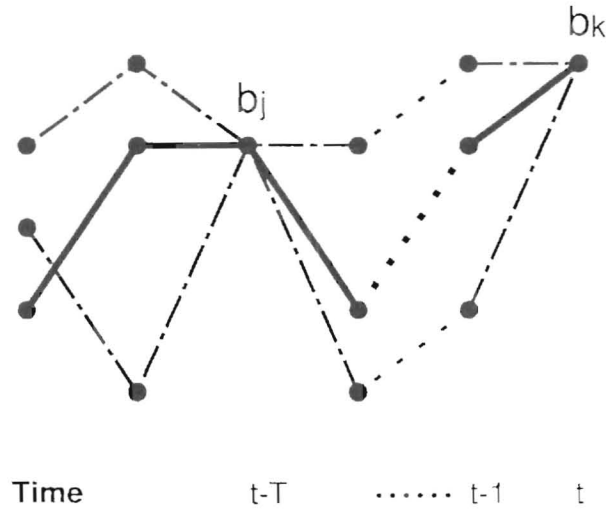


図 3.4 符号間の遷移のトレリス表現

いま時点  $t$  の受信符号を  $\tilde{\gamma}_t$  と表す。この時、時点  $t$  までの受信符号系列は  $R_t = \{\tilde{\gamma}_1, \tilde{\gamma}_2, \dots, \tilde{\gamma}_t\}$  と表される。ここで、時点 0 で受信する符号は  $b_1$  と仮定し、すべての符号系列は  $b_1$  で始まるものとする。また時点  $t$  で符号  $b_j$  を出力する符号系列のうち、時点  $t$  までの受信系列との距離が最小の符号系列を生き残り符号系列  $S_{t,j}$  とする。この時、次の手順で復号を行なうことで、受信符号系列の誤り訂正が可能である。

1. 時点 0 での生き残り系列  $S_{0,1} = \{b_1\}$ , 時点  $t = 1$  とする。
2. 時点  $t - 1$  の生き残り系列の中で時点  $t$  に符号  $b_i$  に遷移可能なすべての系列と時点  $t - 1$  までの受信系列との距離を比較し、距離が最小の系列  $S_{t-1,j}$  を選択する。
3.  $S_{t-1,j}$  に符号  $b_i$  を連結したものを時点  $t$  の符号  $b_i$  における生き残り符号系列  $S_{t,i}$  とする。この時、時点  $t$  までの受信系列  $R_t$  と  $S_{t,i}$  の距離  $\lambda_{t,i}$  は、時点  $t$  までの受信系列  $R_{t-1}$  と  $S_{t-1,j}$  の距離  $\lambda_{t-1,j}$  に  $\tilde{\gamma}_t$  と  $b_i$  の距離を加えたものとする。
4.  $t > \ell_d$  ならば距離  $\lambda_{t,i}$  が最小の生き残り符号系列  $S_{t,i}$  の時点  $t - \ell_d$

における符号を時点  $t - l_d$  の受信符号の判定結果とする.

5.  $t = t + 1$  として2. へもどる.

このようにして得られる符号系列を復号系列とし、対応するコードベクトルを復号ベクトル系列とすることで、通信路で重畳したビット誤りによって生じるコードベクトル誤りを減少し、これに起因して生じる復号品質劣化を低減することが可能となる.

ビタビアルゴリズムでは、受信符号系列とトレリスのパスで表現される符号系列の間の距離尺度には様々な尺度を適用することが可能である. ここでは代表的なものとしてハミング距離とユークリッド距離の二つについて検討を行なう. ここで畳み込み符号のビタビ復号と同様、ハミング距離を用いる場合を硬判定復号、ユークリッド距離を用いる場合を軟判定復号と呼ぶことにする. 硬判定復号ビタビ復号器のブロック図を図 3.5 に、軟判定復号ビタビ復号器のブロック図を図 3.6 に示す. 図中の

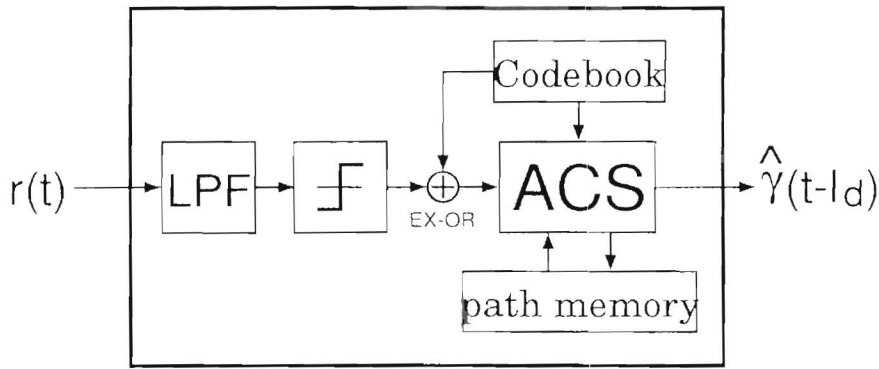


図 3.5 硬判定ビタビ復号器のブロック図  
ACS: Add Compare Select.

ACS は Add Compare Select 回路の略で、ビタビアルゴリズムの距離尺度の加算と比較をおこない、最小の距離尺度のパスを選択するためのものである.

例として、図 3.7 のような 00, 01, 10, 11 の 4 つの符号を出力する符号器を考える. 図中の○は符号を、弧は許された遷移を表す. また、ここで

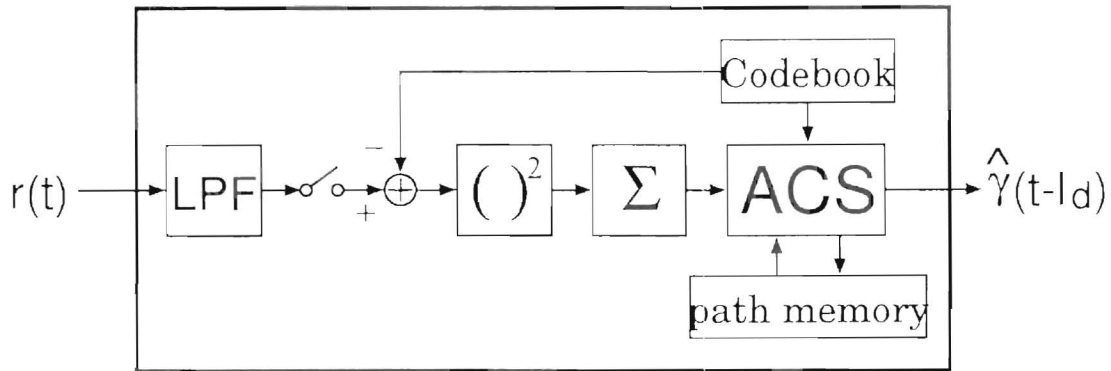


図 3.6 軟判定ビタビ復号器のブロック図

ACS: Add Compare Select

は符号間の距離尺度としてはハミング距離を仮定する. 図 3.7 の符号器からは, 符号 00 の次には 00 か 01 しか出力されないため, 復号器において符号 00 に続いて 10 や 11 が受信された場合には誤りが生じたことが検出できる.

図 3.7 の符号器の出力する符号系列のトレリス上における, 誤り訂正の例を示したものが図 3.8 である. 実線の矢印は各時点でそれぞれの符号を出力する生き残り符号系列  $S_{t,i}$ , 破線の矢印は捨てられた遷移, ○の中の数字はそこへ到達する生き残り系列が表現する符号系列と受信符号系列とのハミング距離  $\lambda_{t,i}$  である. 始めの 2 つの符号に誤りが重畳しているが, 時点 3 以降において  $\lambda_{t,i}$  が最小の符号 (例えば時点 5 の符号 01) から生き残り符号系列をたどる (太い実線の系列) ことで誤りが訂正され, 元の符号系列が復元されることがわかる.

このように LS-CTVQ を用いて符号化を行なった場合には, ビタビアルゴリズムを適用することで通信路において重畳したビット誤りが訂正可能となる. 次節では, 音声の線形予測ケプストラムを符号化する計算機シミュレーションを通じて, LS-CTVQ の能力を検証する.

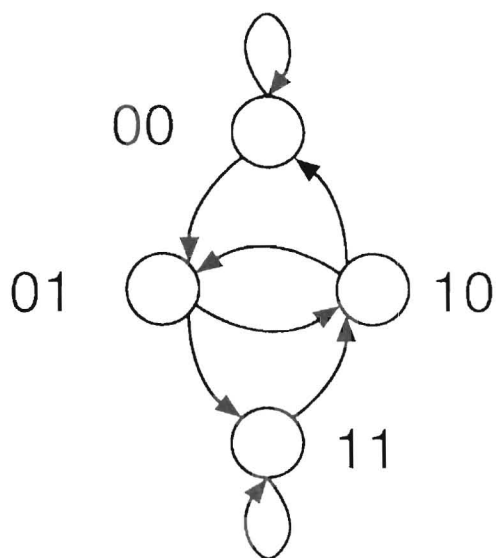


図 3.7 遷移が制約された符号の例

Send seq.	00	01	11	10	01
Error seq.	10	10	00	00	00
Received seq.	10	11	11	10	01
Decoded seq.	00	01	11	10	01

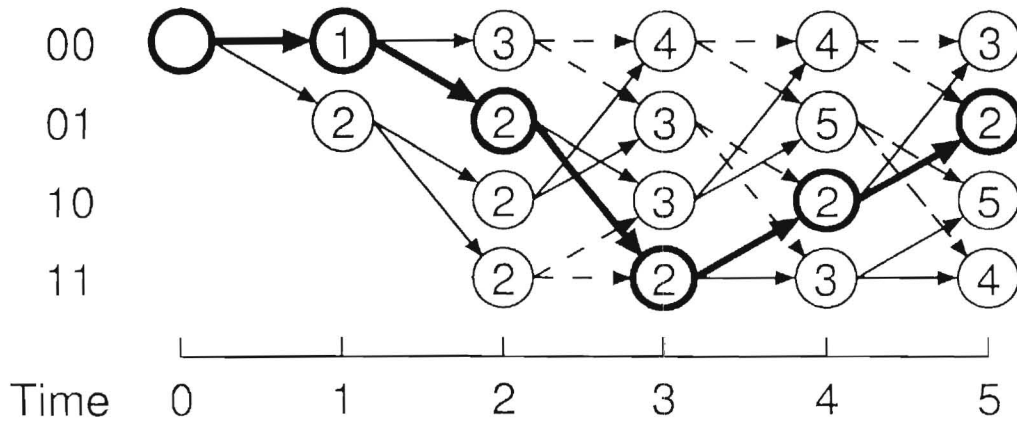


図 3.8 トレリス上での誤り訂正の例

表 3.1 シミュレーション諸元

音声資料	男性2名, 女性4名の発声した50短文
標本化	8kHz, 8bit
分析次数	P=12
分析フレーム長	30ms
分析周期	5ms
エンファシス	なし

### 3.3 計算機シミュレーション

本節では, 種々の音声のスペクトルパラメータのうち, 良好な補間特性を有し, その自乗誤差として定義されるケプストラム距離がヒトの聴覚特性に良く適合しているといわれる線形予測ケプストラムを LS-CTVQ によって符号化し, 符号にビット誤りを重畳したあとに復号する計算機シミュレーションを通して LS-CTVQ の能力を検証する.

本シミュレーションの主な諸元を表 3.1 に示す. 音声資料のうち男女4名の発声分 (計 200 短文) をコードブックデザイン用に用い, 残りの女性2名の発声分 (計 100 短文) によって評価を行った.

ベクトル量子化時のベクトル間の距離尺度としては自乗距離を用いた. すなわち,

$$\delta(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^P \{x(i) - y(i)\}^2 \quad (3.3)$$

である. ここで,  $x(i), y(i)$  はベクトル  $\mathbf{x}, \mathbf{y}$  の  $i$  番目の要素を表す.

#### 3.3.1 コードベクトル間の遷移を制約する影響

ケプストラムは音声のスペクトル包絡を表現するパラメータであり, これをベクトル量子化する場合, コードベクトル系列はスペクトル包絡の



変化を表す. このためコードベクトル間の遷移を制約することは, 表現可能なスペクトルの変化を制限することを意味し, 遷移数を減らすほど量子化ひずみが増加するものと思われる. そこで, LS-CTVQ においてコードベクトル間の遷移を制約することによる影響を調べるために, 遷移数  $M$  によるコードブックのエントロピー  $H_1$  と復号後のケプストラムひずみ  $D_{cep}$  の変化を調べた. この結果を図 3.9 に示す. 図 3.9 の横軸は, 遷移数  $M$  の 2 を底とする対数である.

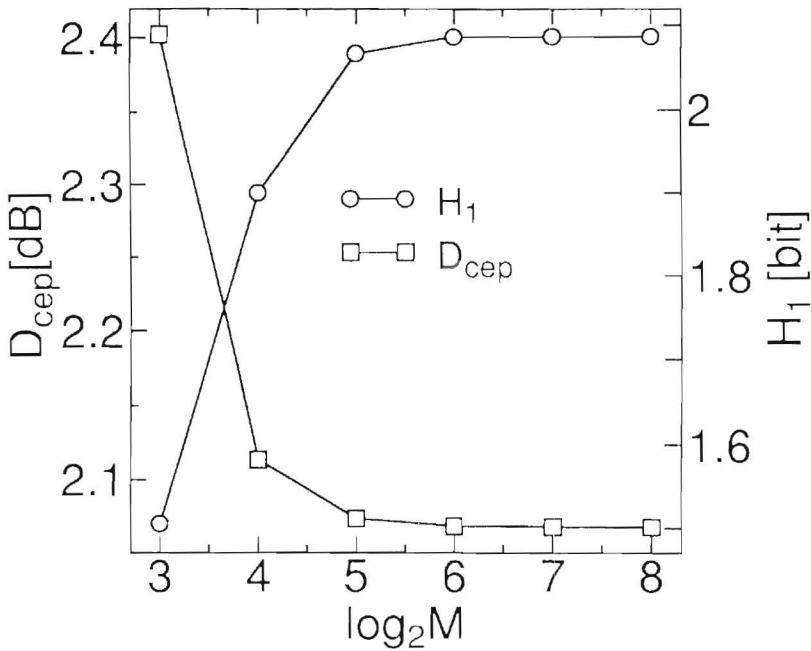


図 3.9 遷移数によるエントロピーと復号ケプストラムひずみの変化

ここでコードブックのエントロピーは, 遷移を制約したコードブック  $\mathcal{C}_j$  をコードベクトル  $\mathbf{c}_k$  の発生確率  $p_k = p(\mathbf{c}_k)$  と遷移確率  $p_{kj} = p(\mathbf{c}_k | \mathcal{C}_j)$  が与えられる, 単純マルコフモデルとみなして推定を行なった.

$$H_1 = - \sum_{i=1}^N p_i \sum_{j=1}^M p_{ji} \log p_{ji} \quad (3.4)$$

ここで, コードベクトル間の遷移確率  $p_{kj}$  はコードブックの作成時に推定したものを用い, 各コードベクトルの発生確率  $p_k$  は音声資料をベクト

ル量子化して求めたヒストグラムより推定を行なった.

$$p_k = p(\mathbf{c}_k) = \frac{n_k}{N} \quad (3.5)$$

$$\sum_{\ell=1}^N n_{\ell}$$

ただし  $N$  はコードブックに含まれる全コードベクトル数.  $n_k$  はコードベクトル  $\mathbf{c}_k$  に量子化されたベクトルの数である.

またケプストラムひずみ  $D_{cep}$  は次式で与えられる.

$$D_{cep} = \overline{d(\mathbf{x}_t, \mathbf{y}_t)} \quad (3.6)$$

$$d(\mathbf{x}_t, \mathbf{y}_t) = \frac{10}{\ln 10} \sqrt{2 \sum_{i=1}^P (x_t(i) - y_t(i))^2} \quad (3.7)$$

ここで,  $\mathbf{x}_t, \mathbf{y}_t$  は時点  $t$  におけるケプストラムベクトルとその量子化ベクトルであり,  $P$  はベクトルの次数,  $\overline{d(\mathbf{x}, \mathbf{y})}$  はひずみの時間平均を表す.

全コードベクトル数  $N = 256$  なので,  $\log_2 M = 8 (M = 2^8 = 256)$  は全コードベクトルへの遷移が可能な通常のメモリレス VQ を表す. 図 3.9 からわかるように,  $\log_2 M \geq 5$  ではエントロピー, ひずみともに大きな変化が見られないが,  $\log_2 M = 4$  ではエントロピーが減少し, ひずみのわずかな増加が見られる. さらに  $\log_2 M = 3$  では, エントロピーが大幅に減少し, ひずみも大幅に増加している. これは遷移の数を減らしすぎたために音声スペクトルの変化を表現しきれなくなったことが原因と考えられる. しかし,  $\log_2 M \geq 4$  ではひずみはほとんど増加していないことから, この範囲ではコードベクトル間の遷移を制約する影響が少なく, LS-CTVQ を適用可能であるといえる.

### 3.3.2 ランダム誤り下での基礎的検討

通信路として誤り率  $P_e$  でランダムビット誤りが生じる二元対称通信路を仮定して, 音声のケプストラムの符号化/伝送/復号のシミュレーションを行なった結果について述べる. このシミュレーションのモデルを図 3.10) に示す. ここでは, 通信路として二元通信路を仮定しているた

め、トレリス上で誤り訂正を行なう際の符号系列間の距離尺度としてハミング距離を用いた。

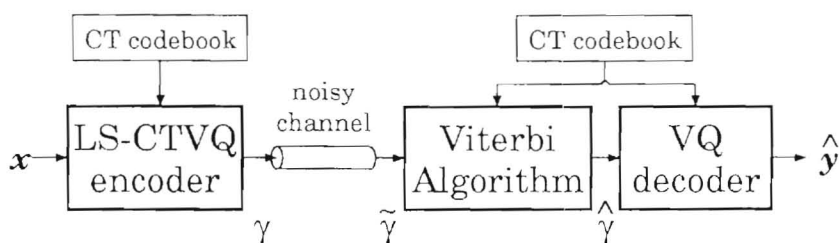


図 3.10 LS-CTVQ を用いた符号化システムのシミュレーションモデル

図 3.11 と図 3.12 は、通信路のビット誤り率  $P_e$  が  $10^{-3}$  から  $10^{-1}$  の場合の復号後のケプストラムひずみを表している。図 3.11 は誤り訂正時の判定遅延  $\ell_d$  が 1 フレームの場合、図 3.12 は 5 フレームの場合である。比較のために、通常のリニア VQ の結果も併せて示してある。どちらの図においても、 $M \geq 16$  の LS-CTVQ ではメモリレス VQ に見られるような大きな復号ひずみは見られず、 $M$  が 16 の場合には誤り率  $P_e = 10^{-1}$  において通常のリニア VQ に比べ 1.5dB 以上ひずみが小さくなっている。これより、LS-CTVQ ではビタビアルゴリズムによる誤り訂正によってビット誤りによるひずみが低減されていることがわかる。簡単な聴取実験を行なった結果では、ビット誤りがない場合とほとんど違いがわからなかった。誤り率の低い領域 ( $P_e \leq 10^{-2}$ ) では、遷移数の不足による量子化ひずみのために  $M = 8$  の LS-CTVQ のひずみがもっとも大きくなっているが、逆に誤り率が高い領域 ( $P_e > 10^{-2}$ ) では最小になっていることがわかる。これらの結果より、LS-CTVQ を用いて符号化を行うことで通信路にて重畳するビット誤りの訂正が可能であり、その誤り訂正能力は遷移数が少ないほど強力になることがわかる。しかし、遷移数の不足は量子化ひずみの増大を生じるため、LS-CTVQ の量子化能力と誤り訂正能力は遷移数を通じてトレードオフの関係にあることがわかる。

図 3.13 は、通信路の誤り率  $P_e$  が  $10^{-3}$  と  $10^{-1}$  の場合について、遷移数による復号ケプストラムひずみの変化を調べたものである。通信路の誤り

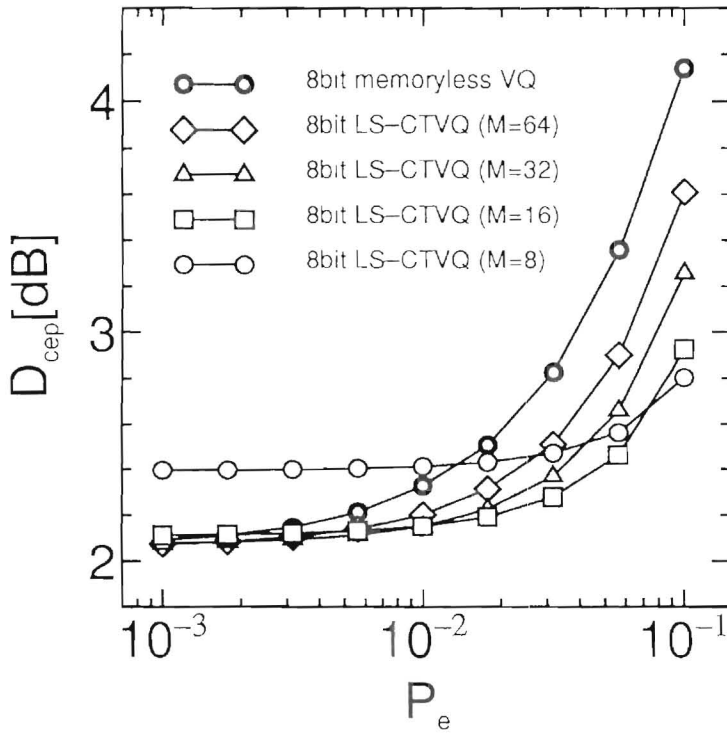


図 3.11 ビット誤り率と LS-CTVQ の復号ケプストラムひずみ ( $\ell_d = 1$ )

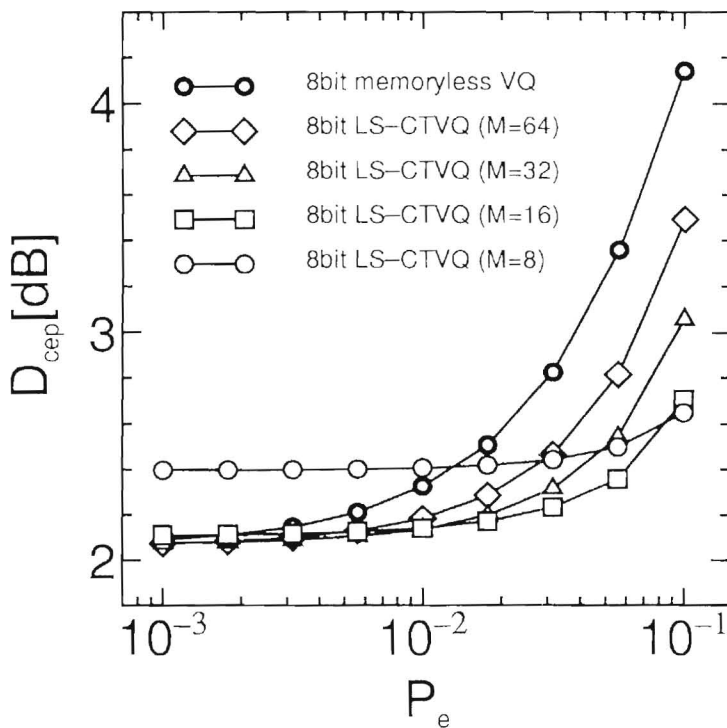


図 3.12 ビット誤り率と LS-CTVQ の復号ケプストラムひずみ ( $\ell_d = 5$ )

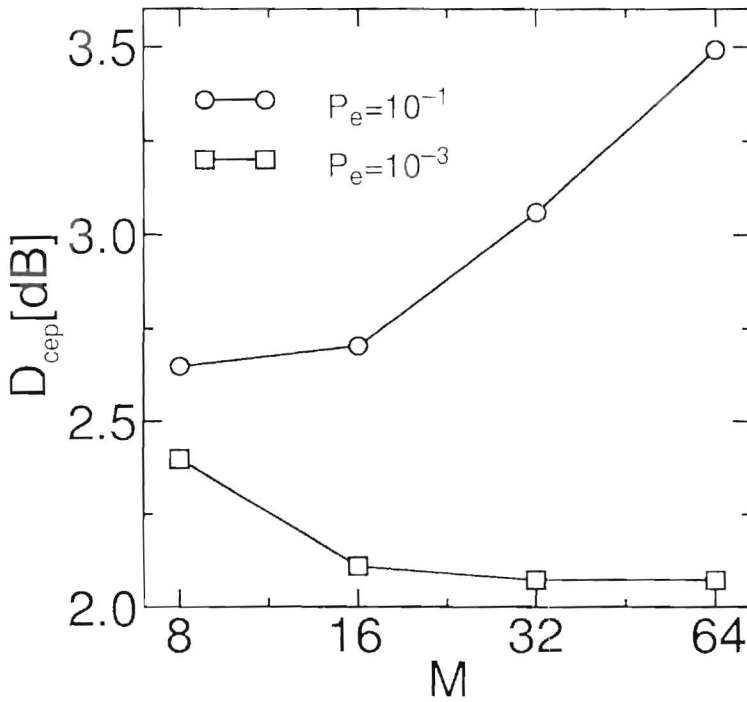


図 3.13 遷移数と復号ケプストラムひずみ  
( $l_d = 5$ )

率が  $10^{-3}$  と低い場合には、遷移数  $M$  が 16 以上ではひずみはほとんど変化しない。これに対して誤り率が  $10^{-1}$  と高い場合には、遷移数  $M$  が 16 を越えると急激にひずみが増加している。これより、各状態からの遷移数  $M$  には誤り率に対して最適値が存在することがわかる。このため LS-CTVQ を符号化に用いる場合には、量子化の対象とするパラメータ、分析周期、コードブックに含まれるコードベクトル数などに応じて最適な遷移数を求める必要があるといえる。ここでは、上記の結果より遷移数  $M$  は 16 が最適であるといえる。

次に判定遅延  $l_d$  に関する検討を行なう。図 3.11 と図 3.12 を比べると、判定遅延の大きい図 3.12 のほうが復号ケプストラムひずみが小さい。これは、判定遅延が大きいほうが誤り訂正能力が高いことを意味する。そこで遷移数  $M = 16$  の場合について、復号時の判定遅延による誤り訂正



能力の相違を調べた。その結果を図 3.14 に示す。誤り率が低い場合に

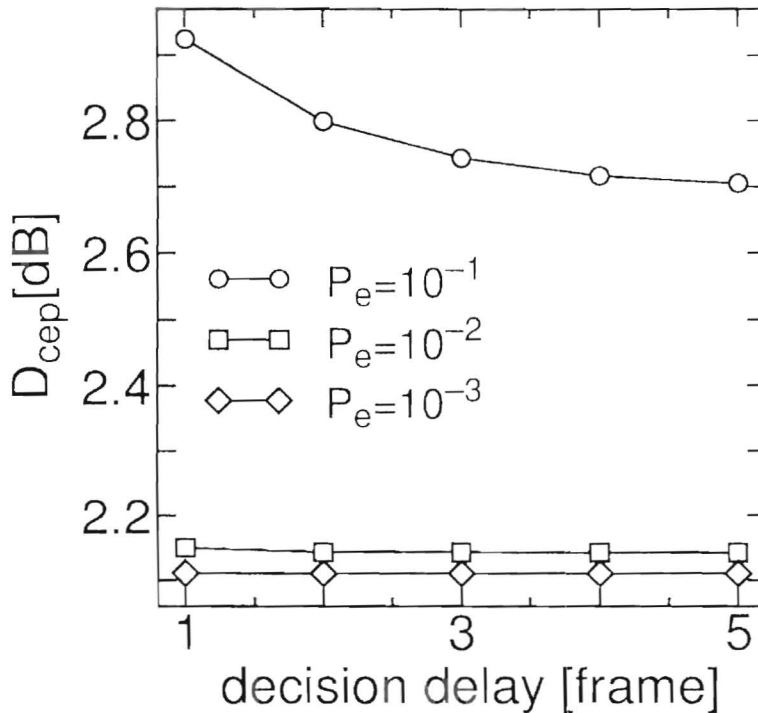


図 3.14 判定遅延と復号ケプストラムひずみ  
( $M=16$ )

は判定遅延を長くしてもひずみの改善は見られないが、誤り率が  $10^{-1}$  と高い場合には判定遅延を長くすることでひずみが改善している。しかしこの場合でも、判定遅延を3フレーム以上に増やしても大きなひずみの改善は見られず、さらに5フレーム以上では全く改善は見られなかった。この結果から、判定遅延は3~5フレーム程度で充分であり、LS-CTVQ 符号の誤り訂正には大きな判定遅延を必要としないことがわかる。

### 3.3.3 フェージング環境下における LS-CTVQ 符号の軟判定復号の効果

3.2.3で述べたように、LS-CTVQ の復号にはビタビアルゴリズムを用いるため、様々な距離尺度を用いることが可能である。通信路として二元対称通信路を仮定している場合には、尺度としてハミング距離を用いる

ことが適切だが、加法性白色ガウス通信路やフェージングのある通信路を仮定する場合には距離尺度としてユークリッド距離を用いた軟判定をすることで誤り耐性が改善することが知られている。そこで本節では通信路にレイリーフェージングが存在するものと仮定して、LS-CTVQの復号に軟判定を導入する効果について検討を行なう。

図3.15は、最大ドップラー周波数が40Hzのレイリーフェージングがある通信路を仮定し、ハミング距離を尺度とする硬判定復号とユークリッド距離を尺度とする軟判定復号の比較を行なったものである。比較のため

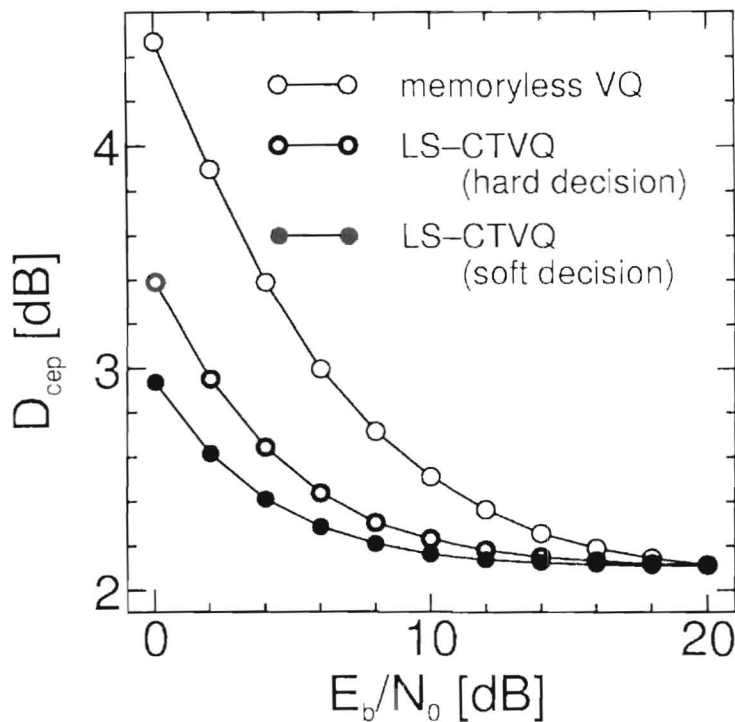


図 3.15 硬判定復号と軟判定復号の比較  
( $N = 256, M = 16, \ell_d = 5, f_d = 40\text{Hz}$ )

めにメモリレス VQ をフェージング通信路に適用した場合の結果も示してある。LS-CTVQ では、硬判定復号、軟判定復号のどちらの場合にもメモリレス VQ に比べてひずみが大幅に低減されている。このことから、フェージングのある劣悪な環境下においても、LS-CTVQ の誤り訂正が

表 3.2 パンクチャド符号の原符号

拘束長	生成多項式の係数	
7	133	171

注) 生成多項式の係数は 8 進表記

表 3.3 (3,2) 符号構成用消去マップ

生成多項式	消去マップ	
133	1	1
171	1	0

有効であることがわかる。さらに、軟判定復号を行なった場合には硬判定復号に比べて約 2dB ひずみが低減されており、LS-CTVQ においても軟判定復号が有効であることがわかる。

### 3.3.4 従来のタンデム符号化との比較

現在、通信路におけるビット誤りに対しては情報源符号化方式に通信路符号化方式を縦列接続したタンデム符号化を行なって対処することが一般的である。このようなタンデム符号化と比べて LS-CTVQ が持つ利点を調べるために、LS-CTVQ とタンデム符号化方式による復号ケプストラムひずみの比較を行なう。ここで情報源符号化には、条件付ヒストグラム法で状態遷移関数を決定した状態数が 256 の状態ラベル FSVQ (LS-FSVQ)<sup>(9)</sup>、通信路符号化には、表 3.2 の (2,1) 畳み込み符号と、それを原符号として表 3.3 の消去マップにしたがって構成した符号化率が (3,2) パンクチャド畳み込み符号<sup>(63)</sup>を用い、畳み込み符号の復号にはビタビアルゴリズム<sup>(61)</sup>を用いた。図 3.16 は、通信路として二元対称通信路を仮定して、 $N = 64$ ,  $M = 16$ ,  $l_d = 5$  の LS-CTVQ を、LS-FSVQ と

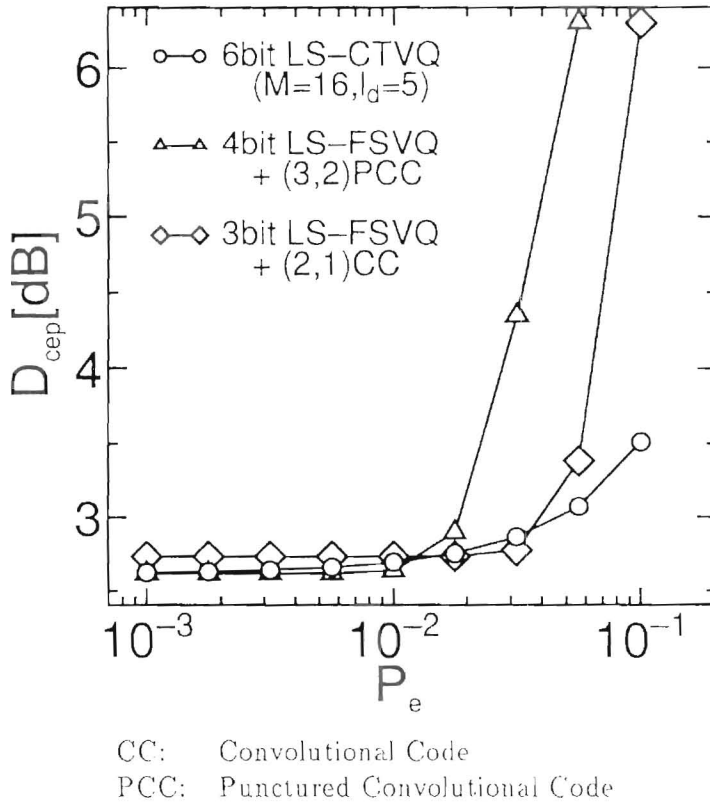


図 3.16 二元対称通信路下でのタンデム符号化方式との比較

畳み込み符号によるタンデム符号化方式と比較したものである。前節までは、LS-CTVQ のコードベクトル数が 256(8bit/vector) の場合について検討を行ってきたが、ここではコードベクトル数を 64(6bit/vector) とする。これによって、LS-CTVQ 符号の誤り訂正に用いるトレリスとパンクチャド符号のビタビ復号に用いるトレリスの状態数を揃え、復号器の複雑さにおいて同じ条件で比較を行なうことができる。また両者のビットレートを揃えるために、LS-FSVQ の各状態で出力可能なコードベクトル数を LS-CTVQ に比べてすくなく設定することで、情報源符号化に必要なビット数を減らし、これによって生じた余剰ビットをパンクチャド符号に割り当てた。すなわち、3bit の LS-FSVQ には符号化率 1/2 の畳み込み符号の原符号を、4bit の LS-FSVQ には符号化率 2/3 のパンクチャド符号を組み合わせて、それぞれ 6bit/vector として比較を



行なった。

4bit の LS-FSVQ を用いたタンデム方式を LS-CTVQ と比較した場合、通信路でビット誤りが重畳しなければ両者の復号ケプストラムひずみは全く同じになる。図から誤り率が低い領域 ( $P_e \leq 10^{-2}$ ) では両者のひずみはほぼ同じである。しかし、誤り率が  $10^{-2}$  を越えるとタンデム方式のひずみは急激に増大しているのに対して、LS-CTVQ ではわずかにひずみが増加するに留まっている。これから LS-CTVQ は (3.2) のパンクチャド畳み込み符号よりも強力な誤り訂正能力を有していることがわかる。また、3bit の LS-FSVQ と (2,1) の畳み込み符号を組み合わせたものを比較した場合、 $P_e \leq 10^{-2}$  では LS-CTVQ のほうがひずみが小さく、それ以上でも LS-CTVQ のひずみが徐々にしか増加していないのに対してタンデム方式のひずみが急激に増大していることを考慮すると、LS-CTVQ のほうが優れているといえる。このタンデム方式の急激なひずみの増大は FSVQ が自己回帰型のベクトル量子化法であることに起因する。すなわち、FSVQ では復号時に過去の復号ベクトルを直接参照するために、一度誤った判定を行なうと永久に誤りが波及するのである。このため、誤り訂正が可能な範囲ではひずみは増大しないが、誤り訂正能力を越えた場合には急激に誤りが増大するのである。この FSVQ における誤りの波及を阻止するためには、補助情報として符号器の状態の伝送や状態を初期状態に戻すなどの操作を一定時間毎に行なう必要があり、伝送情報量やひずみの増加、さらにシステムの複雑化を生じる。これに比べて LS-CTVQ の復号では過去の受信符号は参照するが、過去の復号結果は参照しない。このため誤り判定の波及がおこらないために、誤り率が大きい場合にも急激にひずみが増大することがなく、全体として良好な復号品質が得られるものと考えられる。

同様の検討を、最大ドップラー周波数が 40Hz のレイリーフェージングが存在する通信路を仮定して行なった。図 3.17 に比較結果を示す。図の横軸は通信路の平均  $E_b/N_0$  である。ここで、 $E_b$  は送信 1 ビットあたりのエネルギー、 $N_0$  は通信路のパワースペクトル密度を表す。また、復



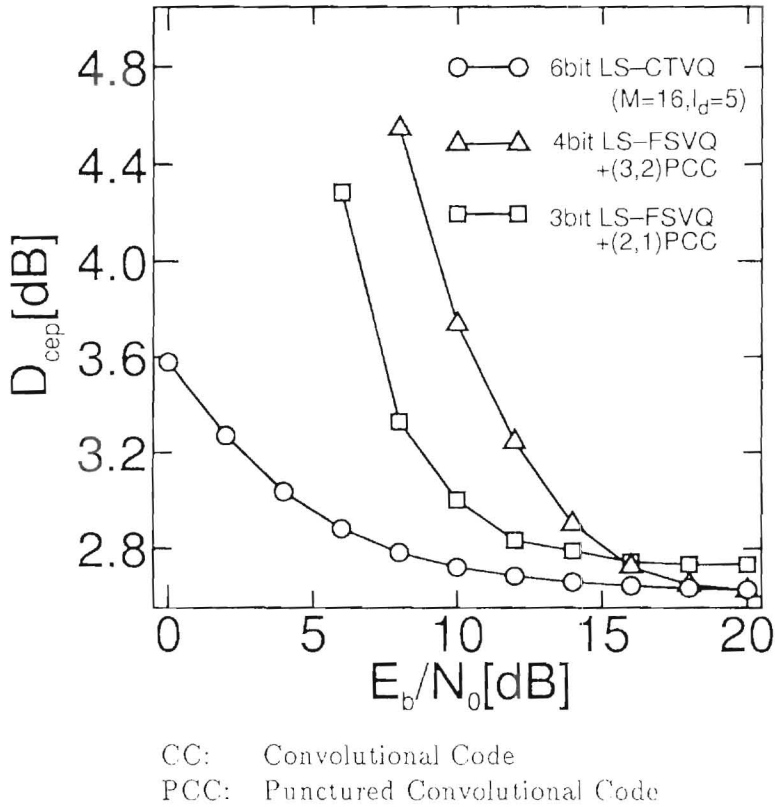


図 3.17 レイリーフェージング下でのタンデム符号化方式との比較

号には LS-CTVQ と畳み込み符号ともに軟判定ビタビアルゴリズムを用いている。この場合にも、LS-CTVQ はタンデム符号化よりも優れた特性を示し、復号後のケプストラムひずみが 3.0dB のところで比較すると、LS-CTVQ では、(3,2) のパンクチャド畳み込み符号を用いる場合に比べて約 9dB、(2,1) の畳み込み符号を用いる場合と比べても約 6dB 劣悪な通信環境において、同程度の品質の通信を行なうことが可能であり、フェージングのある劣悪な環境下においてもタンデム符号化方式よりも優れた誤り耐性を示すことがわかる。

### 3.4 本章のまとめ

本章では、トレリスの状態にラベル付けを行なう遷移を制約をしたベクトル量子化 (LS-CTVQ) を提案し、性能の検討を計算機シミュレーション

ンによって行なった.

音声の線形予測ケプストラムを符号化する計算機シミュレーションによって性能の検討を行ない, 提案する手法ではビット誤りによる復号ケプストラムひずみを大幅に低減できることを示した. また, 通常のベクトル量子化や有限状態ベクトル量子化を誤り訂正符号と組み合わせたタンデム符号化方式と比較し, 提案する手法がタンデム符号化方式より優れた誤り耐性を有することを示した. これらの結果より, 記憶のある情報源の出力をベクトル量子化する場合には, LS-CTVQ を用いることで, 高能率で, しかも通信路誤りに強い符号化が可能であると考えられる.

## 第4章 遷移を制約したベクトル量子化の改良法

### 4.1 まえがき

第3章において、コードベクトル間の遷移を制約することで通信路で重畳する誤りを訂正可能にする LS-CTVQ について述べた。LS-CTVQ ではコードベクトルを状態に対応させたトレリス上でビタビアルゴリズムを適用して誤り訂正を行なうが、ベクトル量子化の量子化精度をあげるためにコードベクトルの数を増やした場合、コードベクトルと状態が対応していることによって復号に用いるトレリスの状態数も増加する。この状態数の増加は、ビタビアルゴリズムの実行に必要な演算量と記憶量が量子化ビット数にたいして指数的に増加してしまうという問題を生じる。ここで、LS-CTVQ の誤り訂正能力はコードベクトル間の遷移が制約されることによって生じるものであることを考えると、コードベクトルとトレリスの状態は必ずしも対応している必要はないと考えられる。

本章では、LS-CTVQ の改良法としてトレリスの状態間の遷移に符号ベクトルに対応させる遷移ラベル型 CTVQ (Labeled-Transition CTVQ : LT-CTVQ) について述べる。コードベクトルをトレリスの状態間の遷移に割り当てた場合にも、各状態から出ていく遷移の数 (各状態で出力可能なコードベクトルの数) を全コードベクトルの数よりも少なく抑えてコードベクトル間の遷移に制約を与えることで、復号器においてビタビアルゴリズムを用いた誤り訂正が可能になると考えられる。本手法では、ベクトル量子化で用いる全コードベクトル数によらずトレリスの状態数を決めることが可能であり、状態数を全コードベクトルの数より少なくすることで LS-CTVQ にくらべてビタビアルゴリズム実行時の記憶量と演算量を少なく抑えることが可能となる。

本章では、まず LT-CTVQ の符号化／誤り訂正／コードブック設計アルゴリズムについて述べる。次に、第3章と同様に、音声の線形予測ケプストラムを誤りのある通信路を通じて符号化／伝送するシミュレーションにより LT-CTVQ の性能評価を行なう。

## 4.2 トレリスの遷移にラベル付けを行なう遷移を制約したベクトル量子化 (LT-CTVQ)

### 4.2.1 符号化アルゴリズム

いま符号化の対象を  $p$  次元ユークリッド空間  $\mathbf{R}^p$  から出力されるベクトルと仮定する. 符号器は  $K$  個の状態  $\mathbf{S} = \{s_k \in \mathbf{R}^p; k = 0, 1, \dots, K-1\}$ ,  $N$  個のコードベクトル  $\mathbf{C} = \{c_i \in \mathbf{R}^p; i = 0, 1, \dots, N-1\}$  および各コードベクトルに対応する符号  $\mathbf{U} = \{u_j \in 2^r; j = 0, 1, \dots, N-1\}$  を持ち,  $p$  次元実数ベクトルを符号長  $r$  bit の二進符号に符号化するものとする. ここで  $N = 2^r$  である. また各状態から出る遷移の数は一定とし, その数を記号  $M$  で表し遷移数と呼ぶことにする. コードベクトル (符号) 間の遷移に制約を与えるために, ここでは遷移数  $M$  は全コードベクトル数  $N$  より小さいものとする ( $M < N$ ). LT-CTVQ ではこれらの遷移にコードベクトルを割り当てるため, 符号化の際には各状態から出ている遷移に割り当てられているコードベクトルを探索して量子化ベクトルを決定することになる. これは状態毎に  $M$  個のコードベクトルを含むサブコードブックを持つことに相当し, これを状態コードブックと呼び以下のように表す.

$$\mathbf{C}' = \{c'_j \in \mathbf{C}; j = 0, 1, \dots, M-1\} (i = 0, 1, \dots, K-1) \quad (4.1)$$

また状態コードブックの各コードベクトルに対応する符号を以下のように表すこととする.

$$\mathbf{U}' = \{u'_j \in \mathbf{U}; k = 0, 1, \dots, M-1\} (i = 0, 1, \dots, K-1) \quad (4.2)$$

ここで, 各状態の状態コードブックには他の状態の状態コードブックと重複するコードベクトルも含まれることに注意する.

符号器の初期状態を  $\sigma_0 = s_0 \in \mathbf{S}$ , 時刻  $n$  における符号器の状態を  $\sigma_n = s_i \in \mathbf{S}$  と表したとき, 入力ベクトル系列  $\{\mathbf{x}_n; n = 0, 1, 2, \dots\}$  は次のように符号化される.

$$\nu_n = \alpha(\sigma_n, \mathbf{x}_n) \quad (4.3)$$

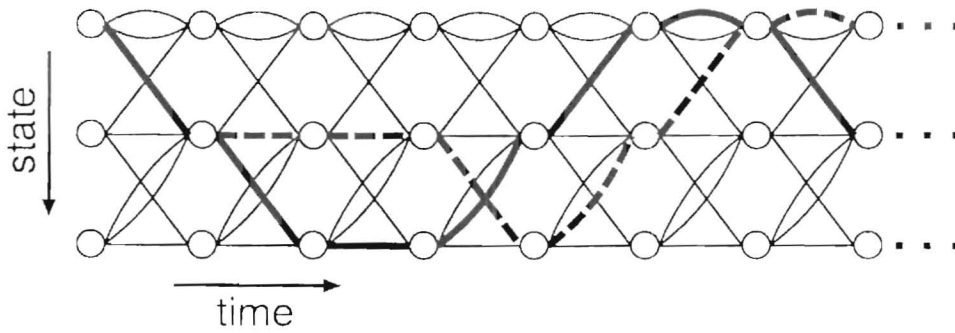


図 4.1 遷移ラベル型 CTVQ のトレリス表現

$$\mathbf{y}_n = \beta(\sigma_n, \nu_n) \tag{4.4}$$

$$\sigma_{n+1} = f(\sigma_n, \nu_n) \tag{4.5}$$

ここで、 $\nu_n \in U$  は  $\mathbf{x}_n$  の符号化出力、 $\mathbf{y}_n$  は  $\mathbf{x}_n$  の量子化ベクトル、 $\alpha / \beta / f$  はそれぞれ、符号器/復号器/状態遷移関数を表す。2つのベクトル  $\mathbf{x}$  と  $\mathbf{y}$  の距離を  $d(\mathbf{x}, \mathbf{y})$  で定義した場合、符号器は次のように表される。

$$\alpha(\sigma_n = s_i, \mathbf{x}_n) = \underset{u_j^i \in U^i}{\operatorname{argmin}} \{d(\mathbf{x}_n, \beta(\sigma_n, u_j^i))\} \tag{4.6}$$

ここで  $\beta(\sigma_n = s_i, \nu_n = u_j^i) = \mathbf{c}_j^i$  である。

LT-CTVQ のトレリス表現の一例を図 4.1 に示す。LT-CTVQ では、トレリスの状態数を全コードベクトル数よりも少なく設定することが可能であり、トレリスの状態間の遷移は必ずしも制約されない。さらに、2つの状態間に複数の遷移が存在することさえある。しかし、このような場合にも符号器が各状態で出力できるコードベクトルの数(遷移数) $M$ を全コードベクトル数よりも少なく抑えることで、コードベクトル間の遷移に制約を与え、トレリス上でビタビアルゴリズムを用いて誤り訂正を行なうことが可能となる。次節では、このトレリスを利用した LT-CTVQ 符号の誤り訂正法について述べる。



#### 4.2.2 遷移ラベル型 CTVQ 符号の誤り訂正

LT-CTVQ の符号器が出力する符号系列  $\{\nu_m; m = 0, 1, 2, \dots\}$  は、トレリス上で1本の連続したパスとして表すことができる。しかし通信路で誤りが重畳した受信符号系列  $\{\tilde{\nu}_m; m = 0, 1, 2, \dots\}$  は、トレリス上で1本の連続したパスにはならず、パスに不連続が生じる。そこで LS-CTVQ と同様に、受信符号系列に対してトレリス上でビタビアルゴリズムを適用して元の受信符号系列の推定を行うことで、受信符号系列に重畳した通信路誤りを低減することが可能となる。

時刻  $n$  に状態  $s_i$  に到達するトレリスのパスのうち、時刻  $n$  までの受信符号系列  $\tilde{U}_n = \{\tilde{\nu}_m; m = 0, 1, \dots, n\}$  に最も近い符号系列  $\hat{U}_{n,i}$  を出力するパスを  $Q_{n,i}$ 、 $\tilde{U}_n$  と  $\hat{U}_{n,i}$  の累積距離を  $\lambda_{n,i}$  と表し、次のように初期化する。

$$\lambda_{n,i} = \begin{cases} 0 & (s_i = \sigma_0) \\ \infty & (\text{otherwise}) \end{cases} \quad (4.7)$$

$$Q_{0,i} = \{s_i\} \quad (4.8)$$

また、状態  $s_i$  に遷移可能な状態の集合を  $A_i$ 、状態  $s_j \in A_i$  から状態  $s_i$  への遷移で出力される可能性のある符号の集合を  $P_{j,i}$  と表す (図 4.2)。すなわち、

$$A_i = \{\sigma; f(\sigma, \nu) = s_i\} \quad (4.9)$$

$$P_{j,i} = \{\nu; f(s_j, \nu) = s_i\} \quad (4.10)$$

通信路から符号を受信した復号器は、符号系列間の累積距離  $\lambda_{n,i}$  を最小化するパス  $Q_{n,i}$  をトレリスの各状態  $\{s_i; i = 0, 1, \dots, K-1\}$  についてビタビアルゴリズムを用いて探索する。

$$\hat{j} = \underset{j, s_j \in A_i}{\operatorname{argmin}} \left[ \lambda_{n-1,j} + \min_{u \in P_{j,i}} \{\Delta(\tilde{\nu}_n, u)\} \right] \quad (4.11)$$

$$\lambda_{n,i} = \lambda_{n-1,\hat{j}} + \min_{u \in P_{\hat{j},i}} \{\Delta(\tilde{\nu}_n, u)\} \quad (4.12)$$

$$Q_{n,i} = Q_{n-1,\hat{j}} \otimes s_i \quad (4.13)$$

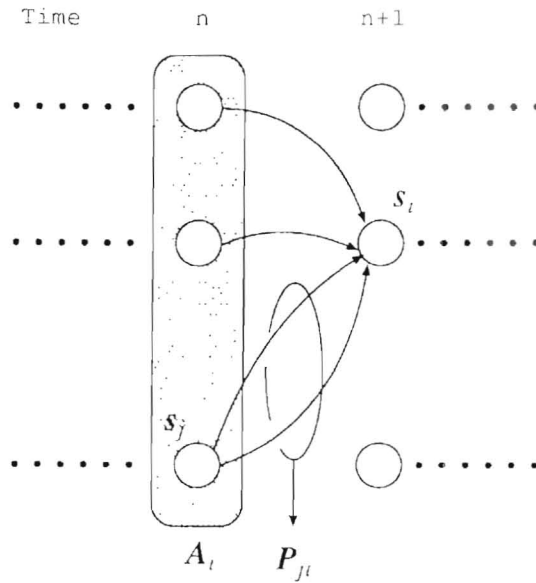


図 4.2 記号の説明

ここで  $\hat{j}$  は  $\lambda_{n,i}$  を最小にするパスの時刻  $n - 1$  における状態へのインデックス,  $\Delta(u, v)$  は符号  $u$  と  $v$  の間の距離を表し,  $\otimes$  はパスに状態を連結して新たなパスをつくる演算子である. こうして得られるパス系列  $Q_{n,i}$  に対応する符号系列  $\hat{U}_{n,i}$  は,

$$\hat{U}_{n,i} = \hat{U}_{n-1,\hat{j}} \otimes \underset{u \in P_{\hat{j},i}}{\operatorname{argmin}} \{ \Delta(\tilde{v}_n, u) \} \quad (4.14)$$

で与えられる.  $\otimes$  は  $\otimes$  同様, 符号系列に符号を連結して新たな符号系列をつくる演算子である.

上記の操作を全ての符号を受信するまで行い, 距離が最小のパスに対応する符号系列を復号系列とすれば良いが, これでは復号の遅延が膨大なものになってしまう. そこでここでは以下のような方法で符号の推定を行う. まずあらかじめ復号遅延  $l_d$  を決めておき, その数だけの符号を受信するまでは符号の推定は行わない. 次に  $n > l_d$  の時刻において符号  $\tilde{v}_n$  を受信した時には, その時点までの符号系列間の累積距離  $\lambda_{n,i}$  が最小になるパス  $Q_{n,i}$  に対応する符号系列  $\hat{U}_{n,i}$  を選択し, この符号系列における時刻  $n - l_d$  の符号  $\hat{v}_{n-l_d}$  に対応するコードベクトルを復号ベクトルと

することで、復号ベクトル系列  $\hat{X} = \{\hat{x}_m; m = 0, 1, 2, \dots\}$  が得られる。

LS-CTVQ の場合と同様、誤り訂正にビタビアルゴリズムを用いているため、符号間の距離尺度  $\Delta(u, v)$  に符号と受信サンプル間のユークリッド距離を用いる軟判定復号が容易に実行できる。軟判定復号の導入によって復号器のシステム構成は複雑になるが、符号間の距離尺度として符号間のハミング距離を用いる硬判定復号よりも誤り訂正能力を改善することができると考えられる。この時、硬判定と軟判定では符号間の距離尺度が異なる以外、誤り訂正のアルゴリズムには大きな違いはない。

### 4.2.3 コードブック設計アルゴリズム

LT-CTVQ のコードブック設計アルゴリズムは、(1) 初期コードブックの作成、(2) 各状態の状態コードブックの決定、という2つの段階に大きく分けることができる。

まず最初に、初期コードブックとして  $K$  個の状態ラベル  $S = \{s_k \in \mathbf{R}^p; k = 0, 1, \dots, K-1\}$  と  $N$  個のコードベクトル  $C = \{c_i \in \mathbf{R}^p; k = 0, 1, \dots, N-1\}$  を LBG アルゴリズム<sup>(6)</sup>などのコードブック設計アルゴリズムを用いて決定する。ここで得られる  $K$  個の状態ラベルは状態遷移関数の決定に用いるもので、符号化の際には不要となる。また、各コードベクトル  $c_i$  にはそれぞれ符号  $u_i$  が割り当てられる。2.3.4で述べたように、各符号への通信路アルファベットの割り当てを工夫することで符号誤りが重畳した場合の劣化を抑える手法がいくつか提案されているが、ここではコードベクトル(符号)間の遷移を制約することで符号誤りを低減することが目的であるため、単純に  $u_i = i$  とした。

次に状態コードブックを決定する必要がある。ここでは、LBG アルゴリズムなどで作成された  $N$  個のコードベクトルの中から、状態毎に  $M$  個のコードベクトルを選択し状態コードブックを得る。まず、全ての状態間の遷移が可能であると仮定し、符号器の状態遷移規則を「符号  $u_i \in U$  を出力した場合には、コードベクトル  $c_i \in C$  に最も近い状態ラベルを持

つ状態に遷移する。」と決める。すなわち、状態遷移関数  $f(\sigma, u_i)$  は、

$$f(\sigma_n = \mathbf{s}_k, \nu_n = u_i) = \underset{\mathbf{s}_k \in \mathbf{S}}{\operatorname{argmin}} \{d(\mathbf{c}_i, \mathbf{s}_k)\} \quad (4.15)$$

となる。このように状態遷移関数を決めた場合、状態コードブック内の各コードベクトルに対する状態遷移先は一意に決定され、各コードベクトルは状態間の遷移と対応することになる。

上記の状態遷移規則の下で、1) 条件付きヒストグラム法、2) 最近傍法の2通りの方法によって状態毎に  $M$  個のコードベクトルを選択した。条件付きヒストグラム法は、学習用データを用いて各状態において各コードベクトルが出力される条件付きヒストグラムを求め、度数が高いコードベクトルから  $M$  個のコードベクトルを状態コードブックの要素とする手法である。これに対して最近傍法は、各状態の状態ラベルと各コードベクトルの距離を調べ、距離の近いコードベクトルから順番に状態コードブックの要素とする。条件付きヒストグラム法は、ベクトル量子化による量子化ひずみを抑えることが可能だが、学習用データに依存しやすく状態コードブックの決定に必要な計算量が多い、という欠点を有する。これに対して最近傍法は、状態ラベルと各コードベクトル間の距離を計算するだけでよいので簡易だが、量子化ひずみが大きくなる可能性がある。以後、どちらの学習法を用いたかを区別する必要がある場合には、条件つきヒストグラム法で学習した場合には CH-LT-CTVQ、最近傍法を用いた場合は NN-LT-CTVQ と呼ぶことにする。

本節で述べた手法により、それぞれ  $M$  個のコードベクトルを含む  $K$  個の状態コードブックと状態遷移関数が決定される。ここで、遷移数  $M$  をコードベクトルの総数  $N$  より少なく抑えることでコードベクトル間の遷移に制約が生じ、ビタビアルゴリズムによる誤り訂正が可能となる。さらに、状態数  $K$  を  $N$  よりも少なく設定することで LS-CTVQ よりもビタビアルゴリズムの実行に必要な計算量と記憶量を低減することが可能である。次節では、音声の線形予測ケプストラムを符号化/伝送する計算機シミュレーションによって LT-CTVQ の性能評価を行なう。

表 4.1 シミュレーション諸元

音声資料	男性2名, 女性4名の発声した50短文
標本化	8kHz, 8bit
分析次数	P=12
分析フレーム長	30ms
分析周期	5ms
エンファシス	なし

### 4.3 計算機シミュレーション

本節では、音声の線形予測ケプストラムを伝送する計算機シミュレーションによってLT-CTVQの能力を検証する。シミュレーションの主な諸元を表4.1に示す。また、ベクトル量子化時のベクトル間の距離尺度には(3.3)式の自乗距離を用いた。

#### 4.3.1 ランダム誤り下での基礎検討

まず最初に通信路としてランダムビット誤りの発生する二元対称通信路を仮定し、符号器の状態数  $K$  の違いによる復号後のケプストラムひずみの変化を調べた。図4.3は条件付きヒストグラム法で作成された  $N = 64$ ,  $M = 16$  のコードブックを用いたCH-LT-CTVQの、通信路のビット誤り率対復号後のケプストラムひずみ特性を表している。ビタビアルゴリズムの判定遅延  $l_d$  は5である。比較のために、メモリレスVQの結果も併せて示してある。すべての  $K$  において、ビット誤り率が低い場合にはメモリレスVQに比べてわずかにひずみが大きくなっている。しかし、 $K = 16$  の場合で誤り率が  $1.0 \times 10^{-2.0}$  以上、 $K = 32$  と  $K = 64$  の場合では誤り率が  $1.0 \times 10^{-2.25}$  以上では逆転し、LT-CTVQのほうがメモリレスVQに比べてケプストラムひずみが大幅に小さくなっている。また、状態数が違っても特性には大きな差はみられない。これより、遷移に



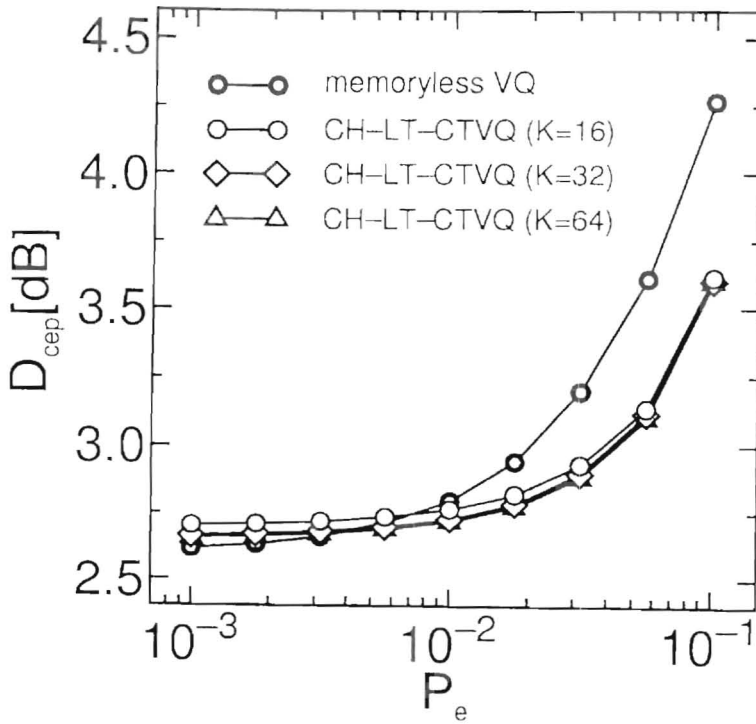


図 4.3 ビット誤り率と CH-LT-CTVQ の復号ケプストラムひずみ  
( $N = 64, M = 16, \ell_d = 5$ )

コードベクトルを割り当てた場合にも、コードベクトル間の遷移が制約されていればビタビアルゴリズムによる誤り訂正が可能であり、復号時に用いるトレリスの状態数の削減が可能となることがわかる。

図 4.4 は最近傍法で作成されたコードブックを用いた NN-LT-CTVQ の、ビット誤り率対ケプストラムひずみ特性を示したものである。コードベクトル数などは図 4.3 と同じである。NN-LT-CTVQ では、CH-LT-CTVQ に比べて状態数が少ない場合の特性が悪い。これは状態数が少ない場合に最近傍法を適用すると、どの状態コードブックにも属さないコードベクトルができてしまい、実際の符号化に用いられるコードベクトルが少なくなってしまうことが原因と考えられる。このため NN-LT-CTVQ は、状態数  $M$  を全コードベクトル数  $N$  に比べて大幅に低減する必要がある場合にはあまり適していないと考えられる。

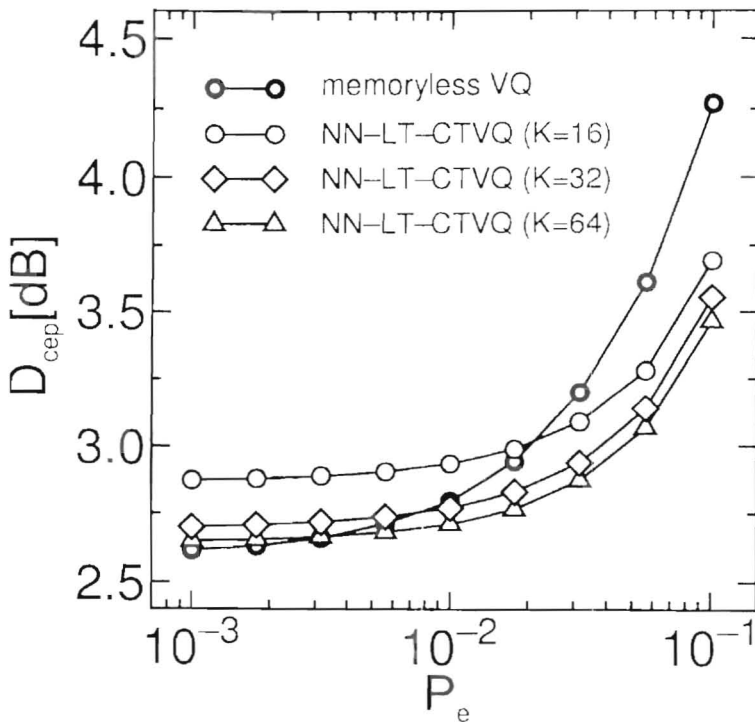


図 4.4 ビット誤り率と NN-LT-CTVQ の復号ケプストラムひずみ  
( $N = 64, M = 16, \ell_d = 5$ )

遷移数  $M$  が少ない場合には、コードベクトル系列に与えられる制約が強くなるために誤り訂正能力は強力になるが、スペクトルの変化に対応しきれずに誤り率が低い場合のひずみが悪くなるものと考えられる。この点を検証するために、遷移数によるひずみの変化を調べた。図 4.5 と図 4.6 は、 $N = 64$  で  $K = 16$  と  $K = 64$  の CH-LT-CTVQ の遷移数とケプストラムひずみの関係を表している。図 4.5 の遷移数が 8 の場合を除けば、ビット誤り率が高い場合には遷移数を減少につれて復号後のケプストラムひずみは減少し、誤り率が低い場合には遷移数の減少につれて復号後のひずみは増加する傾向が見られる。図 4.5 の遷移数が 8 の場合は、状態数、遷移数共に少ないためにどの状態コードブックにも含まれないコードベクトルができてしまい、極端に量子化ひずみが大きくなってしまふことが原因と考えられる。

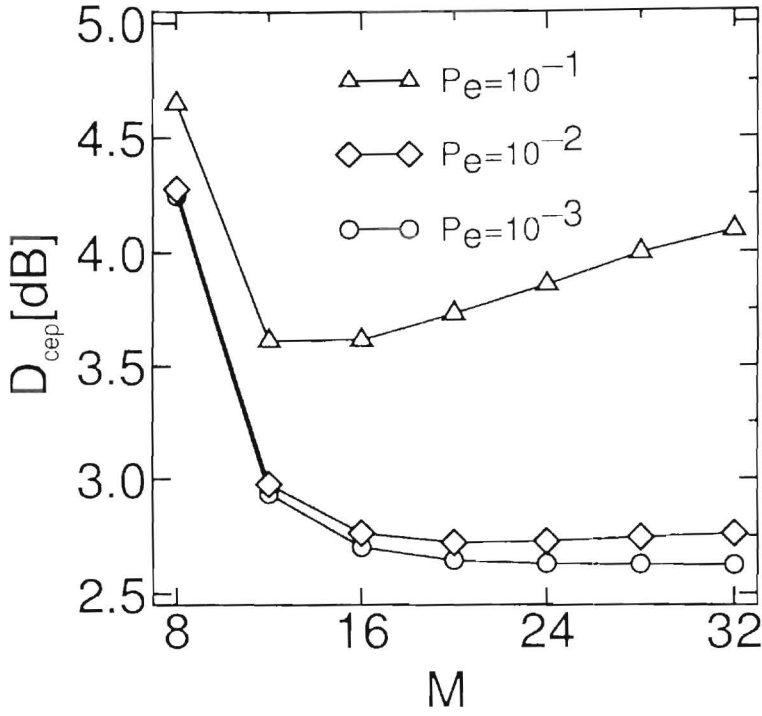


図 4.5 遷移数と CH-LT-CTVQ の復号ケプストラムひずみ ( $N = 64, K = 16, \ell_d = 5$ )

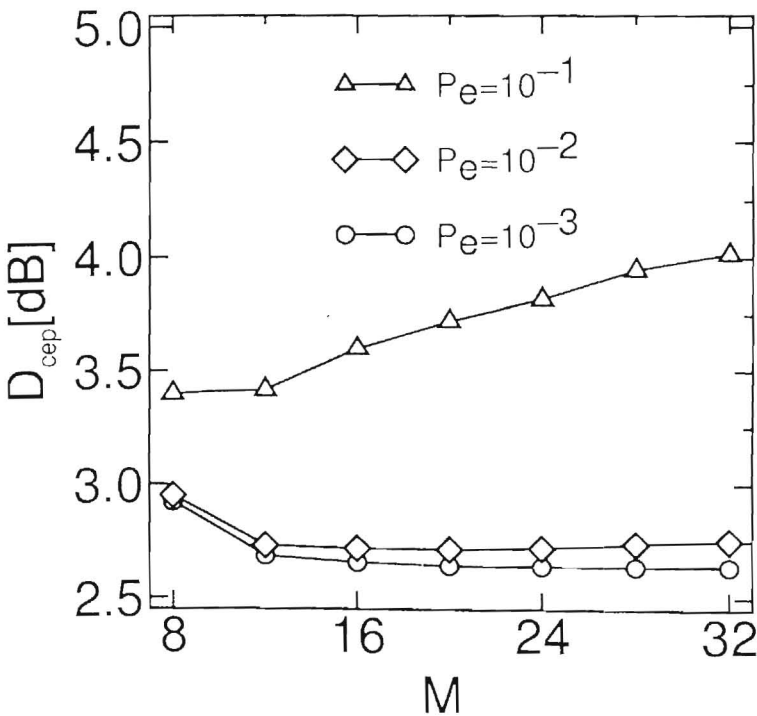


図 4.6 遷移数と CH-LT-CTVQ の復号ケプストラムひずみ

次に最近傍法によりコードブックを作成した場合の遷移数対ケプストラム特性を図4.7と図4.8に示す。NN-LT-CTVQの場合もCH-LT-CTVQの場合と同様の傾向を示しているが、図4.7では遷移数が12以下でひずみが大幅に劣化している。また、図4.7でも遷移数が8で遷移数の不足に起因すると考えられるひずみの増大が見られる。これらは、CH-LT-CTVQの場合と同じくどの状態コードブックにも属さないコードベクトルができてしまうことが原因と考えられ、NN-LT-CTVQを用いる場合にはCH-LT-CTVQに比べて遷移数を多くする必要がある。ビタビアルゴリズム実行時の計算量と記憶量を低減するためにトレリスの状態数を削減することが、LT-CTVQを用いる目的であることを考慮すると、条件つきヒストグラム法が優れているといえる。

#### 4.3.2 フェージング環境下におけるLT-CTVQ符号の軟判定復号の効果

LT-CTVQにおいてもLS-CTVQと同様に、ビタビアルゴリズムの距離尺度にユークリッド距離を用いた軟判定復号が可能である。そこで、通信路に最大ドップラー周波数が40Hzのレイリーフェージングが存在すると仮定し、LT-CTVQにおける軟判定復号の効果調べた。前節の結果をふまえ、学習には条件つきヒストグラム法を用いた。

図4.9は、コードベクトル数 $N = 64$ 、状態数 $K = 16$ 、遷移数 $M = 16$ のコードブックを用いたCH-LT-CTVQにおける、平均 $E_b/N_0$ 対復号後のケプストラムひずみ特性を表している。比較のためにメモリレスVQを用いた場合の特性も示してある。メモリレスVQと比較すると、図4.9ではケプストラムひずみが3.0dBの時に、硬判定ビタビアルゴリズムを用いたCH-LT-CTVQで約2.2dB、軟判定ビタビアルゴリズムを用いたCH-LT-CTVQでは約4.1dBの改善が得られている。この結果から、LT-CTVQの場合にも軟判定の導入によって、復号特性が改善されることが分かる。そこで、以降では軟判定ビタビアルゴリズムを用いて検討を行うこととする。

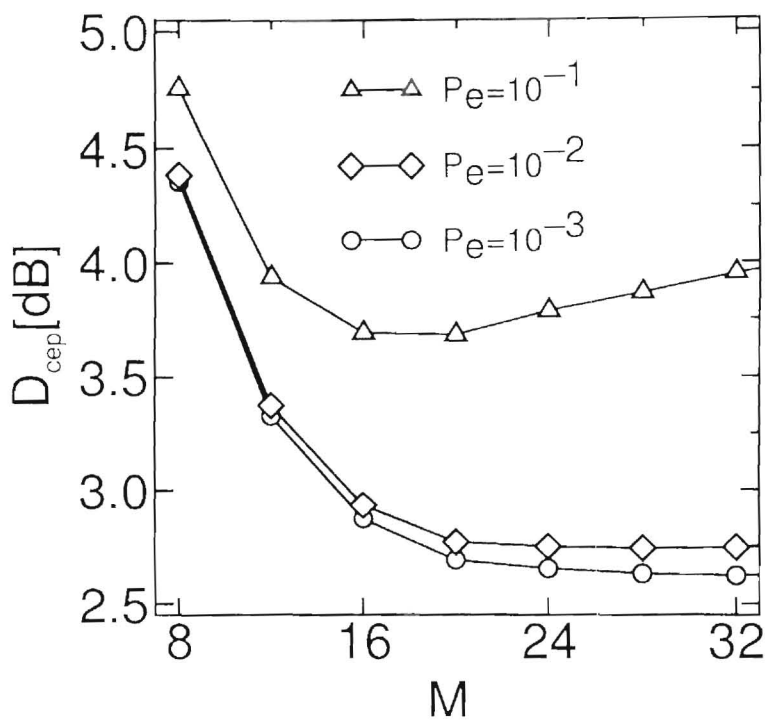


図 4.7 NN-LT-CTVQ における遷移数とケプストラムひずみの関係  
( $N = 64, K = 16, \ell_d = 5$ )

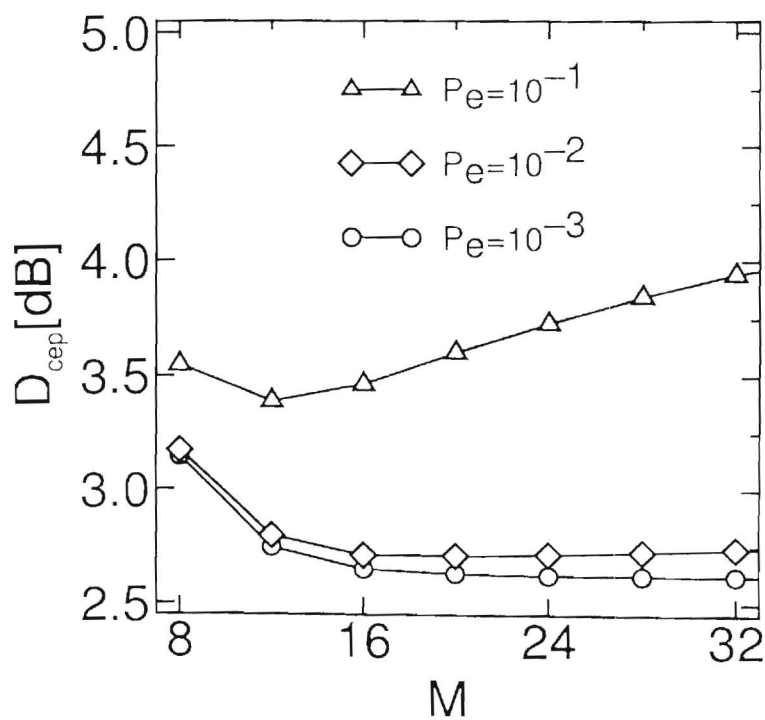


図 4.8 NN-LT-CTVQ における遷移数とケプストラムひずみの関係



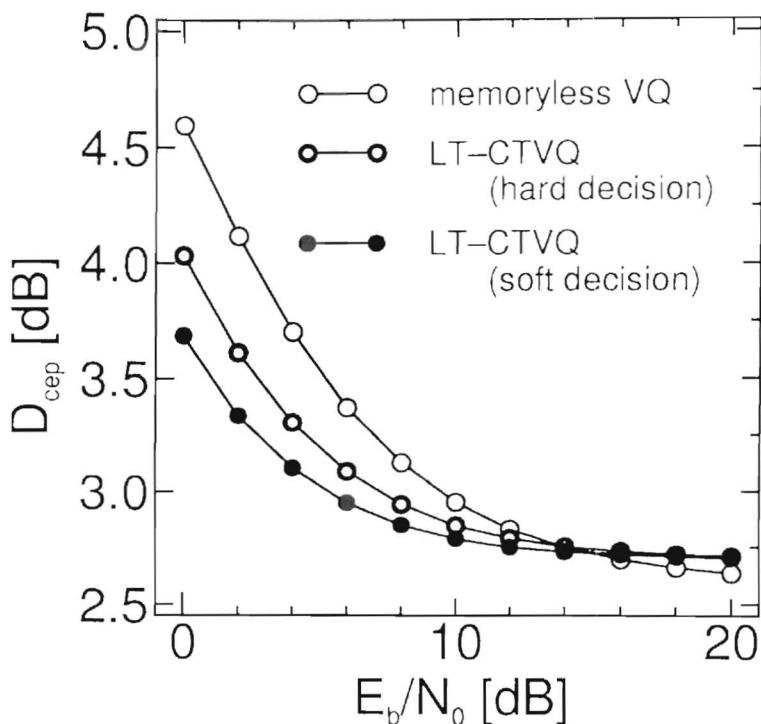


図 4.9 LT-CTVQ における硬判定と軟判定の比較

#### 4.3.3 従来のタンデム符号化方式との比較

3.3.4と同様に、LT-CTVQによる符号化システムとタンデム符号化を用いる符号化システムの比較を行なった。ここでは、情報源符号化に遷移ラベル型の有限状態ベクトル量子化(LT-FSVQ)、通信路符号化に表??の畳み込み符号から表3.3の消去マップによって構成されるパンクチャド畳み込み符号化を用いるタンデム符号化システムのみを構成した。これは、LT-FSVQではLT-CTVQと同様に、遷移数が少ない場合には量子化ひずみが増大してしまい、比較対象となりえないためである。

図4.10は、コードベクトル数  $N = 64$ 、状態数  $K = 16$ 、遷移数  $fM = 16$  のLT-CTVQを比較したものである。ここではLT-CTVQ、パンクチャド畳み込み符号ともに復号には軟判定ビタビアルゴリズムを用いた。LT-FSVQと(3,2)パンクチャド符号を用いたタンデム符号化方式では、 $E_b/N_0$ が15dB以下になると急激にひずみが増加している。これは、タ

表 4.2 パンクチャド符号の原符号

拘束長	生成多項式の係数	
5	23	35

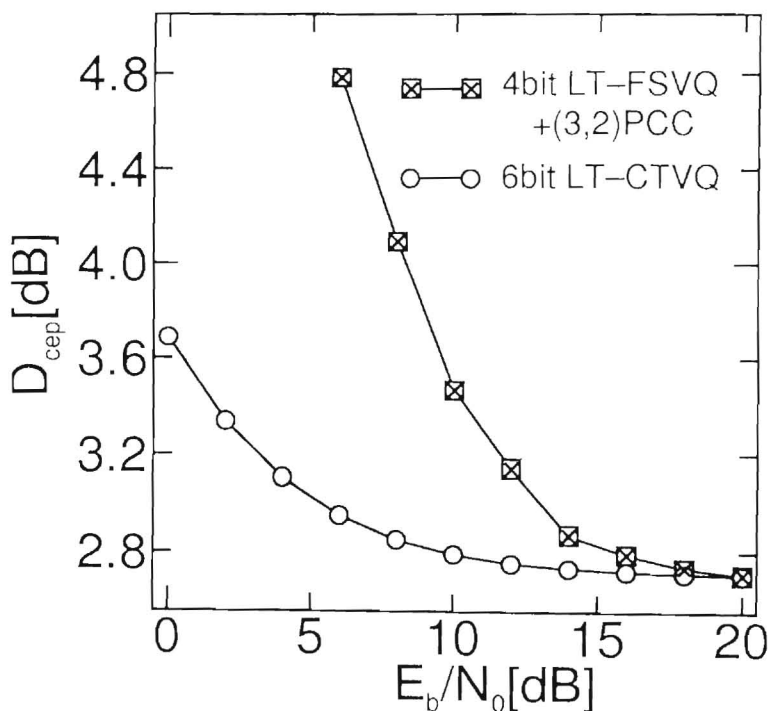


図 4.10 タンデム符号化方式との比較  
 $(N = 64, K = 64, \ell_d = 5)$

ンデム符号化方式で用いている畳み込み符号の拘束長が5と短いために誤り訂正能力が低いことに加えて、FSVQが自己回帰型の復号を行なうことが原因である。これに対して、LT-CTVQでは $E_b/N_0$ が5dB程度までほとんどひずみが増加していない。LT-CTVQでもLS-CTVQと同様に過去の受信符号は参照するが、過去の復号結果は直接は用いていない。このため、誤りの波及がおこらないことが急激な歪みの増大を防いでいるものと考えられる。これからLT-CTVQによっても、従来のタン

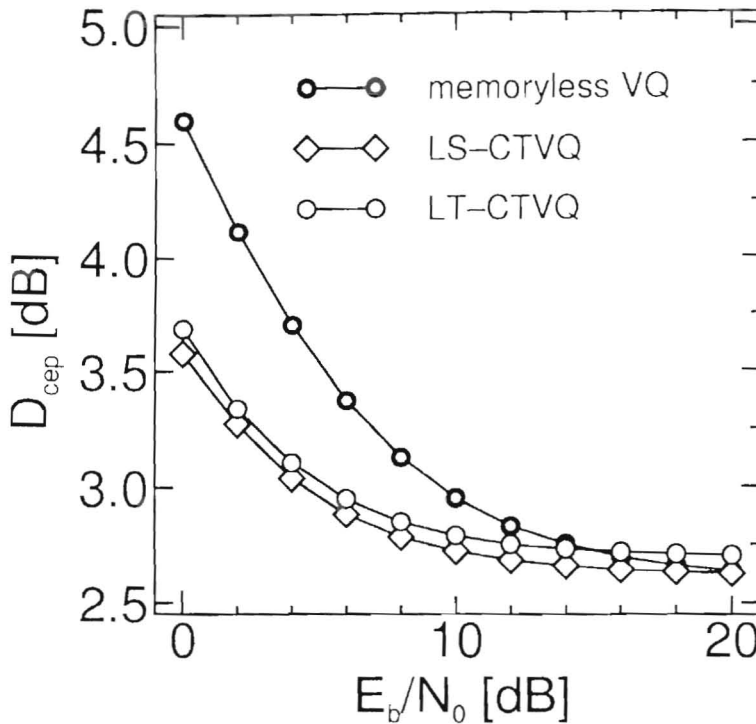


図 4.11 LS-CTVQ と LT-CTVQ の比較

デム符号化よりも誤りに強い符号化を行なうことが可能なことがわかる。

#### 4.3.4 LS-CTVQ と LT-CTVQ の比較

次に LT-CTVQ と LS-CTVQ の比較を行った。図 4.11 は、コードベクトル数  $N = 64$ 、遷移数  $M = 16$  の LS-CTVQ と、コードベクトル数  $N = 64$ 、状態数  $K = 16$ 、遷移数  $M = 16$  の LT-CTVQ の  $E_b/N_0$  対復号後のケプストラムひずみ特性を比較したものである。どちらの場合も復号遅延  $l_d$  は 5 フレームである。ここでも通信路には最大ドップラー周波数が 40Hz のレイリーフェージングが存在するものと仮定した。LT-CTVQ では、LS-CTVQ に比べて約 1/4 の状態数しか持たないにもかかわらず、ほぼ同等のケプストラムひずみを得られており、LT-CTVQ を用いて状態を削減しても、LS-CTVQ を用いたシステムと同等の性能を示す符号器／復号器が構成可能なことを示唆している。

表 4.3 LS-CTVQ と LT-CTVQ の記憶量と演算量の比較

-	演算量 [回数*]	記憶量 [ワード]	復号ひずみ [dB**]
LS-CTVQ	1024	320	2.72
LT-CTVQ	256	80	2.79

\* 1 受信符号あたりの距離尺度計算回数

\*\*  $E_b/N_0 = 10\text{dB}$  の時の復号ケプストラムひずみ

次に用いている LS-CTVQ と LT-CTVQ のビタビアルゴリズム実行時の演算量を比較した。CTVQ の復号では、各状態において出力可能な符号と受信符号の距離尺度の演算を行ない、その中の最小の符号を選択する。このため、ひとつ符号を受信する度に状態数  $K$  と遷移数  $M$  の積  $M \cdot K$  回の符号間距離の演算と最小距離との比較を行なう必要がある。LS-CTVQ ではトレリスの状態数  $K$  は全コードベクトル数  $N$  と同じなので、各時刻において総数  $N \cdot M$  回の距離尺度の演算を行うことになる。一方、LT-CTVQ では  $K \cdot M$  回の演算を行なう。図 4.11 の場合では、LS-CTVQ はコードベクトル数  $N = 64$ 、遷移数  $M = 16$  であるので、ひとつの符号を受信するたびに 1024 回の距離尺度の演算と比較が必要であるのに対して、LT-CTVQ では状態数  $K = 16$ 、遷移数  $M = 16$  なので、距離尺度の演算は 256 回でよい。またビタビアルゴリズムの実行時には、各状態における最小距離を示した符号とその時の距離尺度を保持しておく必要があり、判定遅延  $l_d$  と状態数  $K$  の積  $l_d \cdot K$  ワードの記憶が必要である。図 4.11 の例では、LS-CTVQ で 320 ワード、LT-CTVQ では 80 ワードである。これらをまとめると表 4.3 のようになる。このように、LT-CTVQ では LS-CTVQ と比べて性能をほとんど劣化させることなく、ビタビアルゴリズムの実行時の記憶量と演算量を大幅に低減することが可能である。

#### 4.4 本章のまとめ

本章では, CTVQ の改良法としてトレリスの遷移にラベル付けを行なう手法を提案し, 計算機シミュレーションによって性能評価を行った. 本章で提案した LT-CTVQ では, 前章の LS-CTVQ と比較して性能をほとんど劣化させることなく, 復号時の記憶量と演算量を大幅に低減できることを示した.



## 第5章 遷移を制約したベクトル量子化の2.8kbps 残差励振線形予測音声符号化システムへの 応用

### 5.1 はじめに

近年、デジタルセルラ電話等の移動音声通信システムが急速に普及している。デジタルセルラ電話用の次世代のハーフレート音声符号化システム(誤り訂正込みで5.6kbps)の標準方式として符号励振線形予測音声符号化方式(CELP)<sup>(51)</sup>の発展形のPSI-CELP方式<sup>(52)</sup>が採用されたが、今後ユーザ数がさらに増加した場合には、より高能率に音声を符号化できる符号化システムが必要になると考えられる。しかし、このような高能率の音声符号化システムは概して通信路で重畳する符号誤りに対して脆弱で、通信路符号化として誤り訂正符号化を行なうものの、わずかな通信路誤りが重大な品質劣化を引き起こす可能性がある。この問題に対処するために、高能率でしかも通信路誤りに強い符号化システムの開発が急務である。

これまで、デジタルセルラ電話用のフルレート音声符号化システムとハーフレート符号化システムは、ともにCELP方式をその基幹技術に採用している。CELP方式は非常に高品質な復号音声を得られるが、今後クォータレート音声符号化システム(誤り訂正込みで2.8kbps程度を想定)の検討を行なう場合、CELP方式はその伝送情報量の多さから適用が困難であると思われる。

一方、4.8kbps程度までのビットレートにおける音声符号化システムへの応用を目的として残差励振線形予測音声符号化方式(RELP)<sup>(53,54,56)</sup>の検討がいくつか行われている。RELPはCELPと比較すると復号音声の音質はいくらか劣るが、音源として伝送する情報が低域ろ波後にダウンサンプリングされた残差信号であるため、背景雑音に左右されやすい音源のピッチ周期の抽出が不要となる。このため、伝送誤りによって重大な品質劣化を引き起こす可能性のあるピッチ情報を伝送しないので通信

路誤りに強く、さらにシステムの構成が比較的簡易である、等の利点がある。よって、RELPの低域残差信号を高能率に伝送することができればREL Pを用いてクォータレート符号化システムを構成可能であると考えられる。そこで本章では、CTVQを用いてREL Pの線形予測係数と低域残差信号を符号化する誤り訂正込みで2.8kbpsの音声符号化システムを構成し、その性能の検討を行う。

本章では、まず5.2でCTVQを用いた2.8kbpsのREL P符号化システムについて述べ、5.3で2.8kbps符号化システムを構成するためのパラメータの決定および性能評価を計算機シミュレーションによって行った結果について述べる。

## 5.2 CTVQを用いた2.8kbps 残差励振線形予測音声符号化システム

本節ではREL P符号化システムとLT-CTVQを組み合わせて2.8kbpsで音声を符号化するシステムについて述べる。

提案するシステムの符号器と復号器の基本構成を、それぞれ図5.1と図5.2に、入力音声および分析の条件を表5.1に示す。このシステムでは、ス

表 5.1 入力音声と分析の条件

サンプリング周波数	8 kHz
フレーム長	30 ms
フレーム周期	20ms
分析次数	10 次
遮断周波数	800 Hz
ダウンサンプリング	1/5

ベクトル包絡パラメータとして線スペクトル対(LSP)係数を用い、さらに低域残差信号は、ベクトル量子化時の連続するベクトル間の相関を高めるため、離散コサイン変換(DCT)によって周波数軸上のパラメータに

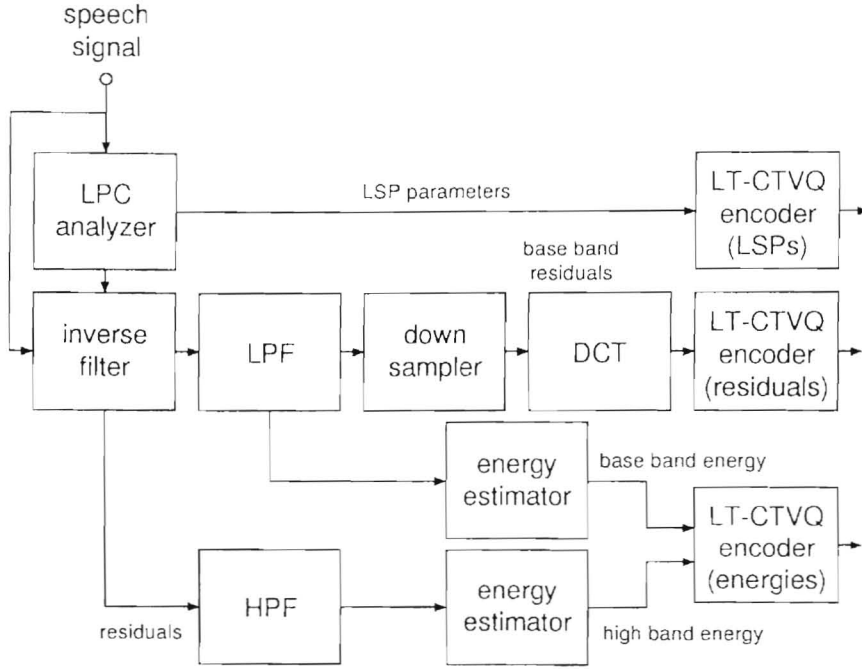


図 5.1 符号器の構成

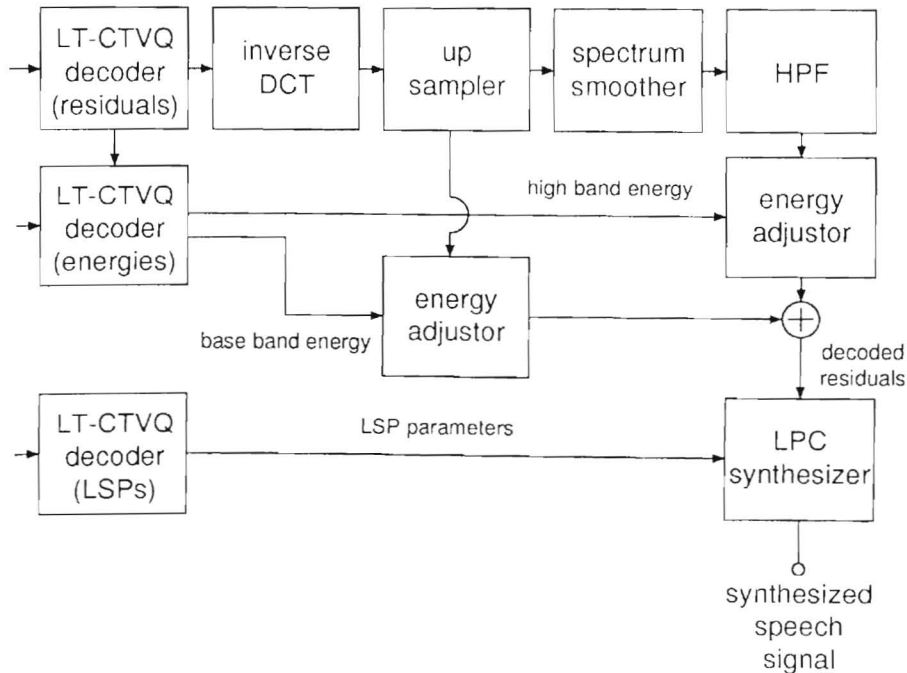


図 5.2 復号器の構成

表 5.2 パラメータへのビット割り当て

LSP係数	9 bit/frame
DCT係数	10 bit/subframe
エネルギー	7 bit/frame
合計	56 bit/frame

変換した後に、第4章の結果をふまえてLT-CTVQを用いてベクトル量子化を行なう。ここで、残差のエネルギーとLSP係数の時間的な変化は残差信号のDCT係数に比べて緩やかであることを考慮して、残差の低域成分に多くのビットを割り当て、表5.2のように各パラメータにビット配分を行なった。符号器では通常のREL Pと同様に入力音声の線形予測分析を行い、線形予測係数と残差の低域成分、低域及び高域のエネルギーを求める。LSPパラメータはLT-CTVQを用いてフレーム当たり9bitで符号化する。音源情報のうち残差の低域成分は時間的な変化が大きいいため、線形予測分析時のフレームをさらに4つのサブフレームに分割する。この各サブフレーム毎にDCT係数に変換し、LT-CTVQを用いて10bitで符号化する。低域と高域のエネルギーは伝送情報量を減らすために一つのベクトルとしてフレームあたり7bitで符号化する。

復号器では伝送された符号をそれぞれのLT-CTVQの復号器で復号する。この時にビタビアルゴリズムによって誤り訂正が行われ、通信路で重畳したビット誤りが低減される。その後、逆離散コサイン変換で得られる残差波形の低域成分をアップサンプリングし、非線形回路を通して高調波成分を発生させる。この高調波成分にスペクトル平滑化、高域フィルタリング、ゲイン調整を行ない、同じくゲイン調整した低域成分と加えあわせて線形予測フィルタを励振し、復号音声を得る。

### 5.3 計算機シミュレーション

本節では、LT-CTVQ を用いた 2.8kbps の計算機シミュレーションを行なった結果について述べる。前節までに見たように、CTVQ を用いる場合には符号器の状態数、各状態からの遷移数、ビタビアルゴリズムの復号遅延によってシステムの性能が変化する。そこで本節では、まずはじめにこれらのパラメータに関して検討を行ない、2.8kbps の RELP システムを構成するための LT-CTVQ のパラメータを決定する。パラメータの決定に際しては、通信路として二元対称通信路を仮定し、受信符号の誤り訂正には硬判定ビタビアルゴリズムを用いた。続いて、決定したパラメータを用いて 2.8kbps の RELP 符号化システムを構成し、移動通信への応用を想定して通信路にフェージングのある性能評価を行なう。ここで、LT-CTVQ のコードブックは男女各 2 名がそれぞれ発声した 50 短文を用いて学習を行ない、学習外の女性話者 2 名の発声した 50 短文を用いて評価を行なった。

#### 5.3.1 線スペクトル対パラメータの符号化

ここでは、LSP パラメータを符号化するための LT-CTVQ 符号器のパラメータについて検討を行なう。LSP ベクトル間の距離尺度としては前章までと同様に自乗距離を用いた。

$$\delta(\mathbf{l}_x, \mathbf{l}_y) = \sum_{j=1}^P \{l_x(j) - l_y(j)\}^2 \quad (5.1)$$

ここで  $P$  はベクトルの次数を表し、 $l_x(j)$  と  $l_y(j)$  はそれぞれ入力と出力の LSP ベクトルの  $j$  番目の成分を表す。また復号ベクトルの評価には次式のケプストラムひずみを用いた。

$$D_{cep} = \frac{1}{N} \sum_{i=1}^N \frac{10}{\ln 10} \sqrt{2 \sum_{j=1}^P (c_x^i(j) - c_y^i(j))^2} \quad (5.2)$$

ここで  $c_x^i$  および  $c_y^i$  は  $\mathbf{l}_x^i$  と  $\mathbf{l}_y^i$  から求めた LPC ケプストラムベクトル、 $N$  は平均操作の数である。



LT-CTVQによってLSP係数を符号化した場合の、通信路のビット誤り率と復号後のケプストラムひずみの関係を図5.3に示す。ここで符号

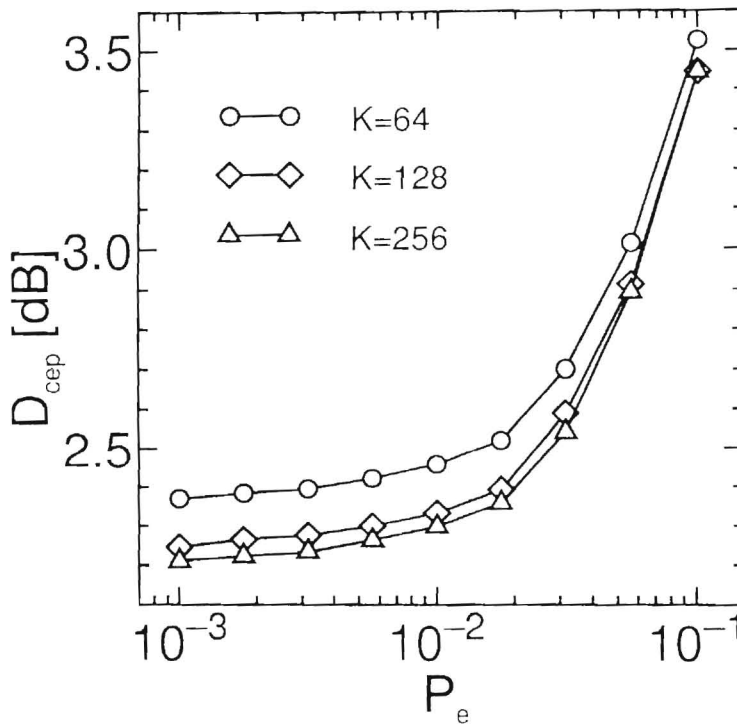


図 5.3 ビット誤り率と LSP 係数の復号ケプストラムひずみ  
( $N = 512, M = 64, l_d = 1$ )

器の全符号語数は 512, 状態数  $K$  は 64, 128, 256, 各状態から出る遷移数  $M$  は 64, ビタビ復号時の復号遅延は 1 フレームである。状態数  $K$  が多いほど誤り率が小さい場合のひずみが小さくなっているが,  $K = 128$  と  $K = 256$  の場合のひずみの差はわずかであり, 誤り訂正能力も大きな差はないものと考えられる。前節で述べたように, 復号時のビタビアルゴリズムの計算量と記憶量は状態数  $K$  と遷移数  $M$  の積  $K \cdot M$  で決定される。このため, 計算量と記憶量を抑えるためには状態数は少ない方が望ましく, LSP 係数の符号化には  $K = 128$  の符号器を用いることにする。

続いて遷移数に関して検討する。図5.4は状態数  $K$  が 64, 128, 256 の符号器において遷移数を変えた場合の遷移数と復号後のケプストラムひ

ずみの関係を、通信路の誤り率  $P_e$  が  $10^{-3}$  と  $10^{-1}$  の場合について示したものである。  $P_e = 10^{-1}$  の場合、遷移数 64 までは遷移数の減少に応じて

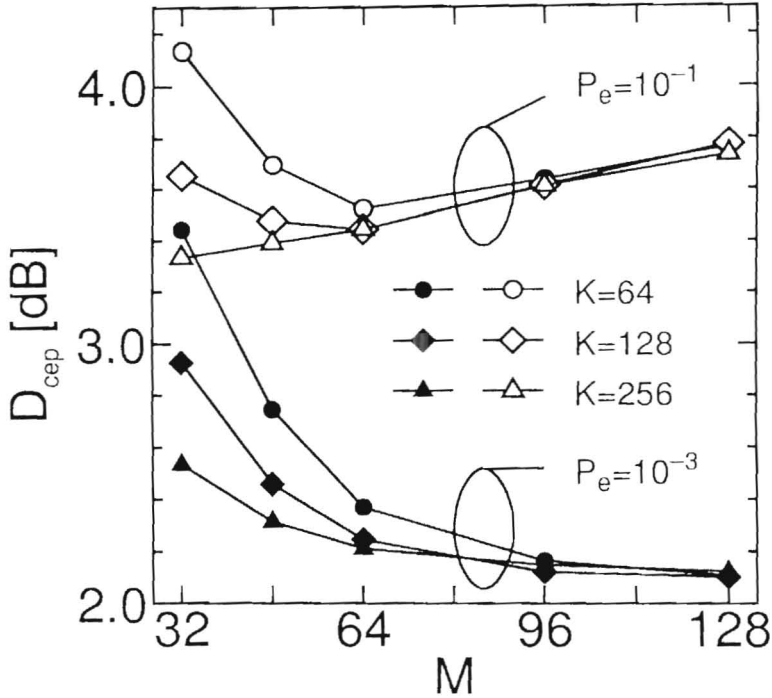


図 5.4 遷移数と LSP 係数の復号ケプストラムひずみ  
( $N = 64, M = 16, \ell_d = 5$ )

誤り訂正能力が強化されケプストラムひずみは減少するが、  $P_e = 10^{-3}$  の場合には遷移数の減少に応じて量子化ひずみが増大し、ケプストラムひずみは逆に増加する。また、遷移数が多い場合には状態数による差異はあまりないが、遷移数が少ない場合には状態数が多いほどひずみが小さくなっていることがわかる。これは遷移数を少なく抑えた場合には状態数を増やすことで量子化ひずみを低減可能なことを意味するが、復号時のビタビアルゴリズムの計算量を減らすためには状態数は少ないことが望ましい。この点からも状態数は 128 が妥当であり、遷移数は誤り訂正能力と量子化ひずみのトレードオフを考慮して 64 とする。

### 5.3.2 残差信号の符号化

ここでは、残差の低域成分の DCT 係数を LT-CTVQ で符号化する場合の符号器のパラメータの検討を行なう。

ベクトル量子化時の DCT 係数ベクトル間の距離尺度には LSP 係数と同様に自乗距離を用い、コードブックの評価には、符号化前の低域残差波形と復号された DCT 係数に逆 DCT を施して得られる低域残差波形とのセグメンタル SNR を用いた。

$$\text{SNR}_{seg} = 10 \log_{10} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{\sum_{j=0}^{L/2} f_x^i(j)^2}{\sum_{k=0}^{L/2} (f_x^i(j) - f_y^i(j))^2} \right\} \quad (5.3)$$

ここで、 $N$  は平均操作の際のセグメントの数、 $L$  はセグメント長、 $f_x^i$  と  $f_y^i$  はそれぞれ元の低域残差波形と合成した低域残差波形の FFT パワースペクトルである。

図 5.5 は LT-CTVQ によって残差の DCT 係数を符号化した場合の、通信路のビット誤り率と復号後のセグメンタル SNR の関係を表している。ここで符号器の全符号語数は 1024、状態数  $K$  は 64, 128, 256、各状態から出る遷移数  $M$  は 32、ビタビ復号時の復号遅延は 1 フレームである。また、 $\text{SNR}_{seg}$  を求める際のセグメント長  $L$  は 64 とした。LT-CTVQ の誤り訂正能力によって、どの状態数でも誤り率  $P_e$  が  $10^{-2}$  までは  $\text{SNR}_{seg}$  に大きな劣化はない。また状態数が多いほど  $\text{SNR}_{seg}$  がよいことがわかる。しかし、LSP 係数の場合と同様に状態数を 128 以上に増やしても大きな  $\text{SNR}_{seg}$  の改善が得られないことと、復号時のビタビアルゴリズムの計算量を抑えるために、符号器の状態数  $K$  は 128 とする。

図 5.6 は  $K = 128$  の符号器の遷移数を変化させたときの、遷移数とセグメンタル SNR の関係を表したものである。 $P_e = 10^{-1}$  では遷移数の減少と共に  $\text{SNR}_{seg}$  は改善されるが、それに伴い誤り率が小さい場合の  $\text{SNR}_{seg}$  が劣化している。これは、残差信号の DCT 係数は LSP 係数と比較して連続するベクトル間の相関が弱いために、遷移数を減らすと

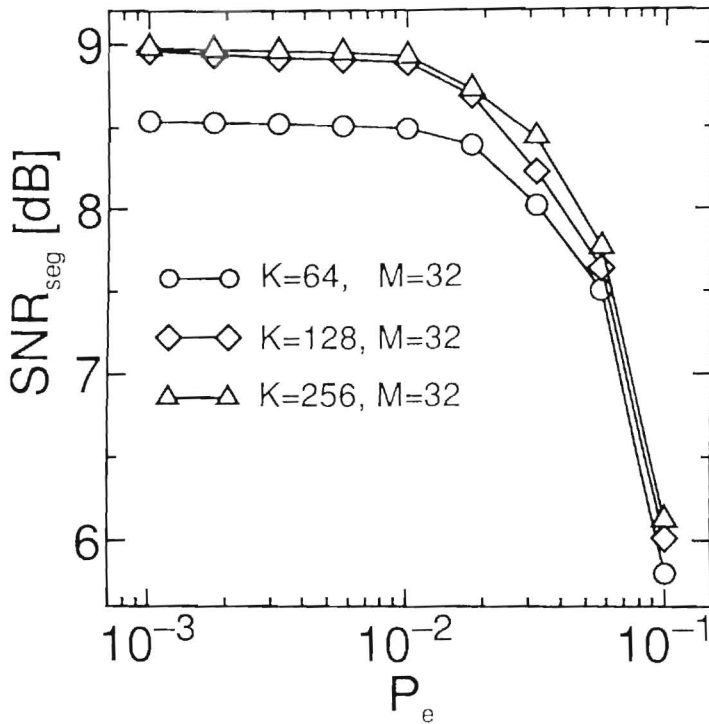


図 5.5 誤り率と低域残差信号のセグメンタル SNR

DCT 係数の変化を表現しきれなくなるためと考えられる。そこで、量子化誤差による  $SNR_{seg}$  の劣化をできる限り抑えたいうえで、ビタビアルゴリズムの計算量と記憶量を減らすために、遷移数は 128 とする。また、残差のエネルギーベクトルの符号化についても同様の検討を行い、状態数 64、遷移数 32 とした。

### 5.3.3 セグメンタル SNR による復号音声の品質評価

前節で決定したパラメータを用いて 2.8kbps 音声符号化システムを構成し、その復号音声の評価を  $SNR_{seg}$  を用いて計算機シミュレーションにより行った。  $SNR_{seg}$  を求める際のセグメント長  $L$  はここでは 256 とした。また、通信路には最大ドップラー周波数 40Hz のレイリーフェージングが存在するものと仮定し、LT-CTVQ の復号には軟判定ビタビアルゴリズムを用いた。

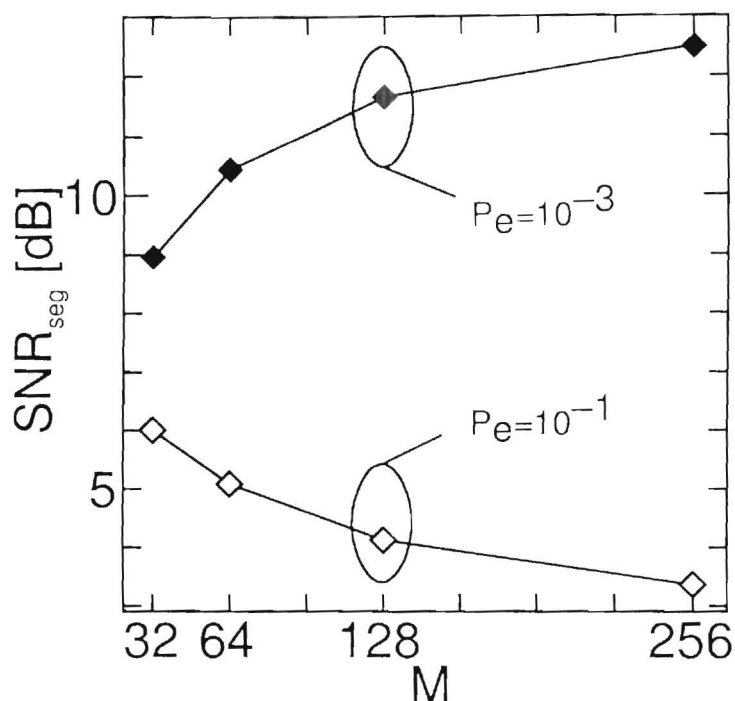


図 5.6 遷移数と低域残差信号のセグメンタル SNR

比較のために LT-FSVQ とパンクチャド畳み込み符号を組み合わせたタンデム符号化システムを構成し、同様のシミュレーションを行った. タンデム符号化システムの主な諸元を表 5.3 に示す. LT-FSVQ には LT-CTVQ と同じ状態数と遷移数のものを用い、ベクトル量子化器の規模と通信路誤りがない場合の復号品質を揃えた.

図?? に通信路の平均  $E_b/N_0$  に対する復号後のセグメンタル SNR を示す. LT-CTVQ を用いた提案システムでは、通信路の平均  $E_b/N_0$  が 10dB 程度までは  $SNR_{seg}$  に大きな劣化はなく、それ以下でも急激には劣化しない. これに対して、LT-FSVQ とパンクチャド畳み込み符号を組み合わせた従来のシステムでは、 $E_b/N_0$  が 20dB では LT-CTVQ の場合とほぼ同程度だが、それ以下になると急激に  $SNR_{seg}$  が低下する.  $SNR_{seg}$  が 6.5dB で比較した場合、LT-CTVQ を用いたシステムはタンデムシステムに対して  $E_b/N_0$  に対して約 7dB 優れている.



表 5.3 タンデム符号化システムの主な諸元

	パラメータ	符号数	状態数	遷移数	ビット数*
LT-FSVQ	LSP 係数	512	128	64	6
	DCT 係数	1024	128	128	7
	エネルギー	128	64	32	5
誤り訂正	拘束長 8	符号化率 2/3			
	パンクチャド畳み込み符号				
	ビタビ復号遅延 50 ビット				
ビットレート					2,925 bps

\* 1フレームあたりの量子化ビット数

表 5.4 パンクチャド符号の原符号

拘束長	生成多項式の係数	
5	345	237

さらに、両システムの復号音声に対して平均主観評点 (MOS) による主観評価実験を行なった。MOS 評価実験とは、被験者に合成音声を聴いてもらい、その主観的な品質を表 5.5 のように評価してもらった平均値を評価値とする評価方法である。10 人の被験者による実験結果を図 5.8 に示す。評価を行なった全ての  $E_b/N_0$  において、CTVQ を用いたシステムの MOS 評点が高くなっている。タンデム符号化システムでは平均  $E_b/N_0$  が 10dB では発声内容が聴取不可能な部分があり、さらに 5dB では全く内容の聴取が不可能であるのに対して、CTVQ システムでは  $E_b/N_0$  が 5dB でも発声内容は完全に聴取可能であった。この結果からも、CTVQ システムがタンデムシステムより優れた誤り耐性を有していることがわかる。

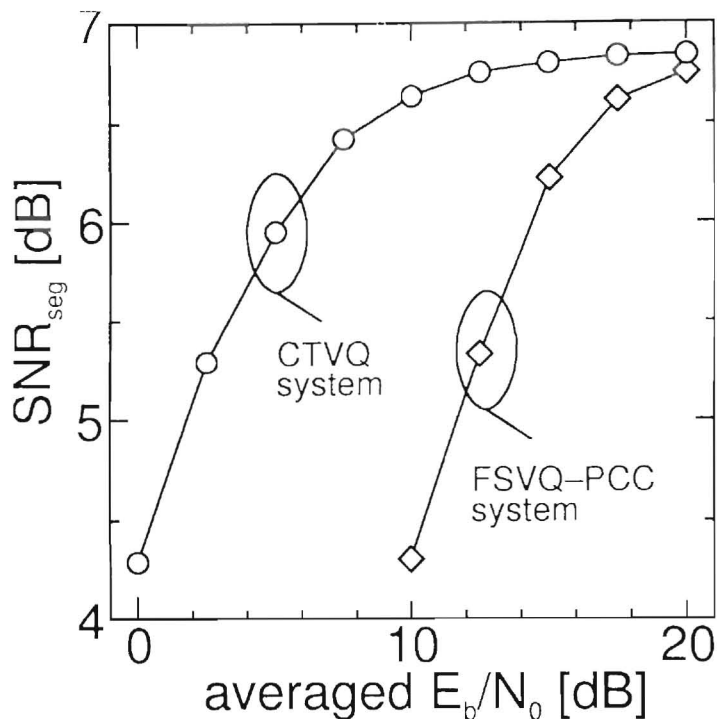


図 5.7 タンデムシステムと LT-CTVQ を用いたシステムの比較

表 5.5 MOS 評点

主観品質	評点
非常に良い	5
良い	4
ふつう	3
悪い	2
非常に悪い	1

上記の結果は、LT-FSVQ が自己回帰型のベクトル量子化法であるため、一度誤った符号が受信されるとその後すべての符号を連鎖的に誤って復号してしまふ可能性があること、復号器の複雑さを LT-CTVQ と同程度にしたうえで 2.8kbps で符号化するという制約のもとでは、あまり

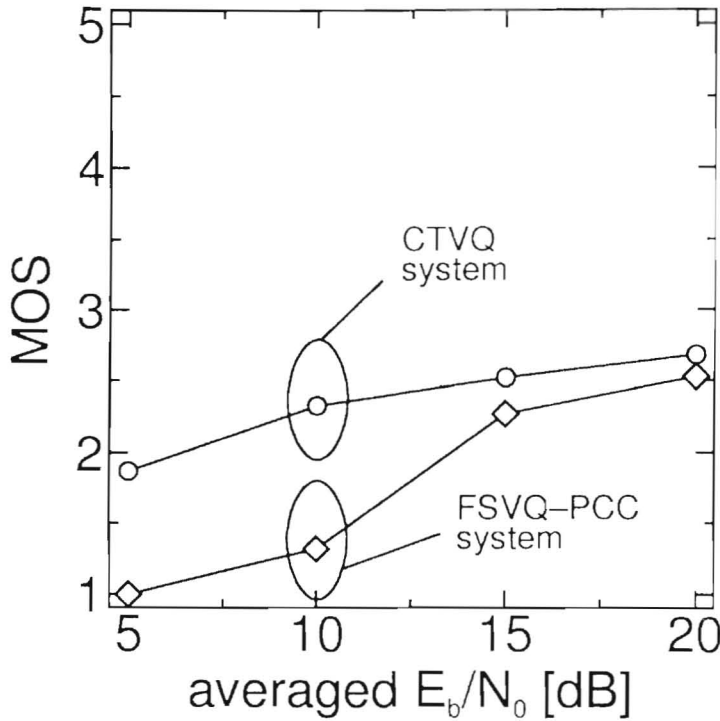


図 5.8 MOS 評価実験の結果

誤り訂正能力の高い畳み込み符号を用いることができないことなどが理由としてあげられる。タンデムシステムでは、より拘束長の長い畳み込み符号などの強力な誤り訂正符号を用いて誤り耐性を強化することもできるが、この場合には復号器の構成はさらに複雑になる。上記の結果から、LT-CTVQ を用いることで2.8kbps という低いビットレートでも劣悪な通信環境において音声伝送が可能なシステムが構成できることがわかった。

#### 5.4 本章のまとめ

本章では LT-CTVQ を RELP 符号化システムに適用して誤り訂正込みで2.8kbps で音声を符号化するシステムを構成し、計算機シミュレーションによる検討を行なった。

LT-CTVQ を用いた RELP システムをフェージングのある劣悪な通信

路に適用し、その場合の復号音声の品質評価をセグメンタル SNR によって行なった。計算機シミュレーションの結果より、LT-CTVQ を用いることで低ビットレートで通信路誤りにも強い音声符号化システムが構成可能なことを示した。

今後の課題としては、さらに話者を増やした大規模な実験による各パラメータの最適化や、システム構成の改良による復号音声のさらなる音質改善などを行なう必要がある。

## 第6章 結論

本論文では、新しい音声の情報源 - 通信路統合符号化法である遷移を制約したベクトル量子化について述べた。まず、情報源 - 通信路統合符号化用のベクトル量子化法として、符号に重畳する通信路誤りを訂正可能なベクトル量子化法を提案した。次に、その復号時の演算量と記憶量を低減する改良法を提案した。さらに、提案したベクトル量子化法の実際の音声符号化システムへ応用について検討した。ここでは、本研究の成果と今後の課題を総括する。

第3章では、新しい情報源 - 通信路統合符号化法として、状態にラベルづけを行なう遷移を制約したベクトル量子化 (LS-CTVQ) を提案した。LS-CTVQ は、ベクトル量子化のコードベクトルとトレリスの状態を対応させ、状態間の遷移を制約してまばらなトレリス上で符号化を行なうことで、ビタビアルゴリズムを用いた符号の誤り訂正を可能にする。これによって、フェージング環境下などの劣悪な通信路において、高能率な音声符号化システムを適用することが可能になる。音声のスペクトル包絡情報を表す線形予測ケプストラムを符号化して伝送する演算機シミュレーションを行ない、二元対称通信路とフェージングのある通信路における性能を評価した。フェージングのある通信路において LS-CTVQ とタンデム符号化を用いた符号化システムを比較した結果、従来のベクトル量子化と誤り訂正符号を組み合わせたタンデム符号化システムでは通信路で誤りが重畳した場合に重大な品質劣化を生じるのに対して、LS-CTVQ を用いた符号化システムでは大きな品質劣化は生じず、6dB 以上劣悪な環境下においてタンデム符号化システムと同じ復号品質が得られることがわかった。CTVQ の誤り訂正に必要な演算量と記憶量は、誤り訂正に用いるトレリスの状態数に比例する。LS-CTVQ ではコードベクトルとトレリスの状態が対応するため、ベクトル量子化の精度を向上するためにコードベクトルを増やした場合、同時に誤り訂正に必要な記憶量と演算量も増加してしまうという問題がある。



第4章では、誤り訂正に必要な演算量と記憶量を LS-CTVQ に比べて低減する改良法として、遷移にラベルづけを行なう遷移を制約したベクトル量子化 (LT-CTVQ) を提案した。LT-CTVQ では LS-CTVQ とは異なりトレリスの状態間の遷移にコードベクトルを対応させ、全コードベクトル数よりも状態数を少なく設定することで、LS-CTVQ に比べて演算量と記憶量を低減することが可能となる。LS-CTVQ と同様のシミュレーションを行ない、LT-CTVQ の性能評価を行なった。これによって LT-CTVQ では、LS-CTVQ に比べてトレリスの状態数を減らしてビタビアルゴリズムの実行に必要な演算量と記憶量を大幅に低減した (例では 1/4) 場合にも、わずかに量子化ひずみが増大するものの、ほぼ同程度の通信路誤り耐性を有することがわかった。LT-CTVQ では、トレリスの状態数と各状態から出る遷移の数によって量子化ひずみと誤り耐性のトレードオフが決まり、さらに状態数と遷移数の積によって復号時の演算量が決まる。また LT-CTVQ では状態遷移関数の選び方によって性能が大きく変化するが、本検討では一例についての検討しか行っていない。良い性能を持つシステムを構成するには、コードブックの作成において適切な状態数、遷移数および状態遷移関数を選ぶことが重要である。本検討では各種のコードブックを作成したうえで、もっとも適切なものを選んでいますが、この方法は大量のデータと長い演算時間を必要とするため、今後の改善が必要である。

続く第5章では、CTVQ の実際の音声符号化システムへの応用例として、残差励振線形予測 (REL P) 音声符号化システムに LT-CTVQ を適用し、次々世代の携帯電話システムへの適用を想定した 2.8kbps 音声符号化システムを構成した。このシステムでは、REL P システムの出力を LT-CTVQ で符号化することによって、高能率化と誤り耐性の強化を同時に行なう。LT-CTVQ を用いて構成したシステムでは、情報源符号化に有限状態ベクトル量子化、通信路符号化に畳み込み符号化を用いたタンドムシステムに比べて、7dB 環境の悪い通信路においても同じ品質の音声通信が可能であることがわかった。さらに、携帯電話等の実際のシ

システムにおいてはその伝送遅延が問題になるが、CTVQでは本質的な復号遅延は数フレーム程度であり、長い遅延を必要としない。これらの点から、LT-CTVQを用いることで、高能率かつ誤りに強い音声符号化システムを構成可能なことを示した。しかし上述したように、LT-CTVQの適切な構成を決定するためには、多くのコードブックを実際に作成して評価を行なう必要がある。本システムでは3つのLT-CTVQの符号化／復号化システムを用いており、そのそれぞれについて適切な構成を決定するために多くのデータと演算量を要した。このように、LT-CTVQを実システムに本格的に応用するためには、簡易なコードブック学習法を早急に確立する必要がある。本章で検討したシステムは次々世代の携帯電話への適用を想定したものだが、符号化システムを実際に移動体通信路に適用するためには、変調方式との適合性などの通信回線の特徴を考慮した検討を行なう必要がある。このような検討は本論文の範囲を越えるとの判断の下に行なっていないが、今後の検討が必要であろう。また、通信路の状態が良い場合のMOS評価値が3に至っておらず、実システムへの適用にはさらなる品質の向上が必要であろう。MOS値をさらに向上させるためには、AbSの手法を導入するなどの改良を行なう必要もあるだろう。

1.1や2.3.4で述べたように、ベクトル量子化による情報源 - 通信路統合符号化法には様々な手法が存在する。本論文で述べたCTVQは従来の手法と併せて用いることも可能であり、併用することでさらに劣悪な通信路に対する耐性を強化することができるものと考えられる。現在、携帯電話などの移動音声通信端末が一般の人々に急速に普及しているが、今後のさらなる普及を支えるために、ベクトル量子化はキーテクノロジーの一つとなっていくと考えられる。将来、CTVQのような劣悪な通信路に強いベクトル量子化法が応用され、より快適な音声通信環境が整備されていくことを期待する。

## 謝辞

本研究は、著者が埼玉大学大学院理工学研究科博士後期課程在学中に、同大学の小林禧夫教授、鈴木誠史教授、羽石操教授、荒木純道助教授、長谷川孝明助教授の御指導のもとで行なったものです。心から感謝致します。学会等を通じ、多くの御意見、御指導をいただきました方々に感謝致します。日頃から有益な御討論、御協力をいただきました研究室諸氏に感謝致します。また、暖かく励まし応援して頂いた家族と友人諸氏に感謝致します。

## 参考文献

- (1) A. Gersho and V. Cuperman : “Vector Quantization: A pattern-matching technique for speech coding” , IEEE Commun. Magazine, **21**, pp.15-21 (Dec. 1983).
- (2) R. M. Gray : “Vector Quantization”. IEEE ASSP Magazine, **1**, pp.4-29 (April. 1984).
- (3) J. Makhoul, S. Roucos, and H. Gish : “Vector Quantization in Speech Coding”, Proc. IEEE, **73**, 11, pp.1551-1588 (Nov. 1985).
- (4) H. Dudley : “Phonetic pattern recognition vocoder for narrow-band speech transmission”, J. Acoust. Soc. Amer., **30**, 8, pp.733-739, (Aug. 1958).
- (5) C. P. Smith : “Perception of vocoder speech processed by pattern matching”, J. Acoust. Soc. Amer., **46**, 6(pt.2), pp.1562-1571, (June 1969).
- (6) Y. Linde, A. Buzo and R. M. Gray : “An Algorithm for Vector Quantizer Design”, IEEE Trans. Commun., **COM-28**, 1, pp.84-95 (Jan. 1980).
- (7) A. Buzo, A. H. Gray Jr., R. M. Gray and J. D. Markel : “Speech Coding Based upon Vector Quantization”, IEEE Trans. Acoust., Speech and Signal Process., **ASSP-28**, 5, pp.562-574 (Jan. 1980).
- (8) B. H. Juang and A. H. Gray, Jr. : “Multiple Stage Vector Quantization for Speech Coding”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1982, pp.597-600 (1982).

- (9) M. O. Dunham and R. M. Gray : “An algorithm for the Design of Labeled-Transition Finite-State Vector Quantizers”, IEEE Trans. Commun., **COM-33**, 1, pp.83-89 (Jan. 1985).
- (10) J. Foster, R. M. Gray and M. O. Dunham : “Finite-State Vector Quantization for Waveform Coding”, IEEE Trans. Inform. Theory, **IT-31**, 3, pp.348-359 (May 1985).
- (11) B. H. Juang : “Design and Performance of Trellis Vector Quantizers for Speech Signals”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1986, pp.437-440 (1986).
- (12) 菅村 昇, 板倉 文忠 : “スペクトルパターンマッチングによる音声情報圧縮”, 電子通信学会論文誌 (A), **J65-A**, 8, pp.834-841 (1982-08).
- (13) S. Roucos, R. M. Schwartz and J. Makhoul : “SEGMENT QUNATIZATION FOR VERY-LOW-RATE SPEECH CODING”. Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1982, pp.1565-1568 (1982).
- (14) D. Y. Wong, B. H. Juang, and A. H. Gray, Jr : “An 800 bit/s Vector QUantization LPC Vocoder”, IEEE Trans. Acoust., Speech and Signal Process., **30**, 5, pp.770-780 (Oct. 1982).
- (15) S. Roucos, R. M. Schwartz and J. Makhoul : “A SEGMENT VOCODER AT 150B/S”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1983, pp.61-64 (1983).
- (16) D. Y. Wong, B. H. Juang and D. Y. Cheng : “VERY LOW DATA RATE SPEECH COMPRESSION WITH LPC VECTOR AND MATRIX QUANTIZATION”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1983, pp.65-68 (1983).

- (17) 関 章光, 宮内 新, 荒井秀一 : “メル LSP ベクトル量子化音声符号化方式”, 電子情報通信学会技術研究報告, **SP86-14** (1986).
- (18) T. Moriya and M. Honda : “Transform Coding of Speech Using a Weighted Vector Quantizer”, *IEEE J. Sel. Area. in Commun.*, **6**, 2, pp.425-431 (June 1988).
- (19) 浅川 吉章, 山崎勝也, 小窪 浩明, 市川 熹 : “相補型ベクトル量子化方式と低ビットレート音声符号化への適用”. 電子情報通信学会技術研究報告, **SP90-87** (1990).
- (20) P. Peterson, P. Jeanrenaud and J. Vandegrift : “IMPROVING INTELLIGIBILITY OF A 300B/S SEGMENT VOCODER”, *Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process.* 1990, pp.653-656 (1990).
- (21) 大室 伸, 間野 一則, 守谷 健弘 : “LSP パラメータのベクトル・マトリクス量子化”, 電子情報通信学会技術研究報告, **SP91-70** (1991).
- (22) P. C. Loizou and A. S. Spanias : “Vector Quantization of Transform Components for Speech Coding at 1200 bps”, *Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process.* 1991, pp.245-248 (1991).
- (23) Y. J. Liu : “A ROBUST 400-bps SPEECH CODER AGAINST BACKGROUND NOISE”, *Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process.* 1991, pp.601-604 (1991).
- (24) P. Jeanrenaud and P. Peterson : “SEGMENT VOCODER BASED ON RECONSTRUCTION WITH NATURAL SEGMENTS”, *Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process.* 1991, pp.605-608 (1991).



- (25) D. P. Kemp, J. S. Collura and T. E. Tremain : “Multi-frame coding of LPC parameters at 600-800 bps”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1991, pp.609-612 (1991).
- (26) K. Rose, E. Gurewitz and G. C. Fox : “Vector Quantization by Deterministic Annealing”, IEEE Trans. Inform. Theory, **IT-38**, 4, pp.1249-1257 (July 1992).
- (27) H. Suda and T. Miki : “An Error Protected 16kbit/s Voice Transmission for Land Mobile Radio Channel”. IEEE J. Sel. Area. in Commun., **6**, 2, pp.346-352 (Feb. 1988).
- (28) M. J. McLaughlin : “Channel Coding for Digital Speech Transmission in The Japanese Digital Cellular System”, 電子情報通信学会技術研究報告, **RCS90-27** (1990-11).
- (29) S. A. Atungsiri, A. M. Kondotz and B. G. Evans : “ERROR DETECTION AND CONTROL FOR THE PARAMETRIC INFORMATION IN CELP CODERS”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1990, pp.229-232 (April 1990).
- (30) R. A. Salami, K. H. H. Wong, R. Steele and D. G. Appleby : “PERFORMANCE OF ERROR PROTECTED BINARY PULSE EXCITAION CODERS AT 11.4 KB/S OVER MOBILE RADIO CHANNEL”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1990, pp.473-476 (April 1990).
- (31) T. Taniguchi, F. Amano and S. Unagame : “COMBINED SOURCE AND CHANNEL CODING BASED ON MULTIMODE CODING”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1990, pp.477-480 (April 1990).

- (32) R. Iacovo and D. Sereno : “CELP CODING AT 6.55KBIT/S FOR DIGITAL MORBILE RADIO COMMUNICATIONS”. Proc. GLOBECOM '90, pp.538-549 (Dec. 1990).
- (33) 山根 延元, 森川 良孝, 澤見 英男, 浜田 博, 畑田 和徳 : “情報源 - 通信路統合符号化におけるビットプレーンへのレート割り当ての方法”, 電子情報通信学会論文誌 (B-I), **J74-B-I**, 1, pp.74-81. (1991-01).
- (34) J. C. Hardwick : “THE APPLICATION OF THE IMBE SPEECH CODER TO MOBILE COMMUNICATIONS”, Proc. ICASSP '91, pp.249-252 (May 1991).
- (35) 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”, 電子通信学会論文誌 (B), **J67-B**, 7, pp.1-8, (1984-01).
- (36) J. R. B. De Marca and N. S. Jayant : “An algorithm for assigning binary indices to the codevectors of a multidimensional quantizer”, Proc. of IEEE Int. Conf. on Commun. 1987, pp.1128-1132 (June 1987).
- (37) J. H. Chen, H. Davidson, A. Gersho and K. A. Zeger : “Speech coding for the mobile satellite experiment”, Proc. of IEEE Int. Conf. on Commun. 1987, pp.756-763 (June 1987).
- (38) K. A. Zeger and A. Gersho : “Zero redundancy channel coding in vector quantization”, IEEE Electron. Lett., **23**, pp.654-655 (June 1987).
- (39) N. Farvardin and V. Vaishampayan : “Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding”, IEEE Trans. Inform. Theory, **IT-33**, 6, pp.827-838 (Nov. 1987).

- (40) W. B. Kleijn : "SOURCE-DEPENDENT CHANNEL CODING FOR CELP", Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1991, pp.1-4 (April 1990).
- (41) N. Farvardin : "A Study of Vector Quantization for Noisy Channels", IEEE Trans. Inform. Theory, **IT-36**, 4, pp.799-809 (July 1990).
- (42) N. Phamdo, N. Farvardin, 守谷 健弘 : "多段ベクトル量子化を用いた LSP パラメータの情報源 - 通信路結合符号化". 電子情報通信学会技術研究報告, **SP90-52** (1990-10).
- (43) R. Laroia, N. Phamdo and N. Farvardin : "Robust and efficient quantization of speech LSP parameters using structured vector quantization", Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1991, pp.641-644 (1991).
- (44) T. Moriya : "Two-Channel Conjugate Vector Quantizer for Noisy Channel Speech Coding", IEEE J. Sel. Area. in Commun., **10**, 5, pp.866-874(June 1992).
- (45) J. Rosebrock : "Shape-Gain Vector Quantization for Noisy Channels with Applications to Image Coding", IEEE J. Sel. Area. in Commun., **10**, 5, pp.918-925(June 1992).
- (46) N. Phamdo, N. Farvardin and T. Moriya : "A Unified Approach to Tree-Structured and Multistage Vector Quantization for Noisy Channels", IEEE Trans. Inform. Theory, **IT-39**, 3, pp.835-850(July 1993).
- (47) 大室 伸, 守谷 健弘, 問野 一則, 三樹 聡 : "移動平均型フレーム間予測を用いる LSP パラメータのベクトル量子化", 電子情報通信学会論文誌 (A), **J77-A**, 3, pp.303-313(1994-03).

- (48) 桂川 浩 : “耐雑音性能を考慮したベクトル量子化テーブルの構成”, 1994 年度電子情報通信学会春季大会, **A-213** (1994-03).
- (49) A. S. Spanias : “Speech Coding: A Tutorial Review”. Proc. IEEE, **82**, 10, pp.1541-1582 (Oct. 1994).
- (50) R. Zelinski and P. Noll : “Adaptive Transform Coding of Speech Signals”, IEEE Trans. Acoust., Speech and Signal Process., **ASSP-25**, 4, pp.299-309 (Aug. 1977).
- (51) M.R.Schroeder, B.S.Atal : “CODE-EXCITED LINEAR PREDICTION(CELP): HIGH-QUALITY SPEECH AT VERY LOW BIT RATES”, Proc. of IEEE Int. Conf. on Acoust. Speech and Signal Process. 1985, pp.937-940 (1985).
- (52) 三樹 聡, 守谷 健弘, 間野 一則, 大室 伸 : “ピッチ同期雑音励振源をもつ CELP 符号化 (PSI-CELP)”, 電子情報通信学会論文誌 (A), **J77-A**, 3, pp.314-324 (1994-03).
- (53) C.K.Un and D.T.Magill : “The Residual-Excited Linear Prediction Vocoder with Transmission Rate Below 9.6 kbit/s”, IEEE Trans. Commun., **COM-23**, 12, pp.1466-1474 (Dec. 1975).
- (54) M. Nakatui, D. C. Stevenson and P. Mermelstein : “Subjective Evaluation of a 4.8 kbit/s Residual-Excited Linear Prediction Coder”, IEEE Trans. Commun., **COM-29**, 9, pp.1389-1393 (Sep. 1981).
- (55) 松山 泰男 : “時空統合ベクトル量子化による情報源符号化と LPC 音声圧縮”, 情報理論とその応用研究会 第 7 回シンポジウム, pp.570-573 (1984-11).
- (56) 来山 征士, 谷戸 文広, 山本 誠一, 樽松 明 : “学習同定型スペクトル平滑法を用いる残差駆動型ボコーダ方式 (LI-RELP)”. 電子通信学

- 会論文誌 (A), **J68-A**, pp.489-495 (1985-05).
- (57) I. A. Gerson : “VECTOR SUM EXCITED LINEAR PREDICTION (VSELP) SPEECH CODING FOR JAPAN DIGITAL CELLULAR”, 電子情報通信学会技術研究報告, **RCS90-26**, (1990-11).
- (58) 小野 茂 : “音声符号化技術の最近の進展”. 日本音響学会誌, **48**, 1, pp.52-59 (1992-01).
- (59) J. Chen, R. Con, Y. Lin, N. Jayant and M. Melchner : “A low-delay CELP coder for the CCITT 16 kb/s speech coding standard”, IEEE J. Sel. Area. in Commun., **10**, 5, pp.830-849 (June 1992).
- (60) 小野 茂 : ” 構造化ベクトル量子化に基づく 3.8kb/s 音声符号化方式の検討”, 電子情報通信学会技術研究報告, **SP92-134** (1993-02).
- (61) A. J. Viterbi and J. K. Omura : Principles of Digital Communication and Coding, McGRAW-HILL.
- (62) J. B. Cain, G. C. Clark, Jr. and J. M. Geist : “Punctured Convolutional Codes of Rate  $(n-1)/n$  and Simplified Maximum Likelihood Decoding”, IEEE Trans. Inform. Theory, **IT-25**, 1, pp.97-100 (Jan. 1979).
- (63) 安田 豊, 平田 康夫, 小川 明 : “ヴィタビ復号の容易な高符号化率畳み込み符号とその諸特性”, 電子情報通信学会論文誌 (B), **J64-B**, 7, pp.573-580 (1981-07).
- (64) 管村 昇, 板倉 文忠 : “線スペクトル対 (LSP) 音声分析合成方式による音声情報圧縮”, 電子通信学会論文誌 (A), **J64-A**, 8, pp.599-606 (1981-08).

- (65) S. E. Levinson, L. R. Rabiner, and M. M. Sondhi : “An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition”, J. Bell Sys. Tech., **62**, 4, pp.1035-1074 (Apr. 1983).
- (66) L. R. Rabiner and B. H. Juang : “An Introduction to Hidden Markov Models”, IEEE ASSP Magazine, pp.4-16 (Jan. 1986).
- (67) 塙 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”, 電子情報通信学会論文誌 (A), **J77-A**, 6, pp.905-914 (1994-06).
- (68) 塙 雅典, 長谷川 孝明 : “CTVQ 符号の軟判定復号について”, 1993 年度電子情報通信学会秋季大会, **A-189** (1993-09).
- (69) M. Hanawa and T. Hasegawa : “Performance evaluation of Constrained Transition Vector Quantization o Rayleigh fading Channels”, Proc. of 1993 Int. Sympo. on Nonlinear theory and its Applications., pp.879-882 (Dec. 1993).
- (70) 塙 雅典, 長谷川 孝明 : “遷移を制約したベクトル量子化の改良法とフェージング通信路への適用について”, 電子情報通信学会技術研究報告, **IT94-19** (1994-05).
- (71) 塙 雅典, 長谷川 孝明 : “遷移を制約したベクトル量子化を用いた 2.8kbps RELP 音声符号化システムについて”, 1994 年度電子情報通信学会秋季大会, **D-401** (1994-05).
- (72) M. Hanawa and T.Hasegawa : “Application of the Labeled Transition CTVQ to the 2.8 kbps RELP speech coding system on fading channels”, Proc of 1994 Int. Workshop on Multi-dimensional Mobile Communications, pp.27-30 (Nov. 1994).