

## LSP 係数の性質を考慮した音声符号化の改善

足立 征士<sup>†</sup> 鈴木 基之<sup>††</sup> 任 福継<sup>††</sup><sup>†</sup> 徳島大学大学院 先端技術科学教育部 <sup>††</sup> 徳島大学大学院 ソシオテクノサイエンス研究部

## 1 はじめに

私たちは音声の性質に応じた符号化に着目し、できるだけ品質を落とさずに低ビットレートな符号化をすることを目指している。

音声信号の符号化において、入力系列の時間方向の情報をも効率的に符号化できるセグメント量子化法が有効である。ここで重要となるのがどのようにセグメント分割を行うかである。これを解決するために、鈴木らは入力系列の類似性に従って分割する ML-BEATS 法 (Maximum Likelihood - Boundary Estimation Algorithm for Time Sequence)[1] を提案した。更に我々はこの方法をもとに、LSP 係数の性質を考慮したセグメント量子化法を提案し比較実験を行った [2]。しかし、符号モデルの状態数が少ない時に HMM 数が極端に少なかったため、十分にケプストラム歪を下げることができなかった。

そこで今回は、状態数と学習サンプルを増やして、また、次元ベクトルによって別々に Huffman 符号化を行い、その効果を評価する。

## 2 LSP 係数の性質を考慮した符号モデル構築

## 2.1 LSP 係数の次元に関する性質

LSP 係数は、ある時刻で次元が縮退・分離し、それ以降の次元がひとつずつずれてしまう性質がある。図 1 はある音声について LSP 係数を示している。この図において、175 フレーム目付近で 1 次元目の LSP 係数が分離している。1 次元目が分離することで、分離する前のフレームでは 2 次元目の LSP 係数だったが、3 次元目の LSP 係数になっていることがわかる。よって、LSP 係数において「次元」という情報は意味をなしていないことがわかる。

ML-BEATS 法を用いた先行研究 [1] では、10 次元の LSP 係数を 3 次、3 次、4 次の 3 つのベクトルとし、それぞれで符号モデルを構築した。結果は、従来法である G.729 よりもビットレートは低くすることができたがケプストラム歪は大きくなった。原因として、

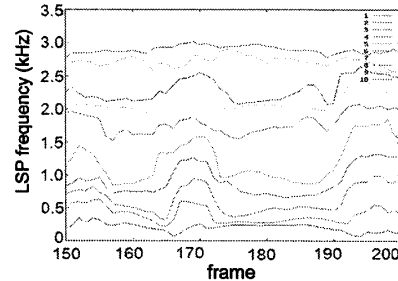


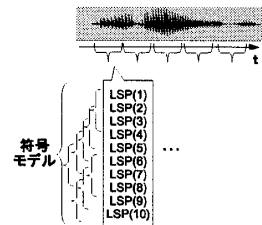
図 1: LSP 係数の次元に関する性質

LSP 係数を数次元ずつベクトルとして捉え、それぞれで符号モデルで構築しているためである。例えば、ある時刻において 1 次元目が分離してしまった場合、それ以前は、1~3 次元ベクトルを学習データとして構築した符号モデルを用いて復号化することができるが、それ以降では本来 3 次元目であった値が 4 次元目の値となっているので、1~3 次元ベクトルを学習データとして構築した符号モデルでは適応できない。そこで、LSP 係数の次元に関する性質を考慮した符号モデルを構築する方法を提案する。

LSP 係数が、とある時刻において次元の縮退・分離が発生した場合、次元がひとつずつずれてしまう。そこで、あらかじめずれを考慮した符号モデルを学習させることで、この問題に対処できると考えられる。

## 2.2 符号モデルの構成

10 次元の LSP 係数に対して、低次元から  $n$  個 ( $1 \leq n \leq 10$ ) ずつのベクトルを 1 次元ずつシフトしていくように分割し、これらすべてを 1 つの符号モデルとして構築する。図 2 は、 $n$  が 3 の時の符号モデル構築を示している。

図 2:  $n=3$  の時の符号モデル構築

このようにして、符号モデルを構築し符号化を行う

An Improvement of Speech Coding considering a Property of LSP Parameter

<sup>†</sup> Masashi ADACHI

<sup>††</sup> Motoyuki SUZUKI

<sup>††</sup> Fuji REN

Graduate School of Advanced Technology and Science, Tokushima University (<sup>†</sup>)

Institute of technology and science, Tokushima University (<sup>††</sup>)

ことで, LSP 係数の次元に関する性質を考慮し, ケプストラム歪を抑えることができる.

### 3 LSP 係数の符号化

#### 3.1 概要

前章で LSP 係数の符号モデル構築について示した. この符号モデルを用いて LSP 係数を符号化・復号化を行う [2]. 符号化する時は, 入力を  $1 \sim n, n+1 \sim 2n, \dots$  と分割し, それぞれ符号化を行う. その後, Huffman 符号化を行って伝送する.  $n$  が 3 の時のような場合は,  $1 \sim 3, 4 \sim 6, 7 \sim 9, 8 \sim 10$  次元情報を送信することになり, 8,9 次元目は重なってしまう. このような次元は平均をとることとする.

#### 3.2 次元別 Huffman 符号化

符号モデルは, LSP 係数のすべての次元を用いて構築している. そのため, ある次元を符号化する際には符号モデルの一部しか利用していないと考えられる. よって, すべての学習データを符号化した時の統計情報を用いて Huffman 符号化するよりも, 符号化を行う次元の学習データを符号化した時の統計情報のみを用いて Huffman 符号化を行った場合の方が, 符号長が短くすることができ, ビットレートが下げられると考えられる.

### 4 実験

提案手法と先行研究 (10 次元の LSP 係数を 3,3,4 次元に分割する方法) の比較実験を行った.

実験条件とし, LSP 係数の算出には, ITU-T G.729 符号化器によって計算される量子化前のものを用いて符号モデルの学習を行った. その他を表 1 に示す.

表 1: 実験条件

サンプリング周波数	8kHz
フレーム周期	10ms
使用データベース	JNAS
学習データ	30 人 $\times$ 50 文章
テストデータ	12 人 $\times$ 50 文章

先行研究の手法では, 1 つの符号モデルの状態数をそれぞれ 40 から 120 まで 40 状態きざみで設定し,  $3^3 = 27$  通りの組み合わせについて実験を行っている. 提案手法については, 符号モデルの状態数が 40 から 120 まで 40 状態きざみで設定し,  $n$  が 2,3,4,5 の時の,  $3 \times 4 = 12$  通りについて実験を行った. 実験結果を図 3 に示す. 塗りつぶした点が送信する次元ベクトルと同じ学習データを符号化した時の統計情報を計算し, 別々に Huffman 符号化を行った方法である.

結果として, 全体的に先行研究の手法と比べてビッ

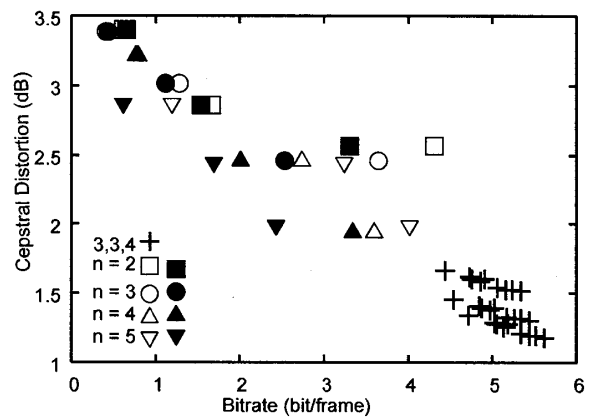


図 3: ビットレートとケプストラム歪の関係

トレートを低くできたが, ケプストラム歪が高くなってしまった. しかし, 状態を分割するごとに格段にケプストラム歪が下がっていることがわかる. このことから, もっと状態数を増やすことで, 先行研究の手法より, ビットレートを低く, かつ, ケプストラム歪の低い符号化を行えると考えられる.

また, 次元によって別々に Huffman 符号を行った方法について, ビットレートが低くなった. これは, 次元ベクトルによって統計情報が違うためだと考えられ, 有効性を示すことができた.

### 5 おわりに

LSP 係数の次元に関する性質を考慮することで, ケプストラム歪を抑えられると考えられるため, ML-BEATS 法を用いて LSP 係数の性質を考慮した符号モデル構築を行い符号化実験を行った. また, 復号化する次元数ベクトルと同じ学習データを符号化した時の統計情報を用いて Huffman 符号化を行うことで, よりビットレートを抑えることができた.

更に符号モデルを状態分割することで, よりケプストラム歪を抑えられることが考えられる.

### 参考文献

- [1] 鈴木基之, 木幡稔, 伊藤彰則, 牧野正三: ML-BEATS 法を用いた LSP 係数の極低ビットレート符号化法の検討, 東北大学電気通信研究所 第 343 回 音響工学研究会 (2006).
- [2] 足立征士, 鈴木基之, 任福継: ML-BEATS 法を用いた LSP 係数のセグメント量子化法の検討, 電気学会電子・情報・システム部門大会論文集, pp. 723-725 (2009).