

A Lattice Predictor Based Adaptive Volterra Filter and Its Convergence Property Analysis

Kenji Nakayama Akihiro Hirano Hiroaki Kashimoto
 Division of Electrical Engineering and Computer Science
 Graduate School of Natural Science and Technology, Kanazawa Univ.
 2-40-20, Kodatsuno, Kanazawa, 920-8667, JAPAN
 e-mail:nakayama@t.kanazawa-u.ac.jp

Abstract— This paper proposes a lattice predictor based adaptive Volterra filter (Lattice-AVF), and its convergence property is analyzed. In the adaptive FIR Volterra filter (AVF), the eigen value spread of a correlation matrix is extremely amplified, and its convergence is very slow for gradient methods. A lattice predictor is employed for whitening the input signal. For stationary colored input signals, the Lattice-AVF can provide a fast convergence and the well reduced residual error. Its convergence is highly dependent on a time constant, used in updating the reflection coefficients. A very large time constant is required. In the case of nonstationary colored input signal, the eigen value spread after the Volterra polynomial is not so highly amplified. This means fast convergence will be expected, and effects of the whitening will be small. These properties are analyzed. A problem of asynchronous updating the reflection coefficients and the filter coefficients observed in linear lattice predictor based adaptive filters, is also observed in the Lattice-AVF.

I. INTRODUCTION

Loud speakers in audio systems and small speakers embedded in a mobile phone have some nonlinearity. When they are used in a remote conference system and as a visual phone, where some echo are caused and are transmitted through such nonlinear parts, nonlinear echo cancellers are very important [1],[2].

An adaptive FIR Volterra filter (AVF), which combines an FIR filter and a Volterra polynomial, which can express general nonlinearity, is one of hopeful candidates [3],[4]. However, the Volterra polynomial generates a huge number of terms, and the same number of filter coefficients are required. Furthermore, eigen value spread of a correlation matrix for the input signal is extremely amplified after the Volterra polynomial, and convergence becomes very slow for gradient methods.

Convergence property of AVFs has been analyzed, and some fast and stable learning algorithms have been proposed [5],[6]. One approach to fast convergence is orthogonalization of the input signal [7],[8],[9],[10]. However, the DCT is not sufficient for orthogonalization. A linear prediction error filter is good for orthogonalization. However, it requires some time delay, and cannot be applied to some applications, such as echo cancellers [11].

In this paper, in order to improve the orthogonalization process without any time delay, a lattice prediction error filter is employed for this purpose. In linear lattice predictor based adaptive filters [12], a problem of asynchronous

update of the reflection coefficients and the filter coefficients is pointed out, and the residual error is not well reduced. A synchronization method has been proposed [13],[14]. In this paper, this asynchronous problem is analyzed. Convergence properties for stationary and nonstationary colored signals are analyzed. Several kinds of structures for orthogonalization of the input signals are taken into account. Convergence property is compared with the DCT based and the linear predictor based adaptive filters.

II. ADAPTIVE FIR VOLTERRA FILTER

A. Structure of AVF

Figure 1 shows a block diagram of an adaptive FIR Volterra filter (AVF). When a second-order Volterra poly-

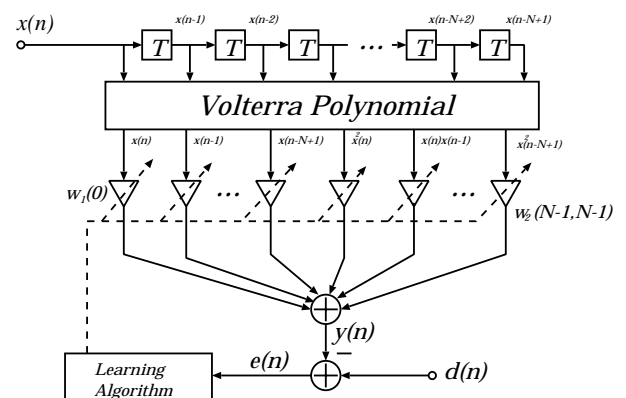


Fig. 1. Adaptive FIR Volterra filter.

nomial is used, the output $y(n)$ is given by

$$y(n) = \sum_{i=0}^{N-1} w_1(i)x(n-i) + \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} w_2(j,k)x(n-j)x(n-k) \quad (1)$$

$$w_2(j,k) = w_2(k,j) \quad (2)$$

B. Eigen Value Spread Amplification by Volterra Polynomial

The eigen value spread $\chi = \lambda_{max}/\lambda_{min}$ of a correlation matrix for the input signal $x(n)$ is extremely amplified by transforming the signals through the Volterra polynomial. Examples are shown here. A 50 tap FIR filter and a 2nd-order Volterra polynomial are used. The number of terms in the Volterra polynomial is 1325. Samples of the colored input signal $x(n-i)$ with $\chi = 780.9$ is transferred into

the Volterra polynomial terms $x(n-i)x(n-j)$ having $\chi = 657100$, which is 841 times as large as that of $x(n)$. For this reason, convergence of the AVF is very slow for gradient methods.

C. Orthogonalization of Input Signal

1) *Discrete Cosine Transform*: The DCT with normalization [10] is used for comparison. The outputs of the tap delay line $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-N+1)]$ are transformed through the DCT to $\mathbf{q}(n) = [q_0(n), q_1(n), \dots, q_{N-1}(n)]$, and they are normalized by its standard deviation $\sigma_{q,i}$, resulting in the outputs $s_i(n)$. $s_i(n)$ are applied to the Volterra polynomial generating 1st-order and high-order terms. These terms are multiplied by filter coefficients, and are accumulated, resulting in the final output $y(n)$. The DCT does not need any time delay. However, orthogonalization is not complete. The adaptive Volterra filter with DCT is denoted 'DCT-AVF' in this paper.

2) *Linear FIR Prediction Error Filter*: When the signal can be modeled as the output of an AR circuit driven by the white noise, a linear FIR prediction error filter is good for whitening. However, it requires some time delay, and its application is rather limited. The prediction error is used as the AVF input. The adaptive Volterra filter with the linear prediction error filter is denoted 'LP-AVF' in this paper.

D. Position of Orthogonalization in Nonlinear Filters

Two kinds of positions for the orthogonalization are taken into account. They are at the input of only the AVF, denoted Type-A, and at the input of both the unknown system and the AVF, denoted 'Type-B'. In the case of nonlinear adaptive filters, any time delay is not allowed in Type-A. On the other hand, in Type B, any kinds of the orthogonalization, with or without time delay, can be employed. However, since the output of the unknown system may be sound from a loud speaker located in a conference room, this type cannot be applied to some applications, such as echo cancellation.

III. LATTICE PREDICTOR BASED AVF

A. Circuit Structure

In practical applications, Type A is important. Therefore, we employ the lattice prediction error filter [12] for the orthogonalization process. The proposed lattice predictor based AVF (Lattice-AVF) is shown in Fig.2. Let the number of taps of the FIR part be N , the order of the Volterra polynomial be M , the order of the lattice predictor be L . Assuming $N > L$, the number of filter coefficients in both AVFs in Fig.1 and Fig.2 are the same. If the unknown system can be modeled by using the FIR Volterra filter, the same transfer function can be realized by using the Lattice-AVF.

B. Reflection Coefficient Update

The reflection coefficients are updated by the following equations [12].

$$\kappa_m(n) = -\frac{2E[b_{m-1}(n-1)f_{m-1}^*(n)]}{E[|f_{m-1}(n)|^2 + |b_{m-1}(n-1)|^2]} \quad (3)$$

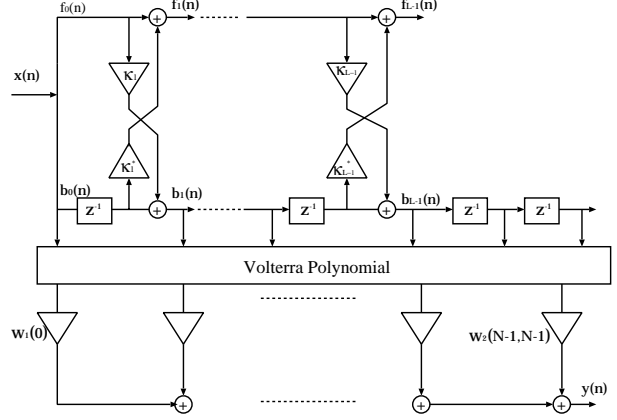


Fig. 2. Lattice predictor based AVF.

$$\kappa_{N,m}(n) = \gamma \kappa_{N,m}(n-1) + b_{m-1}(n-1) f_{m-1}^*(n) \quad (4)$$

$$\kappa_{D,m}(n) = \gamma \kappa_{D,m}(n-1) + |f_{m-1}(n)|^2 + |b_{m-1}(n-1)|^2 \quad (5)$$

$$0 < \gamma < 1$$

$$\kappa_m(n) = -2 \frac{\kappa_{N,m}(n)}{\kappa_{D,m}(n)} \quad (6)$$

C. Synchronization of Updating Reflection and Filter Coefficients

Convergence property of the linear lattice predictor based FIR adaptive filter has been analyzed. Updating the reflection coefficients and the filter coefficients are not synchronized, and some error remain. The synchronized learning algorithm has been proposed [13],[14], which is briefly described here.

The linear adaptive filter with the lattice predictor is equivalent to the circuit shown in Fig.2, except for the Volterra polynomial block. The filter coefficients $\mathbf{w}(n)$ is directly connected to $\mathbf{b}(n)$. The output $y(n)$ is

$$\mathbf{b}(n) = \mathbf{K}(n)\mathbf{x}(n) \quad (7)$$

$$y(n) = \mathbf{w}^T(n)\mathbf{b}(n) \quad (8)$$

$\mathbf{b}(n)$ is a vector of the backward prediction error $b_m(n)$, $\mathbf{K}(n)$ is a matrix consists of the reflection coefficients, $\mathbf{x}(n)$ is the input, $\mathbf{w}(n)$ is the filter coefficients. In the next iteration step, $\mathbf{K}(n)$ is updated to $\mathbf{K}(n+1)$, and $y(n+1)$ and $e(n+1)$ are generated by using $\mathbf{K}(n+1)$ and $\mathbf{w}(n)$. However, $\mathbf{w}(n)$ is optimized for $\mathbf{K}(n)$ not $\mathbf{K}(n+1)$. Therefore, $e(n+1)$ is not guaranteed to be reduced. For this reason, $\mathbf{w}(n)$ is modified sa as,

$$\tilde{\mathbf{b}}(n+1) = \mathbf{K}(n)\mathbf{x}(n+1) \quad (9)$$

$$\tilde{y}(n+1) = \mathbf{w}^T(n)\tilde{\mathbf{b}}(n+1) \quad (10)$$

$$\mathbf{b}(n+1) = \mathbf{K}(n+1)\mathbf{x}(n+1) \quad (11)$$

$$y(n+1) = \mathbf{w}^T(n)\mathbf{b}(n+1) \quad (12)$$

$\tilde{y}(n+1)$ can reduce the output error. Therefore, the filter coefficients $\mathbf{w}(n)$ is modified as follows:

$$\mathbf{K}^T(n+1)\tilde{\mathbf{w}}(n) = \mathbf{K}^T(n)\mathbf{w}(n) \quad (13)$$

$$\tilde{\mathbf{w}}(n) = \frac{\mathbf{K}^T(n)}{\mathbf{K}^T(n+1)}\mathbf{w}(n) \quad (14)$$

$\tilde{\mathbf{w}}(n)$ is used in the next iteration $n+1$, instead of $\mathbf{w}(n)$, for generating $\tilde{y}(n+1)$ and $\tilde{e}(n+1)$. The filter coefficients

are updated to $\mathbf{w}(n+1)$ by using $\tilde{e}(n+1)$ and $\tilde{\mathbf{w}}(n)$. A combination of $\mathbf{K}(n)$ and $\mathbf{w}(n)$ is equivalent to that of $\mathbf{K}(n+1)$ and $\tilde{\mathbf{w}}(n)$. Thus, the output error is guaranteed to be decreased.

The synchronization problem in the Lattice-AVFs is analyzed in the next section.

IV. SIMULATION AND DISCUSSIONS

A. Stationary Colored Signal

The colored signal is generated by using a 2nd-order AR model with the white input signal. A 50 tap FIR filter and the 2nd-order Volterra polynomial are used. The number of the terms in the Volterra polynomial is 1325.

1) *DCT and Linear Predictor*: The learning curves for the AVF without whitening, and with the DCT process in Type A and the linear prediction error filter in Type B are shown in Fig.3. The NLMS algorithm and stepsize=0.1 are employed.

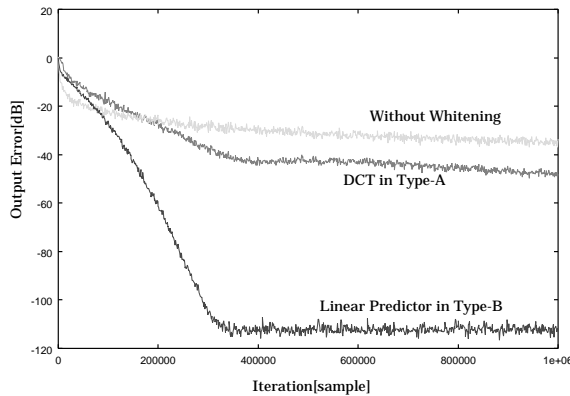


Fig. 3. Learning curves for colored signal, without whitening, and with DCT and linear prediction.

The DCT-AVF in Type-A can improve convergence to some extent. The LP-AVF in Type-B is good, because the whitening is complete.

2) *Lattice-AVF with Fixed and Time Variant Reflection Coefficients*: Time invariant and time variant reflection coefficients are taken into account. Also, they are deviated from the ideal value following

$$\kappa(n) = \kappa_0(1 + a) \quad \text{Time Invariant} \quad (15)$$

$$\kappa(n) = \kappa_0(1 + a \sin(2\pi n/1000)) \quad \text{Time Variant} \quad (16)$$

κ_0 is the ideal reflection coefficients.

Figure 4 shows the learning curves of the Lattice-AVF. If the reflection coefficients are set to the ideal, convergence is fast and the residual error is very small. However, the convergence is sensitive to deviation of the reflection coefficients. Assuming the output error of -50dB, 5% deviation causes a long convergence time, which is 3 times as long as the ideal case.

On the other hand, when the reflection coefficients are time variant, convergence is trapped at -20dB, which is not well reduced compared with the time invariant reflection coefficients. This phenomenon is caused by asynchronous update of the reflection coefficients and the filter coefficients, just the same as the linear lattice adaptive filters [13],[14].

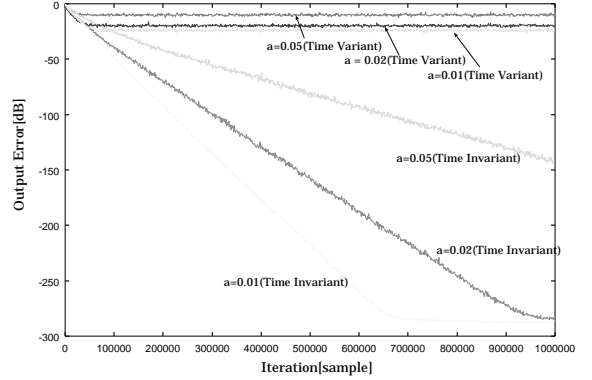


Fig. 4. Learning curve for stationary colored signal by using Lattice-AVF. Both time invariant and time variant reflection coefficients are used.

3) Lattice-AVF with Updated Reflection Coefficients

Figure 5 shows the learning curves, in which the reflection coefficients are updated following Eqs.(3) through (6). Convergence highly depends on a time constant γ , which control the reflection coefficient update. From Eqs.(4) and (5), $\kappa_{N,m}(n)$ and $\kappa_{D,m}(n)$ can be regarded as the output of a 1st-order lowpass filter, with a pole γ on the positive real axis. Therefore, when $\gamma (< 1)$ is very close to unity, the pasband of the lowpass filter is very narrow, and the output, that is the reflection coefficients are very slowly changed. Until $\gamma = 0.999999$, the convergence can be improved. The learning curves for the lattice predictor almost saturate around -50dB. However, it requires only 85,000 iterations until -40dB, while the DCT-AVF and the LP-AVF need 320,000 iterations and 150,000 iterations as shown in Fig.3, respectively.

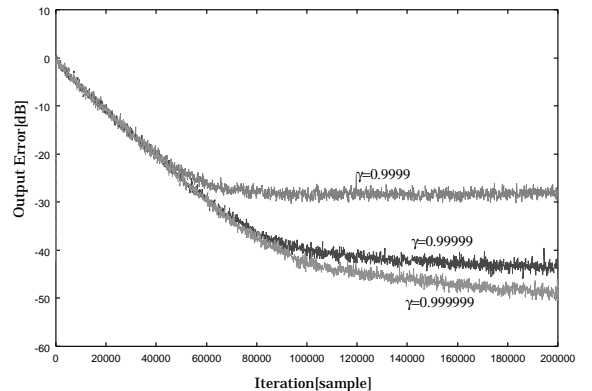


Fig. 5. Learning curve for stationary colored signal by using Lattice-AVF.

B. Nonstationary Colored Signals

A time variant AR model is used for generating nonstationary colored signals. The pole, $re^{\pm\theta}$, of the AR model is controlled by

$$r = 0.9 \quad (17)$$

$$\theta(n) = \frac{\pi}{4} \left(1 + a \sin\left(\frac{2\pi n}{500}\right) \right) \quad (18)$$

1) *Lattice-AVF*: Figure 6 shows the learning curves for the Lattice-AVF with several γ . In this case, $\gamma = 0.999999$ is also useful.

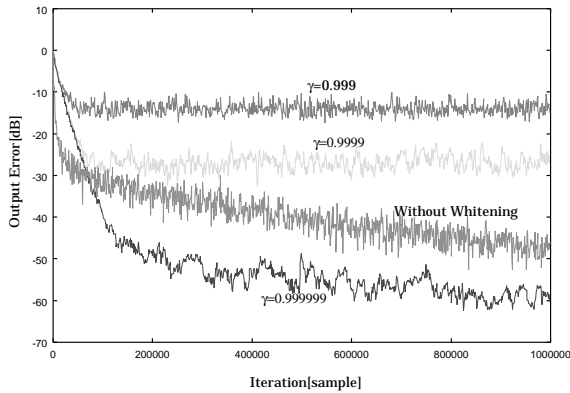


Fig. 6. Learning curves for nonstationary colored signal by using lattice predictor with several γ .

Figure 7 shows the learning curves for the Lattice-AVF, where a is changed as 0.05, 0.1 and 0.2. γ is set to 0.999999. The learning curves are almost the same as in Fig.5, where the AR model is fixed. From these results, the Lattice-AVF is useful for nonstationary input signals.

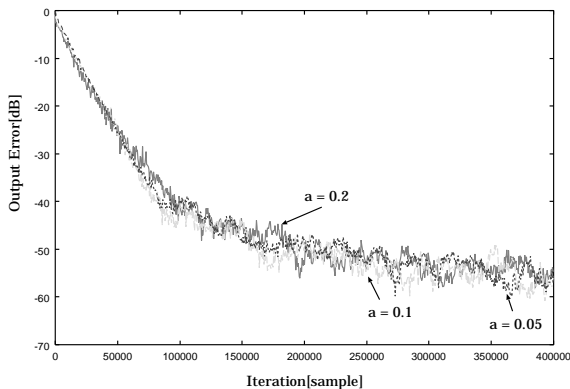


Fig. 7. Learning curve for nonstationary colored signal by using Lattice-AVF, where a is changed.

2) *Comparison among Lattice-, DCT- and LP-AVFs:* Figure8 shows learning curves by using the Lattice-AVF, the DCT-AVF and the LP-AVF. In this figure, 'Stationary' and 'Nonstationary' mean 'Without orthogonalization'. A typical feature is recognized. Convergence is improved for nonstationary colored signal, even though the orthogonalization is not used. In the case of nonstationary colored signal, the eigen value spread of the correlation matrix of $x(n-i)x(n-j)$ is not so amplified from that of $x(n-i)$ compared with stationary colored signal.

In the case of nonstationary colored signal, the Lattice-AVF still provide fast convergence. Convergence of the LC-AVF is slow, while it can reach the more reduced error. However, the LC-AVF is limited to Type-B.

V. CONCLUSIONS

In this paper, the Lattice-AVF has been proposed. Its convergence is highly dependent on the time constant used in updating the reflection coefficients. The asynchronous update of the reflection coefficients and the filter coefficients has been observed. Convergence of AVF for nonstationary colored signal is better compared with stationary

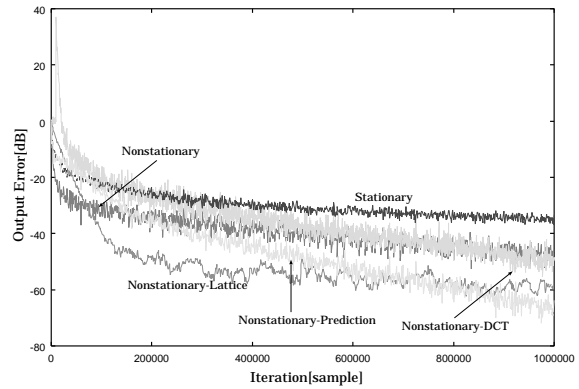


Fig. 8. Learning curves for nonstationary colored signal by using methods, including without whitening, DCT, linear predictor, and lattice predictor.

colored signal. The Lattice-AVF can provide fast convergence and well reduced error.

REFERENCES

- [1] B.S.Nollett and D.L.Jones, "Nonlinear echo cancellation for hands-free speakerphones," *1997 IEEE Workshop on Nonlinear Signal and Image Proc*
- [2] A.Stenger, L.Trautmann and R.Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," *IEEE Proc. ICASSP'99*, AE1.7, pp877-880, 1999.
- [3] M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, New York: Wiley, 1980.
- [4] V. J. Mathews, "Adaptive polynomial filters," *IEEE Signal Processing Mag.*, pp. 10-26, July 1991.
- [5] J.Chao, A.Inomata, T.Kubota, and S.Uno, "Error surface analysis and fast RLS adaptive algorithms of quadratic Volterra ADF," *2001 IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, June 3-6, 2001.
- [6] S.Kinoshita, Y.Kajikawa and Y.Nomura, "Volterra filters using multirate signal processing and their application to loudspeaker," *IEEE Proc. ICASSP'01*, pp.871-874, 2001.
- [7] T.Koh and E.J.Powers, "An adaptive nonlinear digital filter with lattice orthogonalization", *Proc. IEEE, ICASSP'83*, Boston, 1.10, April 1983.
- [8] V.J.Mathews, "Orthogonalization of correlated Gaussian signals for Volterra system identification," *IEEE Signal Processing Letters*, Vol.2, No.10, pp.188-190, Oct. 1995.
- [9] V.J.Mathews, "Adaptive Volterra filters using orthogonalization structures," *IEEE Signal Processing Letters*, Vol.3, No.12, pp.307-309, Dec. 1996.
- [10] X.Li and W.K.Jenkins, "Computationally efficient algorithms for third order adaptive Volterra filters," *IEEE Proc. ICASSP'98*, DSP5.3, 1998.
- [11] A.Stenger and W.Kellermann, "Nonlinear acoustic echo cancellation with fast converging memoryless preprocessor," *Proc. IEEE Workshop on Acoustic Echo and Noise Control*, Pocono Manor, PA, USA, September, 1999.
- [12] S.Haykin, *Adaptive Filter Theory*, 3rd Ed., Prentice Hall Inc., New York, 1996.
- [13] N.Tokui, K.Nakayama and A.Hirano, "A synchronized learning algorithm for reflection coefficients and tap weights in a joint lattice predictor and transversal filter", *IEEE Proc. ICASSP'2001*, Salt Lake City, pp.1472-1475, May 2001.
- [14] N.Tokui, K.Nakayama and A.Hirano, "Block implementation of a synchronized learning algorithm in adaptive lattice filters", *IEEE Proc., ICASSP'03*, Hong Kong, pp.VI-349-VI-352, April 2003.