

音声信号の非線形予測と信号表現に関する研究

Study on Nonlinear Prediction of Speech Signals and Signal Representation

島村徹也 (大学院理工学研究科・教授)

Tetsuya Shimamura (Professor,
Graduate School of Science and Engineering)

1 はじめに

音声の線形予測技術は、音声信号の効率的な表現方法を与える。これまで、音声分析、音声認識、話者認識など、様々な応用に用いられてきた。しかし、応用物理学の分野で、音声には線形成分のみでなく非線形成分も多分に含まれることが知られている。そこで本研究では、音声信号を表現するために、線形の処理のみを施す枠組みを取り払い、非線形処理を施すことを試みる。

2 線形処理と非線形処理の結合

音声の信号表現に、過去のサンプル値から現在または未来のサンプル値を予測する予測フィルタの利用が適している仮定の下に、予測器の係数決定に線形処理と非線形処理とを組み合わせる方法について検討した。具体的には、図1に示されるような線形予測器と非線形予測器の縦続型構成を考えた。

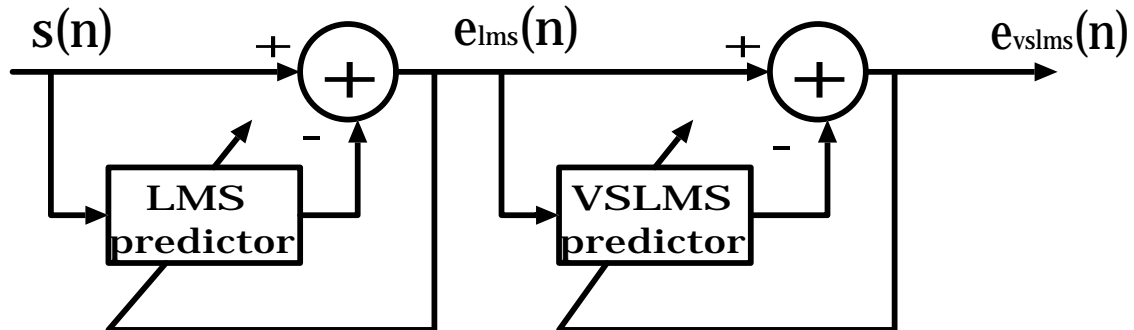


図1 縦続型予測器の構成

実行の容易さ、数値計算に関する安定性などを考慮し、ここでは線形予測器として平均最小自乗 (LMS) 予測器を、非線形予測器としてボルテラ級数展開LMS (VSLMS) 予測器を利用した。図1では、線形予測器が前置フィルタ、非線形予測器が後置フィルタの構成となっているが、フィルタ構成の最適化を考慮し、線形予測器と非線形予測器の配置を逆にする構成をも検討した。

3 実験結果

LMS予測器の次数は13、VSLMS予測器の次数は線形部10、2次の非線形部3とした。またステップサイズパラメータは0.05で共通とした。約10秒の短文に対し、共通して正規化型のLMS

適応アルゴリズムを採用し、LMS予測器、VSLMS予測器、それらの縦続構成、をそれぞれ施したときに得られる信号雑音比(SNR)を算出した。ここでのSNRは、音声信号パワーの予測誤差信号パワーに対する比として定義した。結果が男女別に図2, 3に示される。男性2名、女性2名の平均の結果がそれぞれ示される。

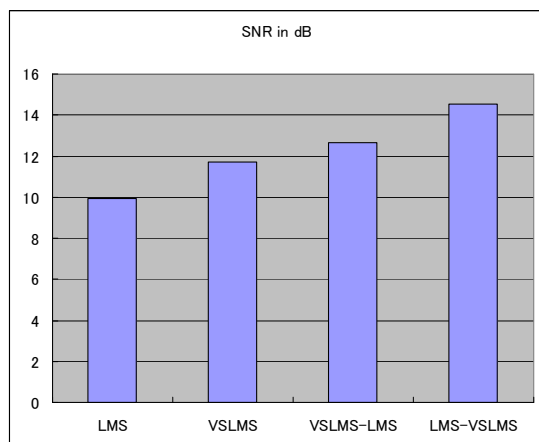


図2 男性の場合の実験結果

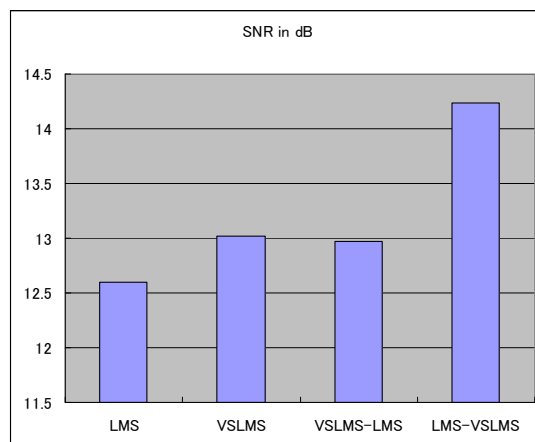


図3 女性の場合の実験結果

図2, 3より、LMS予測器およびVSLMS予測器の単独利用に比べ、縦続型構成はSNRを改善することがわかる。特に、LMS予測器を前処理とする縦続型が効果を発揮する。これは、線形フィルタは高速に収束することが可能であるためで、ボルテラ級数展開フィルタの入力信号としての条件を良好にする作用をもたらすためと考えられる。男性の場合における改善度は4dB程度であり、女性の場合における改善度は2dB程度である。これらの数値は、音声符号化を考慮した場合には、十分知覚される数値であり、縦続型予測器の有効性が図2, 3の結果から検証できる。

4 まとめ

実験において、ときどきVSLMS予測器が適応過程で発散する現象が観察された。これはおそらく数値的不安定性によるものであり、ボルテラ級数が作り出す入力ベクトルに問題があると考えられる。今後は、より数値的に安定な入力ベクトルの導出に努めたい。

本研究に関連し、下記の研究発表を行った。ジャーナル論文1編、国際学会論文2編である。

関連業績

(1)Tanaka H., Ohhashi Y. and Shimamura T., Performance of Adaptive Nonlinear Predictor with Order Statistics in Impulsive Noise, WSEAS Transactions on Signal Processing, Issue 2, Volume 3, 2007. 2.

(2)Tanaka H., Ohhashi Y. and Shimamura T., Adaptive Non-Linear Prediction with Order Statistics for Speech Signals in Mixture Noise, Proceedings of IEEE International Symposium on Intelligent Signal Processing and Communication Systems, pp.295-298, 2006. 12.

(3)Tanaka H., Ohhashi Y. and Shimamura T., Adaptive Non-Linear Prediction with Order Statistics for Speech in Impulsive Noise, Proceedings on WSEAS International Conference on Circuits, Systems, Signal and Telecommunications, pp.125-129, 2007. 1.