

# LF-6 2side PWR法によるVoIPにおけるパケット損失の隠蔽 Packet Loss Concealment in VoIP Using Two-Side PWR Technique

小牧 憲子<sup>†</sup> 青木 直史<sup>†</sup> 山本 強<sup>†</sup>

Noriko Komaki Naofumi Aoki Tsuyoshi Yamamoto

## 1 はじめに

VoIPではパケット損失等の通信上のエラーが音質劣化の要因となっている。本稿では、エラー隠蔽の一手法であるPWR (Pitch Waveform Replication) 法について考察する。従来法は、過去に受信された音声データのみを用いる1sideの手法であるが、提案法は遅延バッファを1パケット分大きくすることで、未来に受信される音声データを用いる2sideの手法となっている。提案法ではピッチ抽出にかかる遅延を小さくすることを目的として、自己相関ではなく相互相関に基づくピッチ抽出について検討した。

## 2 提案法の原理

PWR法は損失フレーム直前の1ピッチ波形を繰り返すことでエラー隠蔽を行う。そのため、ピッチ抽出の失敗が音質低下につながる事が指摘されている[1]。そのほかにも、図1(2)に示すように、同じピッチ波形を繰り返すことによる接続部分での位相ずれ(c)、発話開始や終了部における不適切な補間(a,b)が音質低下の原因となる。

波形のピーク検出に基づく初期のPWR法では非定常部でのピッチ抽出に問題があった。そのため、近年では波形の自己相関によるピッチ抽出が一般的に適用されている[2]。しかしながら、1sideからの補間では上述したそのほかの問題を解決することは困難であり、図1(3)に示すように2sideからの補間が望ましい[1]。そこで、本研究ではPWR法を2sideに拡張する手法について検討した。

2sideで問題となるのは未来に受信される音声データを用いることで遅延が増大することである。一般に自己相関によるピッチ抽出では、最低でも2ピッチ周期が必要であり、ピッチ周期の最大値を15msとすると、30ms必要となる。したがって20msをフレーム長とするVoIPでは、2sideにすると2パケット分の遅延が生じてしまう。本研究ではこれを1パケットに抑えることを目的として相互相関によるピッチ抽出について検討した。

図2に提案法の処理過程を示す。提案法ではピッチ抽出のために、損失フレーム直前の音声サンプルをテンプレート $x(m)$ とし、それよりも過去の音声サンプル $y(m)$ との相互相関 $C(n)$ を計算する。すなわち、

$$C(n) = \sum_{m=0}^{M-1} x(m)y(n+m), \quad (0 \leq n \leq N) \quad (1)$$

ここで、 $C(n)$ を最大にする $n$ が1ピッチ周期となる。また、フレーム長を $F$ とすると、テンプレート長 $M$ と探索長 $N$ には次の関係がある。

$$M \leq F - N \quad (2)$$

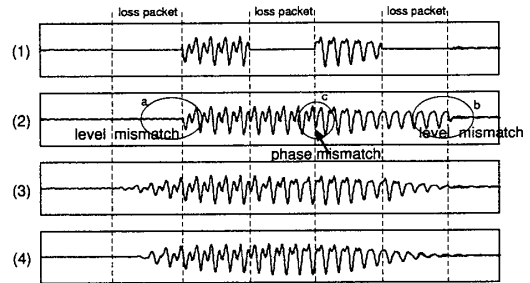


図1 補間結果比較 (1) 補間前 (2) 1side (3) 2side (4) 原音声

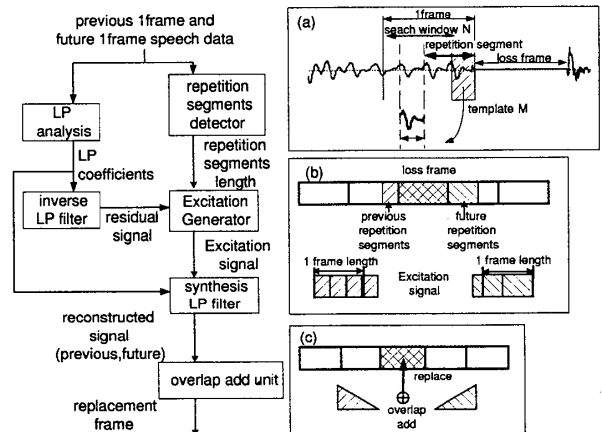


図2 提案法の処理過程

$N$ は提案法によって推定できる最大のピッチ周期であり、これを15msとすると、テンプレート長の最大値は5msとなる。

なお、提案法では線形予測によって得られた残差信号を用いて1ピッチ波形の繰り返しを行い、合成フィルタを通すことで滑らかな補間を行っている[2]。

## 3 自己相関との性能比較

相互相関を用いたピッチ抽出の有効性を調べるために自己相関との比較を行った。テンプレート長を5msとし、1sideと2sideの場合について比較を行った。結果を図3に示す。結果として両者に顕著な違いが見られなかったことから、相互相関を用いてもピッチ抽出には大きな影響を及ぼさないことが推察される。

<sup>†</sup>北海道大学大学院工学研究科

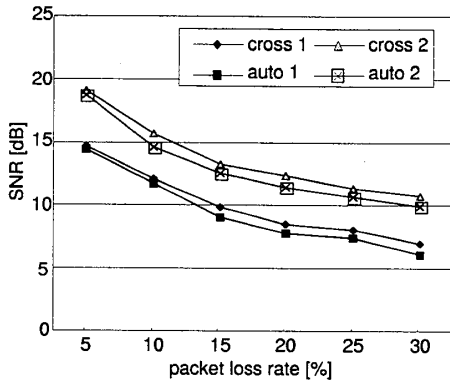


図 3 自己相関と相互相関の SNR 比較

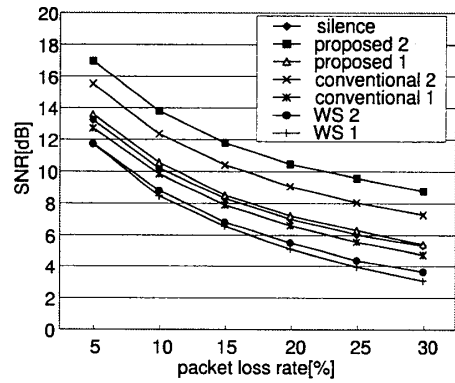


図 5 客観評価実験の結果

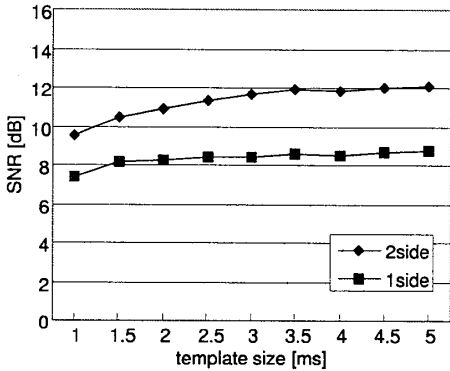


図 4 テンプレート長と SNR の関係

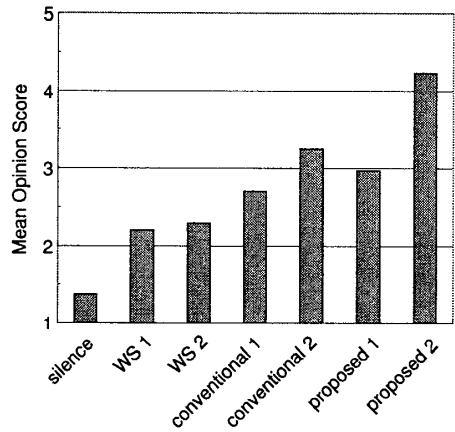


図 6 主観評価実験の結果

#### 4 テンプレート長による性能比較

テンプレート長の違いによる SNR の比較を行った。テンプレート長は 1~5ms までの 0.5ms 刻みとした。パケット損失率 5~30% の SNR を平均した結果を図 4 に示す。結果として 1side と 2side のいずれも極端に短いテンプレートでなければ SNR に大きな影響を及ぼさないことが推察される。

#### 5 客観評価実験

SNR による客観評価によって従来法との比較を行った。なお、ここでは初期の PWR 法を従来法と定義した。また、1 フレーム単位で波形置換を行う Waveform Substitution (WS) 法についても比較を行った [1]。音声データは男女各 1 話者のそれぞれ 5 文とし、5~30% の確率でランダムにパケット損失を施した後、エラー隠蔽を行った。提案法、従来法、WS 法すべて 1side と 2side について実験を行った。

結果を図 5 に示す。同一の手法であれば 2side の方が優れていることがわかる。特に提案法は最も評価が高いことがわかった。

#### 6 主観評価実験

客観評価ではエラー隠蔽を行っていない silence 法が WS 法よりも高い評価を得ている。より正確を期すために主観評価実験として、10 名の被験者による MOS 評価を行った。手法の違いによる品質の優劣はパケット損失率によらないことから、ここでは 15% として呈示刺激音を作成した。

結果を図 6 に示す。silence 法を除き客観評価の順番と一致する結果となっている。提案法はここでも最も評価が高いことがわかった。提案法については発話の開始や終了部におけるエコー感、位相ずれによるクリック音といったアーチファクトが低減されているとの内観報告を得た。

#### 7 むすび

本稿では 2side に拡張した PWR 法について検討した。提案法は遅延バッファを 1 パケット分大きくするだけで従来法よりも適切にエラー隠蔽を行える可能性がある。実際のシステムに提案法を適用し、その有効性について検討することが今後の課題である。

#### 参考文献

- [1] Goodman, "Waveform substitution techniques for recovering missing speech segments in packet voice communications," IEEE Trans. Acoust., Speech, Signal Processing, ASSP-34(6):1440-1448, Dec. 1986
- [2] Committee T1 Telecommunications. "ANSI T1.521-1999, A Packet Loss Concealment Technique for Use with ITU-T Recommendation G.711" ATIS/ANSI, 1999