

## AMR-WB+における線形予測次数の適応可変化\*

張凡 (東大院・情報理工), 守谷健弘, 鎌本優, 原田登 (NTT・CS研),  
西本卓也, 小野順貴, 嵯峨山茂樹 (東大院・情報理工)

## 1 はじめに

本稿では, 音声符号化で用いられる線形予測の次数を適応的に変化させることで情報量を削減する効果的な手法を提案する. これまで有線・無線を問わず品質が担保されているサービスでは符号化ビット列の伝送レートも固定であることが望まれ, 目的のビットレートに応じて揺らぎの無いビット長を伝送していた. 近年普及してきている VoIP (Voice over IP) では多少の揺らぎが許されるため, 必ずしも各フレームにおいて一定のビット長である必要は無く, ある程度の可変ビットレートでの伝送が普及しつつある.

我々は, 音声符号化で用いられる線形予測次数が必ずしも定められた次数まで必要ではないことに着目し, 予測係数を無駄に送らなくても音質が劣化しない場合には, 係数の伝送を省略することによって情報量を削減する方法を考案した. 国際標準化技術である AMR-WB+[1, 2] において提案手法を実装し, 従来法と比較して約 2% の情報量削減が可能であることを検証した.

## 2 AMR-WB+を基にした線形予測次数の適応可変化

AMR-WB+は線形予測分析を用いて, 音声・音楽両方の信号に対応するために開発された国際標準技術である. 線形予測次数は 16 次に固定されているが, 予備実験により無声音や雑音のフレームでは予測次数を下げて同程度の音質を得られることが分かった [3].

そこで我々は, 音質の劣化が無視できる程度の場合に, 線形予測パラメタ (immittance spectral frequencies; ISF) の数を 16 から 8 に減らすことでどれだけ情報量を削減可能かを検証した. 16 次の ISF を伝送するために分割多段ベクトル量子化を用いてフレームあたり 46 bit の情報量を用いているが, 8 次の ISF では 16 bit で十分である. 例えば図 1 に示されているように, 8 次でも十分にスペクトルを表現できる場合は 8 つのパラメタで十分であるし, 16 個のパラメタがなければ表現できない場合もある. どちらを選択す

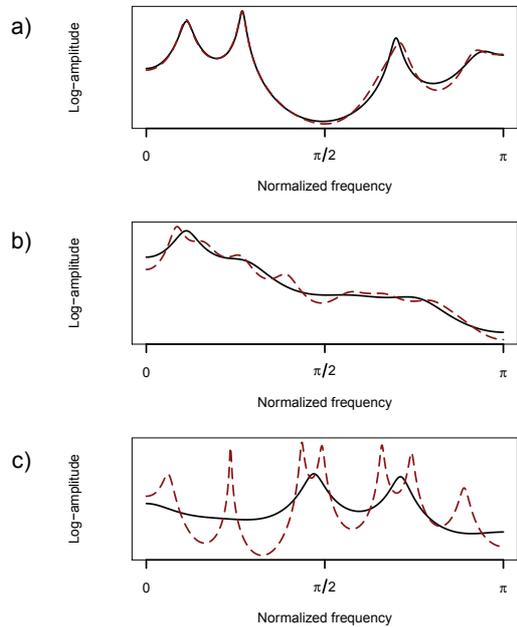


Fig. 1 Comparison of power spectra obtained with LP of order 16 and 8. Dashed line represents the results of the higher prediction order, and solid line the lower order. Three plots show three cases when lower order is acceptable (a), the threshold case, i.e. when switching between orders occur (b), and the case when higher order is chosen (c).

べきが明らかでない場合には, 16 次および 8 次の分析結果の比較規準として SSNR (segmental signal-to-noise ratio) を用いた.

まず各フレームに対し各次数の符号化が行われ, それぞれの SSNR を比較して適切な次数を選択する. 例えば, 8 次の結果から得られる SSNR が 16 次の結果の SSNR の 95% 以上に達した場合に, アドホックではあるが, 予測次数を 8 次にする方策をとった. 図 2 で示されるように, 実線 (16 次) と破線 (8 次) の差が小さければ音質の劣化は少ないと考え予測次数を下げることで情報量削減をおこなう. その結果, 表 1 に示されるように, 各ビットレートにおいて約 2 割のフレームで次数の削減の判断がされることが確認された.

\* Adaptive prediction order scheme for AMR-WB+, by Fan Zhang (The Univ. of Tokyo), Takehiro Moriya, Yutaka Kamamoto, Noboru Harada (NTT), Takuya Nishimoto, Nobutaka Ono, and Shigeki Sagayama (The Univ. of Tokyo).

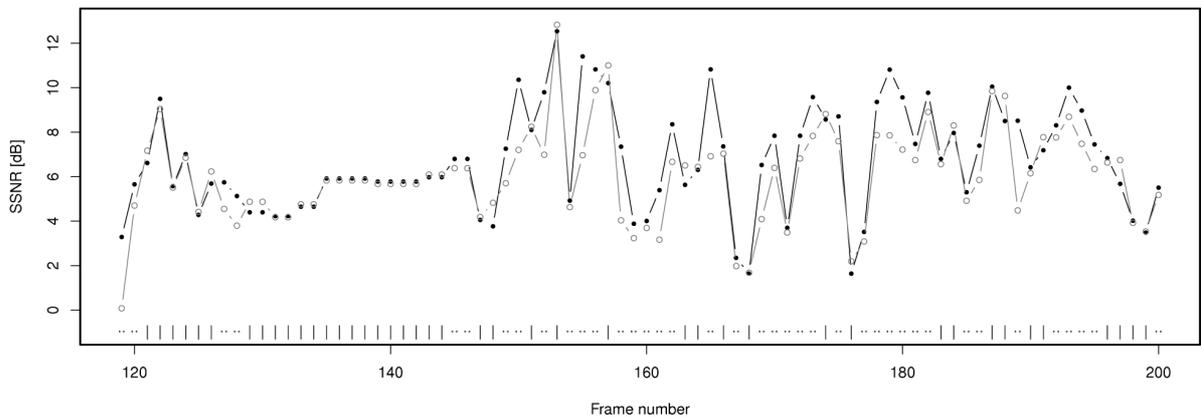


Fig. 2 An visualization of SSNR-based order selection for an example data. Solid line shows the SSNR obtained for sixteenth order prediction, while dash-dotted line for the eighth order. The vertical bars at the bottom indicate frames where lower prediction order is selected, i.e. eighth order SSNR is larger than 95% of the sixteenth order SSNR.

Table 1 Average percentage of frames selected as lower prediction order

bit rate (kbps)	12	16	24	32
Lower order (%)	22.22	23.96	21.77	22.45

Table 2 Degradation of SNR and PEAQ

Sampling rate (kHz)	8	16	
SNR degradation	0.55	0.74	
PEAQ degradation	0.097	0.24	
Sampling rate (kHz)	24	32	48
SNR degradation	0.43	9.0	10.0
PEAQ degradation	0.39	0.92	1.2

### 3 評価実験

各ビットレート及びサンプリング周波数で、提案法による劣化の度合いや情報量の削減率を調べた。音声・音楽・混合音など様々なジャンルの音響信号を入力データとした。ビットレートは 12 kbps, 16 kbps, 24 kbps, 32 kbps とした。各サンプリング周波数における平均 SNR と平均 Perceptual evaluation of audio quality (PEAQ) の劣化度を表 2 に示す。また、ビットレート毎の情報量削減率を図 3 に示す。

低ビットレート (12 kbps ~ 16 kbps) で入力信号のサンプリング周波数も 16 kHz 以上であれば、削減率がおおよそ 2% になることが確認された。以上のことから、我々の提案法により、音質の劣化を抑制したまま、線形予測次数を適応的に変化させることで情報量を削減可能であることが分かった。

### 4 おわりに

本研究では線形予測次数を適応的に変化させることで、ISF に必要な情報量を削減する方法を提案し、国際標準技術である AMR-WB+ に実装した。提案法は音質の悪化を小さく抑えながらビットレートの減少が実現可能であることを実験結果で証明された。今後、MUSHRA による

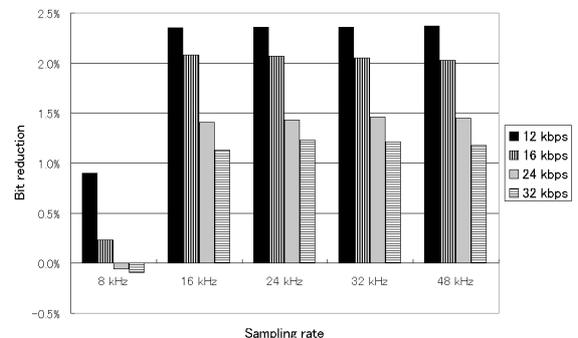


Fig. 3 Results of bit reduction.

主観品質評価を用いて、提案法の品質面での検証を行いたい。

### 参考文献

- [1] R. Salami *et al.*, "Extended AMR-WB for high-quality audio on mobile devices," *IEEE Comm. Magazine*, vol. 44, no. 5, pp. 90–97, May 2006.
- [2] 3GPP TS 26.290: *Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec*.
- [3] F. Zhang *et al.*, "Adaptive Prediction Order Scheme for AMR-WB+," *Proc. IEEE ISICIT 2010*, pp. 39–44, May 2006.