Efficient Adaptive Nonlinear Echo Cancellation Using Sub-band Implementation of the Adaptive Volterra Filter

Dayong Zhou, Victor DeBrunner, Yan Zhai and Mark Yeary
School of Electrical and Computer Engineering
University of Oklahoma, Norman, OK, 73019
{dayong, vdebrunn, yan.zhai-1, yeary}@ou.edu

Abstract

The adaptive Volterra filter has been successfully applied in nonlinear acoustic echo cancellation (AEC) systems and nonlinear line echo cancellation systems, but its applications are limited by its required computational complexity and slow convergence rate, especially for systems with long memory length. In this paper, by leveraging a multi-channel configuration of the Volterra filter and the sampling theory for nonlinear systems, we extend linear sub-band delay-less adaptive filter techniques to develop an efficient sub-band implementation of the adaptive Volterra filter. The developed sub-band configuration of the adaptive Volterra filter can greatly improve the convergence rate and reduce the computational complexity of nonlinear echo cancellers, which is shown by analyses and simulations.

1. Introduction

Acoustic echoes arise from the acoustic coupling between the receive- and transmit-paths of a telecommunication terminal which greatly affects the quality of voice communication in wireless communication systems, VoIP services, etc. Acoustic echo cancellation (AEC) is an effective technique to suppress the echo effect and improve the communication system performance. The configuration of an AEC system is shown in Figure 1. Most AEC techniques do not consider the nonlinear distortion caused by loudspeakers, amplifiers and nonlinear effects in the vibration of the enclosure. However, recently, it has been found that the AEC system performance could be greatly improved by considering the nonlinearities existing in the system. Consequently, nonlinear acoustic echo cancellation (AEC) techniques, e.g. the algorithms in [1]-[3], have been developed. The nonlinear distortion has also been studied in the line echo cancellation configuration [4].

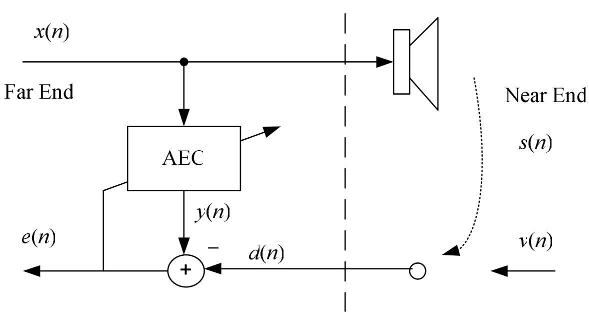

Figure 1. Configuration of acoustic echo cancellation

The adaptive Volterra filter has been applied in nonlinear acoustic echo cancellation (AEC) and nonlinear line echo cancellation systems to identify and track the time-varying nonlinear impulse response (NIR) from the far-end signal to the echo signal, which usually includes the A/D converter, nonlinear loudspeaker, room transfer function and other components [4]. Research results have proven the effectiveness of nonlinear echo cancellation using the adaptive Volterra filter [1]-[4]. However, it is well known that the adaptive Volterra filter suffers from a large computational complexity and is slow to converge. These are significant problems in nonlinear AEC systems due to the long memory length of the NIR and the wide speech spectral dynamics [5]. Many researchers working in this area have proposed different methods and simplified structures to speed convergence and reduce the computational burden, e.g. the affine projection based adaptive Volterra filter [1], cascaded structures implementation [2], and the MMD (multi-memory decomposition) structure [3].

The delay-less sub-band implementations of adaptive linear filters are proposed in [7], [8]. Without introducing the signal path delay, the proposed methods can greatly increase the convergence speed and reduce the computational complexity of large-order linear AEC systems and active noise control (ANC) systems. However, due to the complexity and spectral outgrowth caused by the nonlinearities, no sub-band version of the adaptive nonlinear filters is typically available.

In this paper, we use the sampling theory of nonlinear systems and the multi-channel structure of the Volterra filter to propose a new efficient delayless sub-band adaptive Volterra filter. Our configuration is based on a combination of Morgan's configuration [7] and DeBrunner's configuration [8], which is suitable for adaptive Volterra filter implementation. Like its linear counterparts, our proposed sub-band implementation of the adaptive Volterra filter demonstrates fast convergence and computational efficiency for a large-order system, which has proved to be especially useful in nonlinear AEC systems when the NIR has a large memory and speech is the main reference signal.

2. Preliminary Background

In nonlinear echo cancellation, the nonlinear system can be modeled and approximated by the truncated Volterra series model [1]-[4], [9], i.e. the relationship between the echo signal $y(n)$ and the far-end signal $x(n)$ can be expressed by the N-sample memory p^{th}-order truncated Volterra series expansion:

$$\begin{aligned} y(n) = h_0 &+ \sum_{k=0}^{N} h_1(k)x(n-k) \\ &+ \sum_{k_1=0}^{N}\sum_{k_2=k_1}^{N} h_2(k_1,k_2)x(n-k_1)x(n-k_2) \\ &+ \sum_{k_1=0}^{N}\cdots\sum_{k_p=k_{p-1}}^{N} h_p(k_1,\cdots k_p)x(n-k_1)\cdots x(n-k_p) \end{aligned} \tag{1}$$

We use $h_p\left(k_1,\ldots,k_p\right)$ to represent the p^{th} order Volterra kernel. Note the symmetry of the kernels is considered in (1). Without loss of generality, we can ignore the zero order term h_0. Since we consider the adaptive Volterra filter algorithm, we need indicate the time index n to the Volterra kernel. Thus, we write

$$y(n) = X(n)H(n)^T \tag{2}$$

where the signal $X(n)$ incorporates a special arrangement of the states of $x(n)$ in (1)

$$X(n) = \left[X_1(n), X_2(n), \ldots, X_C(n)\right] \tag{3}$$

Here $X_i(n)$ ($i = 1, 2, \ldots, C$) is a group of states with delayed relationship defined as

$$\begin{aligned}
X_1(n) &= [x(n), x(n-1), \ldots, x(n-N)] \\
X_2(n) &= [x(n)x(n), x(n-1)x(n-1), \ldots, x(n-N)x(n-N)] \\
&\vdots \\
X_C(n) &= [x(n)x(n-N)x(n-N)\ldots x(n-N)]
\end{aligned} \tag{4}$$

By this arrangement, we only need to calculate the first element of each group $X_i(n)$ ($i = 1, 2, \ldots, C$), and then the rest can be obtained by appropriate delay. Correspondingly, we can group the coefficients of the Volterra series as:

$$H(n) = \left[H_1(n), H_2(n), \ldots, H_C(n)\right] \tag{5}$$

with

$$\begin{aligned}
H_1(n) &= [h_1(0;n), h_1(1;n), \ldots, h_1(N-1;n)] \\
H_2(n) &= [h_2(0,0;n), h_2(1,1;n), \ldots, h_2(N-1, N-1;n)] \\
&\vdots \\
H_C(n) &= [h_p(0, N-1, N-1, \ldots, N-1;n)]
\end{aligned} \tag{6}$$

Alternatively, we can write Eq. (1) as:

$$y(n) = \sum_{i=1}^{C} X_i(n) H_i(n)^T = \sum_{i=1}^{C} y_i(n) \tag{7}$$

Eq. (7) represents the multi-channel representation of the Volterra filter as shown in Figure 2 (referred to as the diagonal representation in [10]). The output of the Volterra filter can be regarded as the sum of several linear filter outputs. Here C is the maximum number of groups (channels). $H_i(n)$ ($i = 1, \ldots, C$) is the set of i$^{\text{th}}$ channel coefficients; $H_1(n)$ represents the linear component of the Volterra series and the high order channels $H_i(n)$ ($i > 1$) represent nonlinear components of the Volterra series. The nonlinear echo cancellation problems can be regarded as an adaptive nonlinear system identification problem, i.e. the adaptive identification of the NIR. Using the multi-channel implementation of the Volterra filter, we see that to identify a nonlinear system is to identify each channel's coefficients, the $H_i(n)$ ($i = 1, \ldots, C$) in Fig. 2.

It is well known that to identify a linear system we should sample the input signal at its Nyquist frequency or higher. The input to the high order channels have a frequency much higher than the Volterra input $x(n)$. So it seems we need to sample the input $x(n)$ at a much higher frequency than its Nyquist frequency or we need to up-sample the nonlinear states of the higher order channels. However, a recent research result in [11] proves that, to identify the nonlinear system, we only need to sample the input signal $x(n)$ at its Nyquist frequency – exactly at twice the maximum frequency found in the input signal $x(n)$. Based on this theory, it is not necessary to up-sample the nonlinear states in $X_i(n)$ ($i > 1$). This means we don't need to consider the nonlinear spectral outgrowth to identify the high-order channels, i.e. we can treat each channel in Fig. 2 just like we would a linear filter.

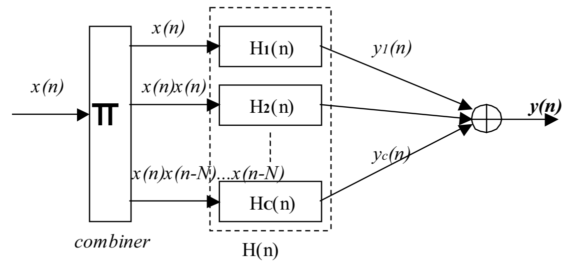

Figure 2. Multi-channel implementation of the Volterra filter

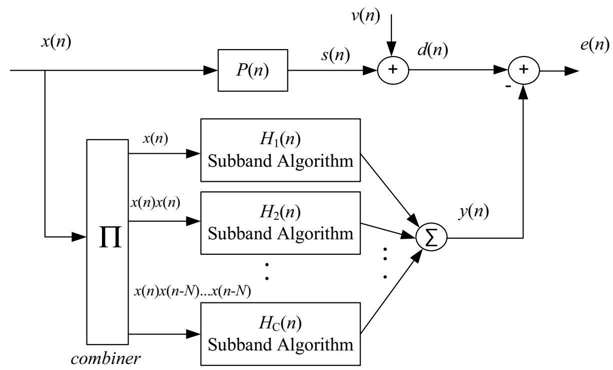

Figure 3. Sub-band adaptive Volterra filter based on multi-channel configuration

3. Sub-band implementation of an adaptive Volterra filter

From the above analysis, in the multi-channel implementation of the Volterra filter we can regard each channel in Fig. 2 as a linear filter without extra consideration concerning the spectral outgrowth. As a result, we can extend the linear sub-band adaptive filter techniques to develop an adaptive sub-band Volterra filter based on the multi-channel structure. The configuration of the developed sub-band adaptive Volterra filter is shown in Fig. 3. If one interprets this figure as a nonlinear AEC configuration, $x(n)$ represents the far-end signal and so is the reference signal provided to the adaptive Volterra filter, which in this case would usually be speech. $P(n)$ is the NIR; $s(n)$ represents the acoustic echo [1].

The conventional linear sub-band adaptive filters induced delay in the signal path through the introduction of the sub-band filters into that path. This delay limits the application of AEC [7]. Morgan *et. al.* [7] proposed the delayless sub-band adaptive linear filter as shown in Figure 4. In that configuration, the coefficients for each sub-band are updated independently and then combined through an FFT to yield the broadband coefficients. The Morgan configuration can greatly reduce the computational complexity when the linear system has a large order. However, in order to have a good approximation for each sub-band frequency response, each sub-band needs to have at least 4 coefficients. Also, to increase the convergence speed, at least 4 sub-bands are required. These limit the application of the Morgan sub-band configuration to applications with large-order channels. DeBrunner *et. al.* [8] introduced another configuration for the sub-band adaptive linear filter as shown in Figure 5 that directly updates the broadband coefficients based on all sub-band signals. Without up-sampling and down-sampling, the DeBrunner configuration has no limitations on the adaptive filter order; however, the computational complexity will increase as the number of channels increases.

In the multi-channel implementation of the Volterra filter, different channels have different lengths for fixed memory Volterra models. As shown in (6), the linear terms of $x(n)$ have a length $N+1$, but the channel with state $x(n)x(n-N)\ldots x(n-N)$ only has length one. As a result, neither the Morgan nor the DeBrunner configuration is suitable for our sub-band adaptive Volterra filter.

In our proposed sub-band adaptive Volterra filter, we combine these two configurations. This means that when the order of the channel is less than 64, we can implement the DeBrunner configuration; otherwise, we implement the Morgan configuration. As a result, we combine Figs. 3, 4, and 5, to yield our proposed sub-band adaptive Volterra filter. By subband decomposition, we can decrease the eigenvalue spread in each sub-band, and each sub-band can be updated using different step sizes. As a result, the convergence speed is greatly improved, and especially for colored inputs such as speech.

4. Sub-band implementation cost

In this section, we consider the implementation cost of the full-band adaptive Volterra filter and our proposed sub-band adaptive Volterra filter, both of which are based on the multi-channel configuration of the Volterra filter. We calculate the number of real multiplications per iteration for updating one channel of the adaptive Volterra filter with a total of M coefficients. We assume that each channel is updated by the simple least mean square (LMS) method in both the full-band and sub-band configurations. Note, however, that other updating methods can also be applied in our algorithm, such as the RLS and affine projection algorithms to further increase the convergence rate. For the full-band LMS based adaptive Volterra filter, we find that the LMS algorithm requires

$$R_F = 2M + 1 \tag{8}$$

real multiplications – one M for filtering and another M is required to calculate the instantaneous gradient, and one more multiplication is required for applying the step size. The number of real multiplications for one channel of length N in Fig. 3 based on Morgan's configuration can be obtained from [7]:

$$R_M = N + \frac{4(P+2N)}{Q} + \frac{16N}{Q^2} + 2\log_2(Q) + 3\log_2(N) \tag{9}$$

where Q is the number of sub-bands and P is the length of the prototype filter. The number of multiplications for one channel based on DeBrunner's configuration can be obtained from [8]:

$$R_D = Q(2L + N) + N + 1 \tag{10}$$

where Q is the number of sub-bands and L is the length of the sub-band filter. Thus, the total number of multiplications required for the sub-band adaptive Volterra filter is a combination of R_M and R_D. This computational complexity can be further reduced by using the recently developed multiple input multipliers [12] and sharing the sub-band error signals for different channels. From these calculations, we find that if most channels in the Volterra filter have large order, our configuration can greatly reduce the computational cost. For a low-order nonlinear system, the sub-band implementation will somewhat increase the computational complexity. However, this cost is offset by the fast convergence rate as shown in our next section. The recently introduced affine projection adaptive Volterra filter can greatly increase the convergence speed of the echo canceller. However, when compared to the full-band normalized LMS algorithm, the computational complexity increase is proportional to the order of the affine projection algorithm.

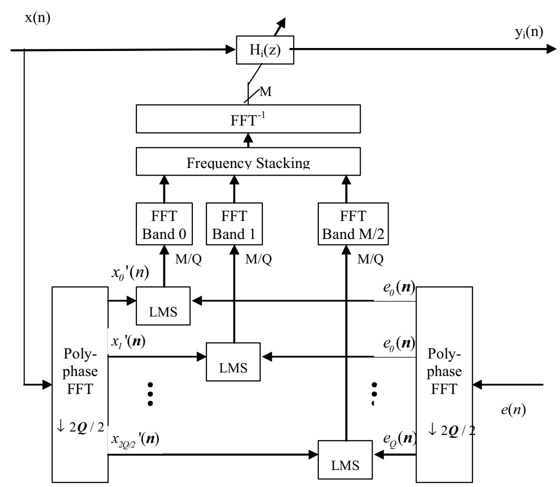

Figure 4. Delayless sub-band adaptive linear filter based on the Morgan configuration

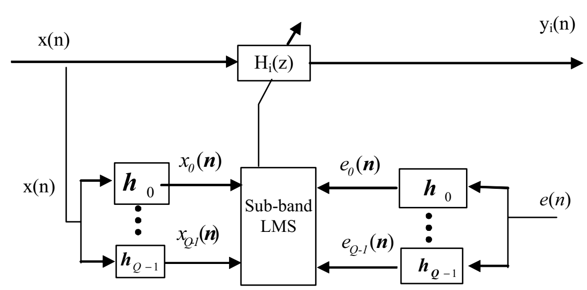

Figure 5. Delayless sub-band adaptive linear filter based on the DeBrunner configuration

5. Simulation of the new algorithm

We simulate the nonlinear AEC using the three approaches: 1) our proposed sub-band adaptive Volterra filter as shown in Fig. 3, 2) the full-band adaptive Volterra filter of [3], and 3) the affine projection adaptive Volterra algorithm with order 3 in [1]. To facilitate our simulation, the NIR is modeled as a Volterra filter with three channels: the first channel is a linear channel with length 128, while the second and third channels are nonlinear channels of length 128 and inputs $x(n)x(n-1)$ and $x(n)x(n-3)$, respectively. These are typical lengths of the impulse response for the hardest receivers. The Volterra channels' coefficients are shown in Fig. 6.

The speech signal $x(n)$ is a speech signal measured in our laboratory with sampling frequency of 20k Hz. The full-band adaptive Volterra filter and our proposed sub-band adaptive Volterra filter are updated using the normalized LMS algorithm. The step sizes of the adaptive Volterra filters for the different channels and the different sub-bands are tuned to ensure that the adaptive filters converge at their fastest convergence rate. The performance of the nonlinear AEC is measured by the echo return loss enhancement (ERLE):

$$ERLE = 10\log_{10}\frac{E(d^2(n))}{E(e^2(n))} \qquad (11)$$

The ERLE versus time for the nonlinear AEC using the full-band NLMS, proposed sub-band NLMS and the affine projection Volterra filters are shown in Fig. 7. Here, we find that the nonlinear AEC based on our proposed sub-band adaptive Volterra filter has a much faster convergence rate when compared to the one using the full-band adaptive Volterra filter. Our proposed algorithm does, however, converge slightly more slowly than the affine projection algorithm. Note, however, that the affine projection algorithm requires far more computational complexity than our proposed subband algorithm.

6. Conclusion

In this paper, an efficient delayless sub-band adaptive Volterra filter algorithm was proposed for the first time for nonlinear echo cancellation. Our simulations and analyses show that our method can increase the convergence rate of the adaptive Volterra filter and reduce the computational complexity for large order systems, which in turn improves the performance of the nonlinear echo cancellation systems. Other applications of our proposed algorithm will be in nonlinear system identification and adaptive nonlinear interference cancellation.

References

[1] A. Fermo, A. Carini, and G. L. Sicuranza, "Low-complexity nonlinear adaptive filters for acoustic echo cancellation in GSM handset receivers," *European Transactions on Telecommunications*, vol. 14, pp. 161-169, 2003.

[2] A. Guerin, G. Faucon, and R. Le Bouquin-Jeannes, "Nonlinear Acoustic Echo Cancellation Based on Volterra Filters," *IEEE Transactions on Speech and Audio Processing*, vol. 11, pp. 672-683, 2003.

[3] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," *Proc. ICASSP*, Mar 1999.

[4] F. Kuch and W. Kellermann, "Nonlinear echo cancellation using a simplified second order volterra filter," *Proc. ICASSP*, Orlando, Florida, 2002.

[5] C. Breining, P. Dreiseitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, "Acoustic echo control, an application of very-high-order adaptive filters," *IEEE Signal Processing Magazine*, vol. 16, pp. 42-69, 1999.

[6] W. A. Frank, "Efficient approximation to the quadratic Volterra filter and its application in real-time loudspeaker linearization," *Signal Processing*, vol. 45, pp. 97-113, 1995.

[7] D. R. Morgan and J. C. Thi, "Delayless sub-band adaptive filter architecture," *IEEE Trans. on Signal Processing*, vol. 43, pp. 1819-1830, 1995.

[8] V. DeBrunner, L. DeBrunner, and L. Wang, "Sub-band adaptive filtering with delay compensation for active control," *IEEE Trans. on Signal Processing*, Oct 2004.

[9] J. Mathews and G. L. Sicuranza, Polynomial Signal Processing: John Wiley & Sons, INC., 2000.

[10] G. M. Raz and B. D. Van Veen, "Baseband Volterra filters for implementing carrier based nonlinearities," *IEEE Trans. on Signal Processing*, vol. 46, pp. 103-114, 1998.

[11] J. Tsimbinos and K. V. Lever, "Input Nyquist sampling suffices to identify and compensate nonlinear systems," *IEEE Trans. on Signal Processing*, vol. 46, pp. 2833-2837, 1998.

[12] Y. Wang, *Efficient multiplierless implementation of adaptive filters in field programmable gate arrays,* Master Thesis, University of Oklahoma, Dec. 2004.

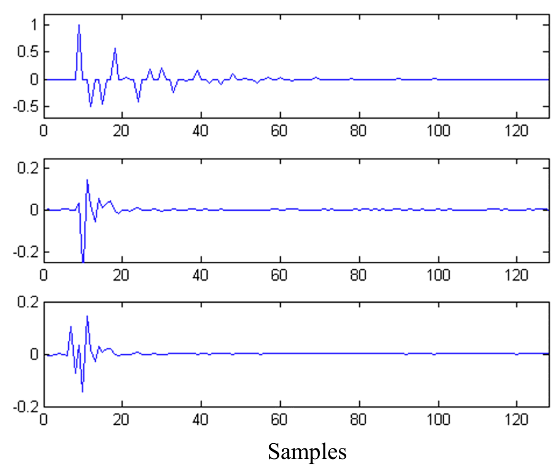

Figure 6. Three-channel Volterra nonlinear model. The top curve is linear channel coefficients; the middle and bottom curves are nonlinear channels.

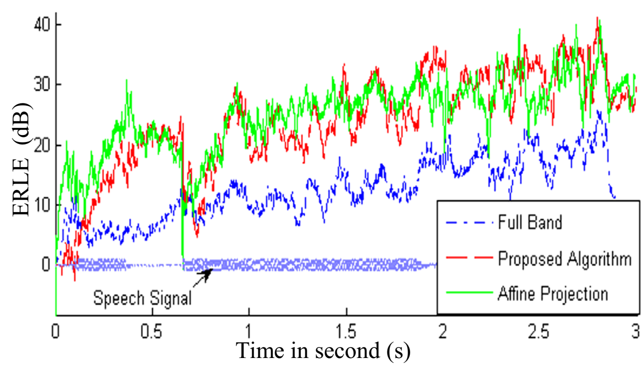

Figure 7. ERLE for different AEC algorithms