Turkish Journal of Electrical Engineering & Computer Sciences

http://journals.tubitak.gov.tr/elektrik/

Research Article

Turk J Elec Eng & Comp Sci
(2016) 24: 4849 – 4857

doi:10.3906/elk-1411-33

A second-order Volterra filter-based nonlinear clipping detector

Jihwan PARK, Joon-Hyuk CHANG*
Department of Electronics and Computer Engineering,

Received: 06.11.2014 • **Accepted/Published Online:** 18.09.2015 • **Final Version:** 06.12.2016

Abstract: In this paper, we propose a novel nonlinear clipping detector based on the second-order Volterra filter for an acoustic echo canceller (AEC). Since the performance of the conventional AEC algorithm drastically deteriorates when nonlinear clipping occurs, the adaptive filter pauses adaptation during the nonlinear clipping periods to avoid unwanted divergence. These nonlinear clipping periods are detected in our method using the magnitude spectrum of the quadratic Volterra filter in lower frequency ranges. Through extensive computer-based simulation results considering an acoustic room environment and various clipping levels, it is found that the proposed approach is effective in detecting the nonlinear clipping periods and improving AEC performance compared to the conventional AEC algorithm.

Key words: Acoustic echo canceller, nonlinear clipping detector, nonlinear echo, Volterra filter

1. Introduction

An acoustic echo often occurs because of acoustic coupling between a loudspeaker and microphone. Since this echo can significantly degrade the conversational quality between speakers, an acoustic echo canceller (AEC) eliminating acoustic echo from the microphone input is essential in mobile and electronic devices [1]. To date, much work has been conducted to develop a high-performance AEC. Acoustic echo cancelling is performed by adaptively modeling the echo path impulse response and subtracting an estimated echo from the microphone input [2–4]. Since most AECs utilize an acoustic echo assuming that acoustic replica is linearly dependent on an input signal of the loudspeaker, many adaptive filter algorithms are suitable to model a linear signal including the normalized least mean squares (NLMS), sign least mean squares (LMS), recursive least mean squares (RLS), proportionate NLMS (PNLMS), affine projection algorithm (APA), and the soft decision method [5,6]. However, the nonlinear characteristic of the amplifier and loudspeaker leads to nonnegligible nonlinear distortion [7], which cannot be sufficiently cancelled by the conventional linear AEC algorithms [8]. Because a nonlinear clipping input is detected as an uncorrelated noise by the AEC, the linear adaptive filter diverges in the nonlinear clipping period. This divergence phenomenon decreases the performance of the AEC not only in the nonclipping period but also in the clipping period. Therefore, a nonlinear clipping detector can be beneficial to detect the nonlinear clipping period and prevent linear filter divergence for a robust AEC.

In this paper, we present a nonlinear clipping detector to identify nonnegligible nonlinear distortion periods using a portion of the second-order Volterra filter [9–11], which efficiently characterizes speaker distortion [12,13]. Thus, the nonlinear clipping detector pauses the linear adaptive filter activity during the nonlinear clipping period so that the linear filter is updated only for the linear echo signal, which is the first approach

*Correspondence: jchang@hanyang.ac.kr

of detecting nonlinear clipping periods without a priori clipping information. It is similar to the double-talk detector, which freezes adaptive filter during the presence of near-end speech [14], but the nonlinear clipping detector operates during the presence of clipping periods and reduces irreducible error and improves speech quality, which is clearly different from the double-talk detector. We review the conventional AEC system in Section 2 and describe the proposed nonlinear clipping detector in Section 3. In Section 4, test environments and computer simulation results are described, while conclusions are drawn in Section 5.

2. Review of an acoustic echo canceller

Acoustic echo cancelling is performed based on a finite-duration impulse response (FIR) adaptive filter and is shown in Figure 1. A microphone input $y(n)$ is given as the sum of a near-end signal $s(n)$, an ambient noise signal $n(n)$, and an echo signal $d(n)$:

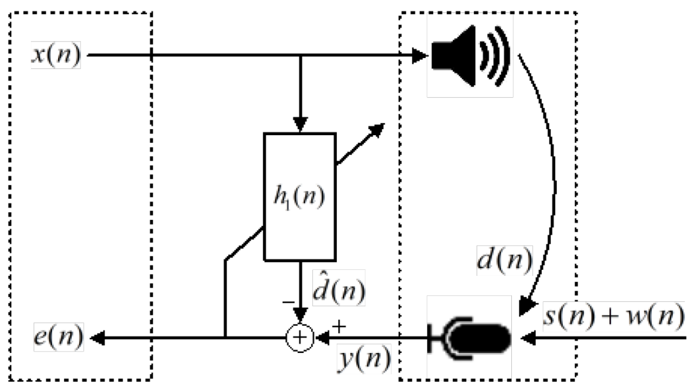

Figure 1. Block diagram of the conventional acoustic echo canceller.

$$\mathbf{y(n)} = \mathbf{s(n)} + \mathbf{n(n)} + \mathbf{d(n)} \tag{1}$$

where an echo signal $d(n)$ occurs when a far-end signal passes through an acoustic impulse response. Adaptive filter modeling utilizes an acoustic room impulse response and forms a replica $\hat{d}(n)$ of the echo signal $d(n)$:

$$\hat{d}\,(n) = \mathbf{h}_1^T\,(n)\,\mathbf{x}\,(n)\,, \tag{2}$$

where $\mathbf{h}_1\,(n) = [h_{1,0}\,(n)\,, h_{1,1}\,(n)\,, \cdots, h_{1,N_1-1}\,(n)]^T$ is a $(N_1 \times 1)$ coefficient vector, and $\mathbf{x}\,(n) = [x\,(n)\,, x\,(n-1)\,,$ $\cdots, x\,(n-N_1+1)]^T$ is a $(N_1 \times 1)$ far-end input vector. T represents the vector transpose. When this estimated echo signal $\hat{d}(n)$ is subtracted from the microphone input $y(n)$, the error signal is given as

$$e\,(n) = y\,(n) - \hat{d}\,(n)\,, \tag{3}$$

where the error signal of the AEC $e(n)$ is used to update the coefficients of the adaptive filter $\mathbf{h}_1(n)$ in order to model the realistic acoustic impulse response. To update these coefficients, the normalized least mean square (NLMS) algorithm is used in this paper. This NLMS-based adaptive filter algorithm is a widely used method because of its relative simplicity and good performance [1–5]. Indeed, adaptation of these coefficients according to the NLMS is performed such that

$$\mathbf{h}_1\,(n+1) = \mathbf{h}_1\,(n) + \frac{\mu_1}{P_{\mathbf{x}(n)} + \beta}\mathbf{x}\,(n)\,e\,(n)\,, \tag{4}$$

where μ_1 is a constant step-size to control adaptive filter convergence, β is a stabilization factor to ensure filter stability, and $P_{\mathbf{x}(n)}$ is the power of the far-end input vector $\mathbf{x}(n)$. However, the conventional linear AEC algorithm based on the NLMS scheme is not sufficient to eliminate the nonlinear clipping echo unlike the linear acoustic echo [8], as shown in Figure 2. During the nonlinear clipping period, the conventional AEC algorithm results in irreducible echo because of divergence of the NLMS adaptive filter used. Accordingly, we devise a way to prevent the adaptive filter from diverging in order to achieve a robust AEC technique in the following section. It has some advantages in that speech quality can be increased compared to the conventional AEC algorithm.

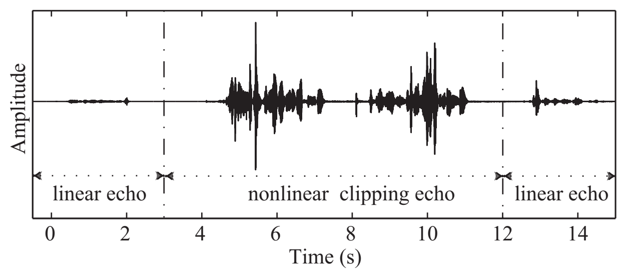

Figure 2. Output of the conventional NLMS adaptive filter for acoustic echo canceller during linear and nonlinear clipping period.

3. Proposed nonlinear clipping detector for AEC

In this section, we introduce a basic concept of the second-order Volterra filter in order to implement the nonlinear clipping detector in a framework of the AEC algorithm as shown in Figure 3. The Volterra filter is originally identical to a generalized Taylor series representation of a function with memory; hence, the Volterra filter is able to characterize a wide range of nonlinear systems [11]. Specifically, the second-order Volterra filter consists of linear and quadratic filters, and its output can be represented as

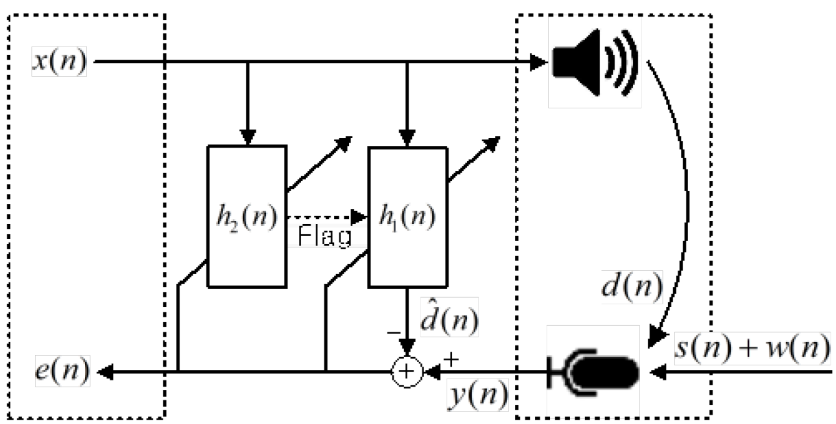

Figure 3. Block diagram of the nonlinear clipping detector based on the second-order Volterra filter for acoustic echo canceller.

$$y(n) = \sum_{i=0}^{N_1-1} h_1(i)\, x(n-i) + \sum_{i=0}^{N_2-1}\sum_{j=0}^{N_2-1} h_2(i,j)\, x(n-1)\, x(n-j), \tag{5}$$

where $x(n)$ denotes input signal, $y(n)$ represents output signal, and N_1 and N_2 denote the memory length of the linear (h_1) and quadratic (h_2) Volterra filters, respectively. The quadratic Volterra filter models the nonlinear characteristic when the linear Volterra filter cannot sufficiently cancel the echo signal from an unknown microphone input. We use the quadratic Volterra filter to classify the nonlinear clipping periods from the microphone input signal. This nonlinear clipping cannot be modeled by a linear filter such as moving average (MA), autoregressive moving average (ARMA), etc. [7,8]. As an example, the magnitude spectrum of the quadratic Volterra filter utilized during the nonlinear clipping and nonclipping periods is shown in Figures 4(a) and 4(b), respectively. In the case of the nonlinear clipping period, the magnitude spectrum of the quadratic Volterra filter is significantly increased in the lower frequency ranges compared to that of the nonclipping period. The magnitude spectrum of the quadratic Volterra filter is obtained by

$$\mathbf{H}_{2,n}\left(m_1, m_2\right) = F_{2D}\left\{\mathbf{h}_2\left(n\right)\right\}, \tag{6}$$

where

$$\mathbf{h}_2\left(n\right) = \begin{bmatrix} h_2(0,0) & \cdots & h_2(0, N_2-1) \\ \vdots & \ddots & \vdots \\ h_2(N_2-1, 0) & \cdots & h_2(N_2-1, N_2-1) \end{bmatrix} \tag{7}$$

Here F_{2D} represents the $(M \times M)$ two-dimensional discrete Fourier transform (2-D DFT), $\mathbf{H}_{2,n}(m_1 m_2)$ denotes 2-D DFT of the quadratic Volterra filter $h_2(n)$, and $0 \leq m_1 m_2 \leq M-1$. Based on the low frequency boosting characteristics of the quadratic Volterra filter, we classify the nonlinear clipping period in the unknown microphone input. For this, we use the summation of the magnitude spectrum from the quadratic Volterra filter in lower frequency ranges as an indicator. If this indicator is higher than a constant threshold, we classify this period as a nonlinear clipping period. As a result, the decision rule is devised by

$$\sum_{m_1=M/2-v}^{M/2+v} \sum_{m_2=M/2-v}^{M/2+v} \left|\mathbf{H}_{2,n}(m_1, m_2)\right| \;\overset{1}{\underset{0}{}}\; \eta, \tag{8}$$

where ν is a summation range factor of the magnitude spectrum of the quadratic Volterra filter and η denotes a given threshold. The nonlinear clipping detector is efficiently implemented by comparing the summation of a portion of the quadratic Volterra filter coefficient matrix with the constant threshold. By using this nonlinear clipping detector, we can obtain information of the nonlinear clipping periods in the microphone input signal. This nonlinear clipping detector information not only informs the linear filter to prevent filter updates at the nonlinear clipping periods, but also controls the overdriven amplifier. Notice that the quadratic Volterra filter is not used to cancel the acoustic echo but is only employed for the nonlinear clipping detector. As in Figure 4, the quadratic Volterra filter should be updated adaptively in order to model the nonlinear behavior from the output of the AEC algorithm. Moreover, the linear filter coefficients are updated to minimize error caused by subtracting the estimated echo from the unknown microphone input. However, the quadratic Volterra filter coefficients are adjusted in order to minimize the irreducible error, which cannot be achieved using a linear filter. For this reason, the sign LMS algorithm is adopted to adjust the quadratic Volterra filter coefficients such that

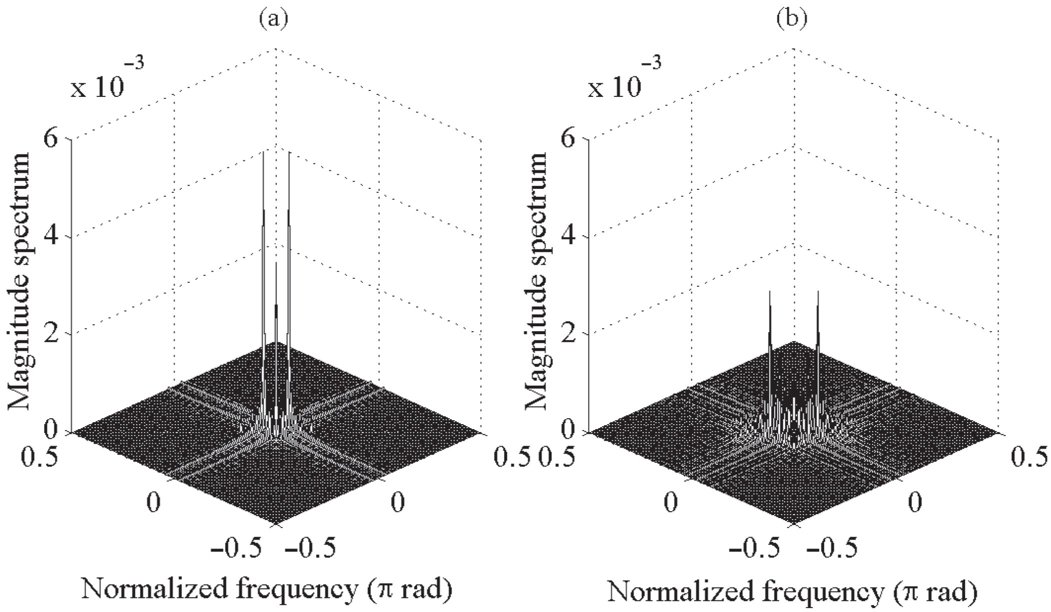

Figure 4. Frequency response of the quadratic Volterra filter: (a) Frequency response of the quadratic Volterra filter during the nonlinear clipping period, (b) Frequency response of the quadratic Volterra filter during the nonclipping period.

$$\mathbf{h}_2(n+1) = \mathbf{h}_2(n) + \mu_2 \mathbf{x}_2(n)\,\mathrm{sgn}\left[e(n)\right], \tag{9}$$

where

$$\mathbf{x}_2(n) = \begin{bmatrix} x^2(n) & \cdots & x(n)\,x(n-N_2+1) \\ \vdots & \ddots & \vdots \\ x(n-N_2+1)\,x(n) & \cdots & x^2(n-N_2+1) \end{bmatrix} \tag{10}$$

and

$$\mathrm{sgn}(\mathrm{a}) = \begin{cases} 1, & a > 0 \\ 0, & a = 0 \\ -1, & a < 0 \end{cases}. \tag{11}$$

4. Simulation and results

In order to measure the performance of the proposed nonlinear clipping detector based on a second-order Volterra filter for AEC, we carried out objective tests under various acoustic environments. In the experiments, the parameters were set as follows: $N_1 = [64, 128, 256, 512]$, $\mu_1 = 0.3$, $\beta = 1.0$, $N_2 = [64, 128, 256, 512]$, and $\mu_2 = 0.3$. For these tests, 20 sentences, which were sampled at 8 kHz and represented in 16-bit format from two male and two female speakers, were used to construct speech data files. Uncontaminated and nonclipped speech was used for the far-end signal. The near-end signal was artificially obtained by the following procedures. The clean far-end signal was normalized and amplified to various volumes than normalized signal where out of range from -2^{15} to 2^{15} corresponding to the nonlinear clipping and truncated by $\pm 2^{15}$. Next the nonlinear clipping far-end signal was passed through a room impulse response [15]. This room impulse response was designed to model a small office room having a size of $5 \times 4 \times 3m^3$. Here the location of speaker and microphone were set as 0.5 m, 2 m, 1.5 m and 3.5 m, 2 m, 1 m, respectively. From these conditions, simulated room impulse response was generated as shown in Figure 5. Lastly, the volume of the room impulse response's output was decreased by as much as 3.5 dB in a mobile communication environment [14]. When employing the algorithm, all tests were conducted in a single-talk situation. For the purpose of objective assessments, we compared the performance of the proposed method and the conventional NLMS AEC algorithm. The performance of the proposed algorithm was measured in terms of echo return loss enhancement (ERLE) and speech attenuation (SA) [14]:

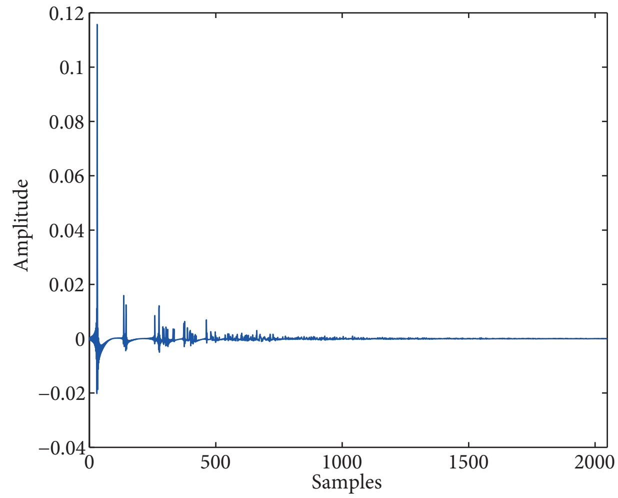

Figure 5. Simulated room impulse response (reverberation time = 200 ms).

$$ERLE\,(n) = 10\log_{10}\left[\frac{E\left[y^2\,(n)\right]}{E\left[e^2\,(n)\right]}\right] \tag{12}$$

$$SA = \frac{1}{N}\sum 10\log_{10}\left[\frac{E\left[s^2(n)\right]}{E\left[e^2(n)\right]}\right], \tag{13}$$

where n is the sample index in the discrete-time domain, $E[\cdot]$ denotes the expectation operator, N is the number of samples during the nonlinear clipping periods, and $s^2(n)$ denotes the power of the clean far-end speech component in the output signal power $e^2(n)$. Actually, ERLE was used only in the nonclipping period for measuring the performance of the acoustic echo cancellation and SA was used only in the nonlinear clipping period in order to assess the signal degradation performance. In Figure 6, an example of the AEC including the nonlinear clipping detector is represented. Figure 6(a) illustrates the microphone input and Figure 6(b) shows the summation of the magnitude spectrum of the quadratic Volterra filter coefficients in lower frequency ranges in conjunction with a constant decision threshold of 0.005. The result of the nonlinear clipping decision was compared with the manual marking of the nonlinear clipping period as shown in Figure 6(c). This figure illustrates that this proposed decision rule can be used as the nonlinear clipping detector. Furthermore, Figure 6(d) shows ERLE and SA results of the proposed method compared to the conventional NLMS method. The result of the ERLE test shown in Figure 6(d) indicates that the proposed nonlinear clipping detector based on the second-order Volterra filter well preserves the speech signal from the nonlinear clipping phenomenon during the nonlinear clipping periods compared to the conventional NLMS method. Given the various clipping levels, overall results for the ERLE and SA scores were averaged to yield the final mean score results, which

Table. Comparison of ERLE and SA score in various clipping levels and filter lengths.

		$N_1, N_2 = 64$		$N_1, N_2 = 128$		$N_1, N_2 = 256$		$N_1, N_2 = 512$	
Clipping Level (dB)	Method	ERLE (dB)	SA (dB)	ERLE (dB)	SA (dB)	ERLE (dB)	SA (dB)	ERLE (dB)	SA (dB)
15	Conventional	8.08	7.60	11.89	10.56	22.11	15.62	22.51	15.67
	Proposed	**9.20**	**7.04**	**13.38**	**10.32**	**23.45**	**15.17**	**23.69**	**15.07**
16	Conventional	7.44	6.05	10.38	8.59	21.09	14.58	21.40	14.29
	Proposed	**7.83**	**5.83**	**11.44**	**8.16**	**22.91**	**13.65**	**22.60**	**13.37**
17	Conventional	6.09	4.81	8.66	7.14	19.72	13.40	19.73	12.76
	Proposed	**6.72**	**4.55**	**10.00**	**6.37**	**21.23**	**12.91**	**21.39**	**12.12**
18	Conventional	4.86	3.79	6.95	5.18	18.02	11.51	18.21	10.79
	Proposed	**5.46**	**3.91**	**8.11**	**5.15**	**19.40**	**11.34**	**19.07**	**10.64**
19	Conventional	3.97	2.64	5.47	3.92	16.21	10.49	16.42	9.66
	Proposed	**5.03**	**2.24**	**7.08**	**3.24**	**17.64**	**9.69**	**17.24**	**9.64**
20	Conventional	3.49	2.19	4.35	2.79	14.49	9.05	14.71	8.04
	Proposed	**3.82**	**1.78**	**5.38**	**2.37**	**15.79**	**8.79**	**16.04**	**7.35**

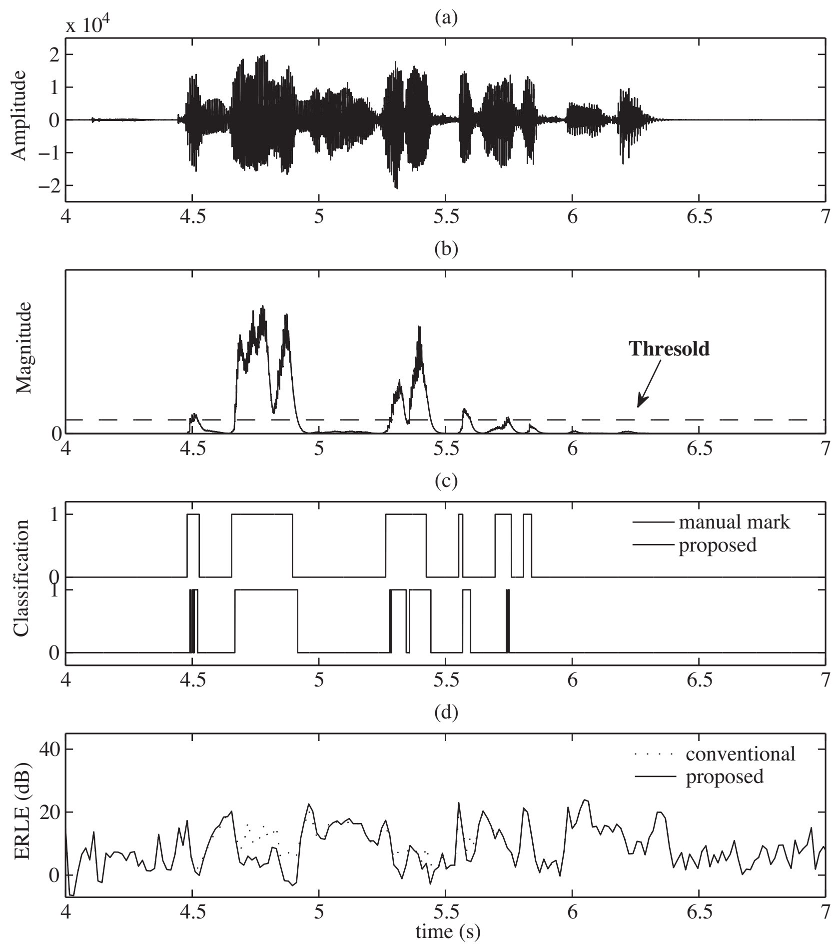

Figure 6. Performance comparison of the proposed algorithm and the conventional method based on ERLE test: (a) Microphone input waveform, (b) Summation of the magnitude spectrum of the quadratic Volterra filter coefficients in lower frequency ranges and the decision threshold, (c) Nonlinear clipping detector result (down) compared to the manual mark (up), (d) Results of the ERLE test of the proposed algorithm and the conventional AEC algorithm.

are shown in the Table. From the Table, it is evident that the proposed AEC employing a nonlinear clipping detector based on a second-order Volterra filter yielded a higher ERLE and lower SA at various clipping levels and filter lengths compared to the conventional technique. From this, we conclude that the proposed algorithm is superior to the previous scheme at all of the tested clipping levels.

5. Conclusions

In this paper, we have proposed a new nonlinear clipping detector for an AEC based on the second-order Volterra filter. The proposed nonlinear clipping detector for the AEC is composed of two parts in which the conventional AEC and the quadratic Volterra filter are used to eliminate linear components of acoustic echo and characterize the nonlinear clipping periods, respectively. For characterizing the nonlinear behavior of the unknown input signal, the quadratic Volterra filter is updated in order to minimize the estimation error of the conventional AEC by using the sign LMS algorithm. The summation of lower frequency ranges of the quadratic Volterra filter coefficients, which is higher than a constant threshold, is employed as an indicator freezing the adaptive filter of the conventional AEC during nonlinear clipping periods.

The performance of the AEC including the nonlinear clipping detector was superior to that of the conventional AEC in objective tests of signal distortion and echo cancellation during the nonlinear clipping period. However, this algorithm demands high computational cost because adaptation of the second-order Volterra filter needs many more operations than the adaptive linear filter does. Therefore, we further need to study the frequency-domain second-order Volterra filter-based nonlinear clipping detector for decreasing computational complexity.

References

[1] Hänsler E, Schmidt G. Acoustic Echo and Noise Control: A Practical Approach. New York, NY, USA: Wiley, 2004.

[2] Breining C, Dreiseitel P, Haensler E, Mader A, Nitsch B, Puder H, Schertler T, Schmidt G, Tilp J. Acoustic echo control. An application of very-high-order adaptive filters. IEEE Signal Proc Mag 1999; 16; 42-69.

[3] Gay S, Benesty J. Acoustic signal processing for telecommunication. Boston, MA, USA: Kluwer, 2000.

[4] Sondhi M. The history of echo cancellation. IEEE Signal Proc Mag 2006; 23; 95-98.

[5] Diniz PSR. Adaptive Filtering: Algorithms and Practical Implementation. New York, NY, USA: Springer, 1997.

[6] Park YS, Chang JH. Frequency domain acoustic echo suppression based on soft decision. IEEE Signal Proc Let 2009; 16: 53-56.

[7] Stenger A, Stenger E, Kellermann W. Adaptation of a memoryless preprocessor for nonlinear acoustic echo cancelling. Signal Process 2000; 80: 1747-1760.

[8] Costa M, Bermudez J, Bershad N. Statistical analysis of the LMS algorithm with a zero-memory nonlinearity after the adaptive filter. In: IEEE 1999 International Conference on Acoustics, Speech, and Signal Processing, 15–19 March 1999; Phoenix, AZ, USA: IEEE. pp. 1661-1664.

[9] Stenger A, Rahenstein R. Adaptive Volterra filters for acoustic echo cancellation. In: IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, 20–23 June 1999; Antalya, Turkey: IEEE. pp. 679-683.

[10] Fermo A, Carini A, Sicuranza G. Simplified Volterra filters for acoustic echo cancellation in GSM receivers. In: Signal Processing Conference, 4–8 September 2000; Tampere, Finland: IEEE. pp. 1-4.

[11] Gurin A, Gaucon G, Bouquin-Jeannes RL. Nonlinear acoustic echo cancellation based on Volterra filters. IEEE T Acoust Speech 2003; 11: 672-683.

[12] Frank W. An efficient approximation to the quadratic Volterra filter and its application in real-time loudspeaker linearization. Signal Process 1995; 45; 97-113.

[13] Burton TG, Goubran RA, Beaucoup F. Nonlinear system identification using a subband adaptive Volterra filter. IEEE T Instrum Meas 2009; 58: 1389-1397.

[14] Lee KH, Chang JH, Kim NS, Kang S, Kim Y. Frequency domain double-talk detection based on the Gaussian mixture model. IEEE Signal Proc Let 2010; 17: 453-456.

[15] McGovern SG. Fast image method for impulse response calculations of box-shaped rooms. Appl Acoust 2009; 70: 182-189.