

7E-1

人間の聴覚特性を考慮したLSPパラメタの能率的符号化

荒井 秀一 陳 静 三井田 惇郎
 千葉工業大学 工学部 情報工学科

1. はじめに 低ビットレートにおける音声信号の符号化においては、線形予測分析に基づく分析合成系がよく用いられている。この際の分析法としてはLPC, PARCOR, LSP等が提案されているが、パラメタの符号化効率の面ではLSP分析が優れており、この分析法を用いた符号化が幾つか提案されている⁽¹⁾⁻⁽⁴⁾。しかし、これらの方法も伝送符号量対音質のコストパフォーマンスはまだ十分とはいえない。本稿では人間の聴覚特性を考慮した符号化法を提案し、LSPパラメタ符号化の高能率化を図る。

2. 人間の聴覚特性 ここでは以下のような人間の聴覚特性を考慮に入れて符号化法を構成することを考える。1) パワーの大きい信号と小さい信号が混在している場合、パワーの大きい信号にパワーの小さい信号がマスクされる。2) 信号の時間的な変化については、変化のない時点(定常部)より変化のある点(遷移部)に敏感である。

3. LSPパラメタと聴覚特性 音声のスペクトル包絡とLSPパラメタの関係を図1に示す。LSPパラメタはホルマントと密接な関係があり、2個の近接したLSPパラメタによりスペクトル包絡の極を表現している。しかし全てのLSPパラメタが極の近傍に存在しているわけではなく、極とは離れた位置にもLSPパラメタは存在する。従来

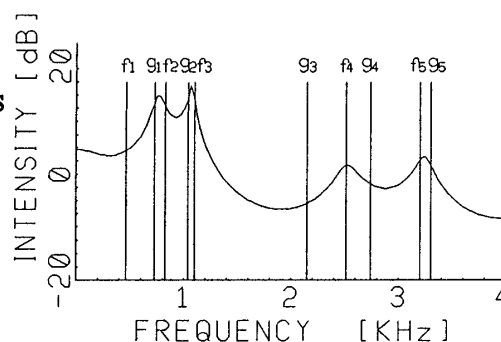


図1 スペクトル包絡とLSP

の符号化法においては全てのLSPパラメタを同様に扱っていたが、ここでは極に近いパラメタ、すなわちパワーの大きい部分を表現すると考えられるパラメタほど量子化精度を上げる事を考える。これにより人間の聴覚特性(1)を考慮した符号化を行う。また、フレーム間のスペクトル包絡の変化を尺度とし、変化の小さいフレームは可能な限り伝送符号量を削減し、変化の大きいフレームで忠実度の高い符号化を行うことを考える。これにより人間の聴覚特性(2)を考慮した符号化を行う。

4. 符号化法 本符号化法の構成図を図2に示す。まず最初に第iフレームm番目のLSPパラメタ $f_i(m)$ のスペクトル包絡への寄与率 $\alpha_{i(m)}$ を以下のように計算する。

$$\alpha_{i(m)} = \exp(-a \cdot (\text{distance}(m) + k \cdot \text{band_width}(m))) \quad (1)$$

但し、 k は定数であり、 $\text{distance}(m), \text{band_width}(m)$ は、スペクトル包絡のピーク周波数とバンド幅をそれぞれ $\text{freq_p}(n), \text{band_p}(n) \quad n=1, 2, \dots, N$ とすると

$$\begin{aligned} \text{pair}(m) &= n \text{ of } \min(|f_i(m) - \text{freq_p}(n)|) \quad n=1, 2, \dots, N \\ \text{distance}(m) &= |f_i(m) - \text{freq_p}(\text{pair}(m))| \\ \text{band_width}(m) &= \text{band_p}(\text{pair}(m)) \end{aligned} \quad (2)$$

として求める。

An Efficient Coding of LSP Parameters Using Characteristics of Human Hearing System

Shuuichi ARAI, Jing CHEN, Yoshiro MIIDA

CHIBA Inst. of Tech.

次に前フレームの送信パラメタ $f_{i-1}(m)$ と $f_i(m)$ よりスペクトル変化 $diff$ を求め、この変化量の大小により、VQコードを伝送するか否かを定める ($Vmode_i$)。 $diff$ が大きい場合にはVQを以下の距離計算に基づいて行う。

$$d = \sum_{m=1}^M \{ \alpha_i^2(m) (f_i(m) - fc_i(m))^2 \} \quad (3)$$

但し、 $fc_i(m)$ はVQコード c_i のセントロイドであり、これによりスペクトルのピークに重みをおいたマッチングを行う。以上の処理により選択された $f'_i(m)$ を用いて、次に $f_i(m)$ とのスペクトル包絡歪 $dist$ を計算する。この歪量により付加コードを伝送するかを決定する ($Amode_i$)。付加情報を伝送する場合にはスペクトル包絡への寄与率 $\alpha_i(m)$ に応じて符号を割り当てる。以上のような手続きにより符号化を行い、 $Vmode_i, Amode_i$ 、必要に応じて $add_i(m), c_i$ が伝送される。

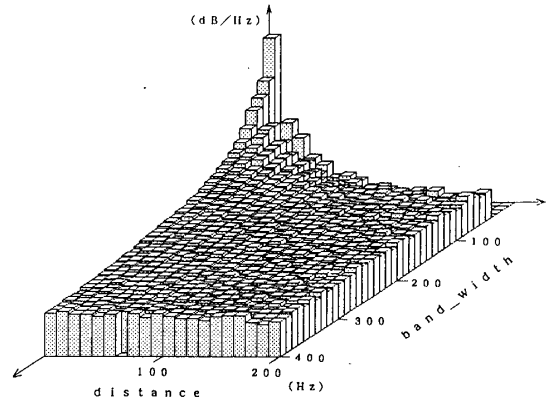


図3 スペクトル感度

5. 実験 まず(1)式のスペクトル寄与率 α の妥当性を検証するために以下のような実験を行った。20000フレームのLSPパラメタ $f_i(m)$ に関し(2)式の $distance(m), band_width(m)$ を求めそのスペクトル感度を算出した。その結果を図3に示す。これより $distance(m), band_width(m)$ が小さいほどスペクトル感度が高く、量子化精度を上げれば良い事が分かる。この実験より(1)式の a, k は0.01, 0.5に定めた。

次に提案する方法により符号化実験を行った。実験に用いた音声は、8kHzサンプリング12bit男女各2名の文書朗読である。LSP分析次数は10、フレーム周期は5msecとした。合成音の

評価を聞き取りによる主観評価により行ったところ良好な結果が得られた。

6. まとめ 人間の聴覚特性を考慮したLSPパラメタの符号化法を提案し、符号化効率の向上を実現した。

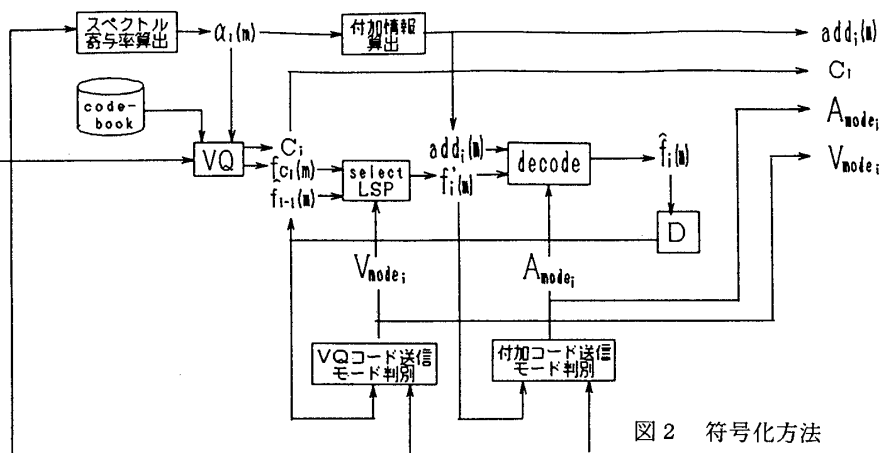


図2 符号化方法

参考文献

(1) 管村, 板倉: "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", 信学論(A), J64-A, 8, pp599-606
 (2) 管村, 板倉: "可変フレームによる低ビット擬ホルマントボコーダ", 音響学会音声研資, S77-36
 (3) 荒井, 小沢: "差分LSPパラメータ列相互の相関に着目した能率的符号化", 信学論(A), J71-A, 1, pp1-8
 (4) Arai, Ozawa: "An efficient coding for LSP parameters using fuzzy reasoning", Trans. IEICE, Vol. E72, No. 4, april 1989.