

CHARACTERIZATION OF A CASCADE LMS PREDICTOR

D.Y. Huang, X.Su† and A. Nallanathan†

Multimedia Signal Processing Laboratory/Media Division
Institute for Infocomm Research
21, Heng Mui Keng Terrace
Singapore 119613

† Department of Electrical and Computer Engineering
National University of Singapore
Singapore 119260

ABSTRACT

In this paper, an FIR cascade structure for adaptive linear prediction is studied in which each stage FIR filter is independently adapted using LMS algorithm. The theoretical analysis shows that the cascade performs a linear prediction in a way of successive refinement and each stage tries to obliterate the dominant mode of its input. Experimental results show that the performance of the cascade LMS predictor are in good agreement with our theoretical analysis.

1. INTRODUCTION

It is well known that linear prediction techniques have been widely used in speech, audio and video coding [1, 2]. Recently, an FIR cascade structure has been proposed for adaptive linear prediction to speed up the convergence and to get much smaller mean square error (mse) than single stage LMS predictor [4].

Later, such cascade structures like a cascade of LMS or RLS-LMS predictors are successfully applied for lossless audio coding and showed that their predictive coding performances may outperform the conventional linear prediction coding technique [5, 6].

Theoretical analysis of this cascade is made only for a low-order example [4]. Computer simulations for synthetic signals or real speech and audio signals have, however, shown that the cascade predictor behaviors more effective particularly during the initial, transient phase of the adaptation and results in a smaller final MSE.

In this paper, we shall carry out a theoretical analysis for this phenomenon and demonstrate formally that each stage of the cascade structure attempts to cancel the dominant mode of its input signal. The paper is organized as follows; Section 2 will review the cascade structure. Section 3 describes the LMS algorithm and its performance in

terms of filter weight and MSE. Section 4 presents a mathematical analysis for the characterization of the cascade LMS predictor and its property. Section 5 contains the simulation results for synthetic and real audio signals to support our theoretical analysis. Finally, we give a conclusion and point out our future research.

2. THE CASCADE STRUCTURE

The structure of the cascade for the linear prediction problem can be shown in Fig.1. In the cascade, each stage of the M sections is uses an independently adapting FIR predictor of order $l_k, k = 1, \dots, M$. Let $x_k(n)$ and $e_k(n)$ be the input to stage k , and the corresponding prediction error, respectively; it is

$$e_k(n) = x_k(n) - \sum_{m=1}^{l_k} h_k^{(m)}(n)x_k(n-m) \quad (1)$$

where $h_k^{(1)}(n), \dots, h_k^{(l_k)}(n)$ are the time-varying taps of the k th predictor. Each stage of the cascade structure satisfies that $x_{k+1}(n) = e_k(n); x_1(n) = x(n)$, where $x(n)$ is the signal to model. The error of the last stage, $e_M(n)$ is the global prediction error of the structure. After convergence, $h_k^{(m)}(n) = h_k^{(m)}$ and the global predictor transfer function can be expressed as

$$\hat{H}(z) = \prod_{k=1}^M H_k(z) \quad (2)$$

where

$$H_k(z) = 1 - \sum_{m=1}^{l_k} h_k^{(m)} z^{-m} \quad (3)$$

In fact, this structure is inadequate for general input signals since the resulting prediction filters has only strictly real zeros. However, there is a report that the speed of convergence

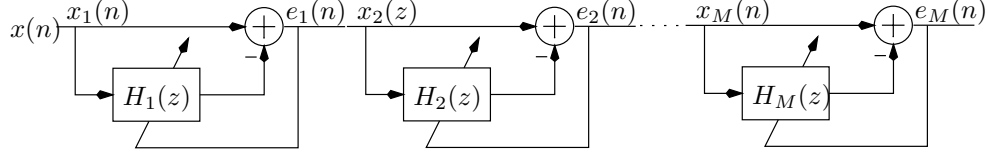


Fig. 1. Cascade predictor.

of the cascade filter with LMS adaptation is such that its initial MSE is usually smaller than those of equivalent-order LMS and lattice LMS predictors [5]. For each stage, the cost function is defined as

$$J_k = E[e_k^2(n)] = E[x_{k+1}^2(n)] \quad (4)$$

3. THE LMS PREDICTOR

Each stage performs the prediction by passing the past values through an l_k -tap FIR filter, where the filter weights are updated through the LMS weight update equation

$$\mathbf{h}_k(n+1) = \mathbf{h}_k(n) + \mu e_k(n) \mathbf{X}_{l_k}(n-1) \quad (5)$$

where T denotes the transpose operator, $\mathbf{X}_{l_k}(n) = [x(n-1), x(n-2), \dots, x(n-l_k)]^T$, and $\mathbf{h}_k(n) = [h_k^{(1)}(n), h_k^{(2)}(n), \dots, h_k^{(l_k)}(n)]^T$.

The weight update equation is derived through a minimization of the mean-square error (MSE) between the desired signal and the LMS estimate, namely,

$$E[e_k^2(n)] = E[(x_k(n) - \hat{x}_k(n))^2] \quad (6)$$

For the simplicity, the performance of the LMS predictor can be analyzed with "independence assumption" [3] and can be bounded by that of the finite Wiener filter, where the filter weights are given in terms of the autocorrelation matrix of the reference signal \mathbf{R}_k , and the cross-correlation vector between the past value and desired signals \mathbf{r} . Explicitly, the weights are

$$\mathbf{h}_k(n) = \mathbf{R}_k^{-1} \mathbf{r}_k \quad (7)$$

where $\mathbf{R}_k = E[\mathbf{X}_{l_k}(n) \mathbf{X}_{l_k}^H(n)]$ and $\mathbf{r}_k = E[\mathbf{X}_{l_k}(n) x_{l_k}(n)]$.

The MSE of the LMS predictor (5) under these assumptions is therefore bounded by the MSE of the finite Wiener filter, which is

$$E[e_k^2(n)] = E[(x_k(n) - \hat{x}_k(n))^2] = E[x_k^2(n)] - (\mathbf{R}_k^{-1} \mathbf{r}_k)^T \mathbf{r}_k \quad (8)$$

Referring Eq.(4), we are able to write this in spectral density function as

$$J_{k_{opt}} = \int_{-\pi}^{\pi} S_{x_k x_k}(\lambda) d\lambda - \int_{-\pi}^{\pi} |H_k(\lambda)|^2 S_{x_k x_k}(\lambda) d\lambda \quad (9)$$

4. CHARACTERIZATION OF THE CASCADE LMS PREDICTOR

In this section, we try to prove that the cascaded adaptive FIR filters operates a linear prediction in terms of successive refinements. The cascaded adaptive FIR operation can be described in the following the theorem:

Theorem 1 *In the cascaded FIR filters, each stage attempts to cancel the dominant mode of its input signal, i.e., to place its zeros close to the dominant poles of the AR model. It performs a linear prediction with a successful progressive refinement strategy, i.e.,*

$$J_M(h_M) \leq J_{M-1}(h_{M-1}) \leq \dots \leq J_1(h_1) \quad (10)$$

Proof: Assume N the minimum description length (MDL) of the AR model, the time series $x(n), x(n-1), \dots, x(n-N)$ can be realized by an autoregressive (AR) of order N as it satisfies the difference equation

$$x(n) + a_1^* x(n-1) + \dots + a_N^* x(n-N) = v(n) \quad (11)$$

where a_1, \dots, a_N are constants called AR parameters and $v(n)$ is white noise. The corresponding system generates $x(n)$ with the white noise $v(n)$, whose transfer equation equals

$$H(z) = \frac{1}{\sum_{i=0}^N a_i^* z^{-i}} \quad (12)$$

This function is completely defined by specifying the location of its poles, as shown by

$$H(z) = \frac{1}{(1 - p_1 z^{-1})(1 - p_2 z^{-1}) \dots (1 - p_N z^{-1})} \quad (13)$$

The parameters p_1, p_2, \dots, p_N are poles of $H(z)$; they are defined by the roots of the characteristic equation

$$1 + a_1^* z^{-1} + \dots + a_N^* z^{-N} = 0 \quad (14)$$

For the system to be stable, the roots of the characteristic equation (13) must all lie inside the unit circle in the z -plane, e.g., $|p_k| < 1$, for all $k = 1, \dots, N$. The cascade FIR filter (3) of order N of each stage consists to estimate w_1^*, \dots, w_N^* in the linear prediction problem

$$x(n) = w_1^* x(n-1) + w_2^* x(n-2) + \dots + w_N^* x(n-N) + v(n) \quad (15)$$

such that $w_k = -a_k$. The analyzer function $H(z)$ can be expressed into cascade form

$$\begin{aligned} H(z) &= 1 - w_0^* z^{-1} - w_1^* z^{-2} - \dots - w_N^* z^{-N} \\ &= (1 - \sum_{m=1}^{l_1} h_1^{(m)} z^{-m})(1 - \sum_{m=1}^{l_2} h_2^{(m)} z^{-m}) \dots \\ &= \prod_{k=1}^M H_k(z) \end{aligned} \quad (16)$$

where $\sum_{k=1}^M l_k = N$. We have the output at first stage of the cascade FIR structure while the LMS predictor converges to its steady-state value (Eq.7),

$$\begin{aligned} e_1(n) &= \sum_{m=1}^M w_m^* x(n-m) + v(n) \\ &\quad - \sum_{m=1}^{l_1} h_1^{(m)}(n) x(n-m) \\ &= \underbrace{\sum_{m=1}^{l_1} (w_m^* - h_1^{(m)}) x(n-m)}_{e_0(n)} \\ &\quad + \sum_{m=l_1+1}^M w_m^* x(n-m) + v(n) \end{aligned} \quad (17)$$

The cost function at the first stage becomes

$$\begin{aligned} J_1(n) &= E[|e_1(n)|^2] \\ &= E[|e_0^2(n) + v^2(n) + 2e_0(n)v(n) \\ &\quad + 2 \sum_{m=l_1+1}^M w_m^* x(n-m)(e_0(n) + v(n)) \\ &\quad + (\sum_{m=l_1+1}^M w_m^* x(n-m))^2|] \end{aligned} \quad (18)$$

According to the principle of orthogonality, at the steady-state, $E[e_0(n)v(n)] = 0$ and $E[x(n-m)(e_0(n) + v(n))] = 0$. The cost function becomes

$$J_1(n) = E[|e_0(n)|^2 + \sigma_v^2(n) + |\sum_{m=l_1+1}^M w_m^* x(n-m)|^2] \quad (19)$$

where $\sigma_v^2(n)$ is the variance of the white noise $v(n)$. We see that $J_1(n)$ achieve its minimum, if and only if, the following two terms are minima

$$J_{1e0}(n) = E[|e_0(n)|^2] \quad (20)$$

and

$$J_{1w}(n) = E[|\sum_{m=l_1+1}^M w_m^* x(n-m)|^2] \quad (21)$$

It means that the first stage attempts to cancel the dominant mode of its input signal, i.e., to place its zeros close to the dominant poles of the AR model.

We look at the sufficient condition: if $J_{1e0}(n)$ and $J_{1w}(n)$ are minima, the dominant component of input signal is removed. In fact, $H_1(z)$ can be decomposed as

$$\hat{H}_1(z) = (1 - \hat{p}_1 z^{-1})(1 - \hat{p}_2 z^{-1}) \dots (1 - \hat{p}_{l_1} z^{-1}) \quad (22)$$

The zeros $|\hat{p}_k| < 1, k = 1, \dots, l_1$ are close to the poles $p_k, k = 1, \dots, l_1$ in Eq.(13), which dominate the component of the input. The rest poles $p_k, k = l_1, \dots, N$ have minimum component of the input, resulting in the minimum $J_{1w}(n)$.

For necessary condition: only if, we can assume that the zeros $|\hat{p}_k| < 1, k = 1, \dots, l_1$ are close to the poles $p_k, k = 1, \dots, l_1$ in Eq.(13), which are not the dominant component of the input. There are the poles in $p_k, k = l_1, \dots, M$, which gives the dominant component of the input. Therefore, there are a set of \hat{w}_k such that

$$|\sum_{m=l_1}^N \hat{w}_m^* x(n-m)|^2 > |\sum_{m=l_1}^N w_m^* x(n-m)|^2 \quad (23)$$

The $J_{1w}(n)$ is not minimum. This is contradictory to the initial assumption, i.e., the cost function $J_1(n)$ achieve its minimum, resulting in the minimum $J_{1w}(n)$. Thus the first stage will attempt to cancel the dominant mode of its input signal, i.e., to place its zeros close to the dominant poles of the AR model. The proof for second stage is done in same way, and so on.

Referring to Eq.(4) and Eq.(9), it is easy to verify,

$$\begin{aligned} J_k(h_k) &= E[e_{k-1}^2(n)] - \underbrace{\int_{-\pi}^{\pi} |H_k(\lambda)|^2 S_{x_k x_k}(\lambda) d\lambda}_{a^2} \\ &= J_{k-1}(h_{k-1}) - a^2 \leq J_{k-1}(h_{k-1}) \end{aligned} \quad (24)$$

Where $a^2 < E[e_{k-1}^2(n)]$. Therefore,

$$J_M(h_M) < J_{M-1}(h_{M-1}) < \dots < J_1(h_1) \quad (25)$$

the theorem is proved.

With above theorem, we can derive the following property of the cascade LMS predictor.

Lemma 1 If each stage of the cascade LMS predictor converges to its steady-state value, the cascaded FIR filters possess the following property:

$$\chi(R_M) < \chi(R_{M-1}) < \dots < \chi(R_1) \quad (26)$$

where

$$\chi(R_k) = \frac{\lambda_{kmax}}{\lambda_{kmin}} \quad (27)$$

Proof: The condition given in lemma 1 means that the optimum cascaded FIR filters satisfy theorem 1. The output of the first stage $e_1(n)$ can be used only $(N - l_1)$ -by-1 tap-input vector $\mathbf{x}(n)$ to characterize. In other word, after first stage adaptation and convergence, the input signal dynamic range to the second stage is reduced. The ratio between the peak and average of the spectral density of the input signal is decreased. Thus

$$\chi(R_2) < \chi(R_1) \quad (28)$$

For the output of the second stage $e_2(n)$, in the same reason, it can be estimated using only $(N - l_1 - l_2)$ -by-1 tap-input vector $\mathbf{x}(n)$. Therefore, the eigenvalue spread of the input to the third stage satisfies

$$\chi(R_3) < \chi(R_2) \quad (29)$$

and so on until the last stage, the input can be estimated using l_k -by-1 tap-input vector and satisfies

$$\chi(R_M) < \chi(R_{M-1}) \quad (30)$$

The lemma is proved.

5. SIMULATION RESULTS

In above demonstration of the theorem and lemma, we assume that each stage of the cascade LMS predictor converges to its steady-state value. However, LMS convergence speed suffers from both the length of the filter and the eigenvalue spread of the input covariance matrix. In practice, the first stage can use a low-order filter as a pre-whitening adaptive filter to reduce the eigenvalue spread. The second stage adopts a long LMS predictor, which works well for general signals and it is different from the cascade short-size filters [4]. We carried out the simulations for synthetic and real signals to evaluation the behavior of the cascade LMS predictor. Here we show the results of three stages cascade LMS predictor for an audio linear prediction to support our theoretical analysis. From Fig.1 (a) (b) and (c), we observe that the cascade LMS predictor removes the dominant component of input in recursive way.

We applied the cascade structure to audio clips sampling at different rates for 51 audio clips provided by MPEG-4 with different lossless audio codecs' predictor (Monkey 3.97, cascade LMS and TUB LPC predictor). The average results of SNR for different predictor are shown in the Table 1. It shows that the proposed predictor in such cascade structure gives best predictive gain among them.

6. CONCLUSION

In this paper, we gave a formal proof that the cascade LMS predictor performs a linear prediction in terms of successive

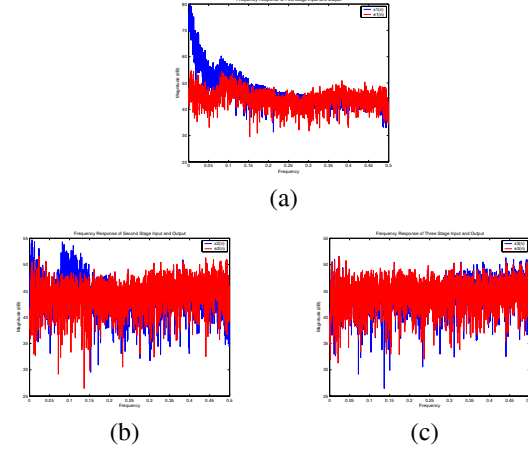


Fig. 2. Frequency Response of three stages cascade LMS predictor: (a) First stage $x_1(n)$ and $e_1(n)$; (b) Second stage $x_2(n)$ and $e_2(n)$; (c) Third stage $x_3(n)$ and $e_3(n)$

Table 1. The results of SNR for different lossless codecs' predictor for mono signals

Track	Monkey 3.97	cascade LMS	TUB (LPC)
48kHz/16bit	28.8(dB)	29.5(dB)	28.1 (dB)
48kHz/24bit	28.9(dB)	30.0(dB)	28.6(dB)
96kHz/24bit	51.5(dB)	53.9(dB)	53.5(dB)
192kHz/24bit	63.4(dB)	65.4(dB)	65.3(dB)

refinement. The simulation results for synthetic and real audio signals confirm our theoretical analysis and the cascade adaptive linear predictor may lead to better predictive gain than LPC technique. We shall study the performance of this structure in near future.

7. REFERENCES

- [1] A. Berg and W. Mikhael. A Survey of Techniques for Lossless Compression of Signals. *Proc. 37th Midwest Symposium on Circuits and Systems*, 1995, pp. 943 - 946.
- [2] ISO/IEC JTC1/SC29/WG11 M9781 T. Liebchen. Detailed technical description of the TUB 'lossless only' proposal for MPEG-4 Audio Lossless Coding.
- [3] S. Haykin. *Adaptive Filter Theory*. 3rd edition, Prentice-Hall, 1996.
- [4] P. Prandoni and M. Vetterli. An FIR cascade structure for adaptive linear prediction. *IEEE Trans. on Signal Processing*, vol. 46, no.9, pp. 2566-2571, Sep. 1998.
- [5] G. Schuller *et al.*. Lossless Coding of Audio Signals using Cascade Prediction. *Proc. ICASSP 2001*, pp. 3273 - 3277, 2001.
- [6] R. Yu, C.C. Ko, S. Rahardja, and X. Lin. An RLS-LMS Algorithm for Lossless Audio Coding. *Asimolar Conference on Signals, Systems and Computers 2003*.