AN ANALYSIS OF LOUDSPEAKER DISTORTION IN THE CONTEXT OF ACOUSTIC ECHO CANCELLATION

Trevor Burton[1] and Rafik Goubran[1]
[1]Department of Systems and Computer Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada, K1S 5B6

1. INTRODUCTION

Acoustic echo is inherent in all hands-free communication systems and can corrupt the conversation between parties if it is not sufficiently suppressed. The conventional approach to removing echo is with a digital echo canceller (EC), which is typically implemented as a linear adaptive filter. A block diagram of a simplified echo cancellation system is shown in Figure 1. Here the reference signal, $x(n)$, is the far-end signal that is played through the loudspeaker and the response signal, $d(n)$, is the signal captured by the microphone. The microphone signal is comprised of the echo, $y(n)$, local noise, $\eta(n)$, and local talker, $v(n)$, signals. The echo signal is formed from the direct loudspeaker to microphone signal along with the reflections from the walls and objects within the acoustic environment. The EC determines an approximation of the loudspeaker-enclosure-microphone-system (LEMS) transfer function via an adaptive filtering algorithm, and produces an echo signal estimate, $\hat{y}(n)$, that is subtracted from $d(n)$ to cancel the unwanted $y(n)$. The resulting error signal, $e(n)$, is then transmitted back to the far-end of the communication system. It should be noted that the EC is only adapted under quiet local talker conditions (i.e. $v(n)=0$). For practical acoustic echo cancellation (AEC), a doubletalk detector is used to determine if a local talker is active or not and controls adaptation of the EC accordingly.

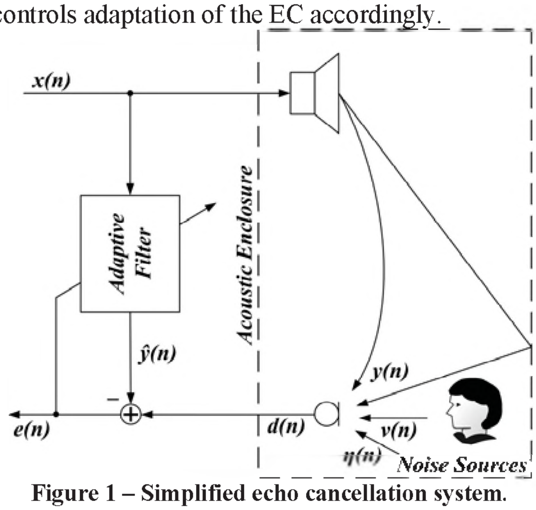

Figure 1 – Simplified echo cancellation system.

The performance of an EC is limited by many factors including undermodeling of the echo path, time variations and background noise within the hands-free environment, and by echo path nonlinearities [1]. Given that the EC accurately models the linear portion of the echo path and that the hands-free environment is stationary with an acceptable amount of background noise, the main factor limiting its performance becomes the echo path nonlinearities. These nonlinearities include loudspeaker distortion, vibrations within the hands-free device, and amplifier saturation [1], [2]. Assuming that the amplifier in the hands-free device is not overdriven, the main source of echo path nonlinearity can become loudspeaker harmonic distortion which is effectively modeled by Volterra series expansions [3]. Thus, a nonlinear EC is required to prevent loudspeaker distortion from degrading the quality of hands-free communication between parties.

This work presents an analysis of loudspeaker distortion based on experimental measurements obtained from several hands-free systems under various operating conditions. The results of the analysis reveal trends in the frequency domain nature of the loudspeaker distortion, which provides insight into designing computationally efficient nonlinear echo cancellers.

2. LOUDSPEAKER MODELING IN AEC

To improve AEC performance compared to a linear EC, adaptive Volterra filters can be used to model the linear portion of the unknown acoustic system along with the nonlinear loudspeaker distortion. The general discrete time P-th order finite memory Volterra series expansion of an input signal, $x(n)$, is given by [3]:

$$y(n) = h_0 + \sum_{p=1}^{P} \left[\sum_{m_1=0}^{L_p-1} \sum_{m_2=m_1}^{L_p-1} \cdots \sum_{m_p=m_{p-1}}^{L_p-1} h_p(m_1, m_2, \ldots, m_p) \times x(n-m_1)x(n-m_2)\cdots x(n-m_p) \right] \quad (1)$$

where h_0 is a constant term and $h_p(m_1,\ldots,m_p)$ are the p-th order Volterra kernels or impulse responses that characterize the nonlinear system with memory lengths L_p. The biggest drawback to using adaptive Volterra filters is their very large complexity requirement, which can be prohibitive for practical AEC especially as the order of the Volterra series increases. Thus, reduced complexity nonlinear AEC structures based on adaptive Volterra filters are highly desirable, particularly with the movement towards wideband telephony systems. As wideband systems operate at a 16 kHz sampling rate, twice as many taps are required to model the same LEMS as a standard 8 kHz narrow band system.

3. EXPERIMENTAL RESULTS

In this section a harmonic distortion analysis is presented for a telephone set with a 2.5 inch loudspeaker and for a cellular telephone with a miniaturized loudspeaker, that were both operated under high volume conditions in hands-free mode. The distortion data was collected by recording the microphone signal, under maximum preamplifier gain, at a 48 kHz sampling rate for various test tones between 50 Hz and 7 kHz for the telephone set, and between 300 Hz and 7 kHz for the cellular telephone. Each test tone was applied to the telephone set loudspeaker at a root-mean-square (RMS) voltage of 1.4 and to the cellular telephone loudspeaker at 1.1. Ten harmonics were included in the distortion calculation or as many that could fit up to the Nyquist rate for fundamental frequencies higher than 2.4 kHz.

As shown in Figure 2 the loudspeaker of the telephone set exhibits significant second harmonic distortion of up to 52% for low fundamental frequencies between 50 and 300 Hz. At frequencies beyond this range the second harmonic distortion decreases rapidly to much smaller values. Significant third harmonic distortion of up to 28% occurs as well for frequencies between 50 and 150 Hz, before decaying towards zero at higher frequencies. Fourth and higher harmonic distortion of up to 25% occurs for frequencies between 50 and 100 Hz and around 400 Hz.

The loudspeaker of the cellular telephone reaches up to 25% third harmonic distortion at low fundamental frequencies between 300 and 500 Hz. At higher frequencies the third harmonic distortion tends to be quite small with the exception of around 800Hz and 3 kHz. Considerable second harmonic distortion of up to 55% occurs for fundamental frequencies near 1 and 1.5 kHz, and between 3 and 4 kHz. Only a small amount of fourth and higher harmonic distortion of up to 12% is displayed between 300 and 500 Hz and around 1.5 kHz.

Based on these harmonic distortion results, nonlinear AEC structures with adaptive Volterra filters up to the third order and possible higher would be required to effectively compensate for the resulting nonlinear echoes under these high signal level conditions. As hands-free devices tend to be operated under high volume conditions, it is important that the underlying EC models and removes the resulting distortion to ensure a high quality of communication. With the increased signal bandwidth of wideband hands-free systems, the need for nonlinear ECs will become more essential.

One approach to reduced complexity nonlinear ECs based on Volterra filters is with subband adaptive filtering structures. In this case complexity is reduced by performing the adaptive filtering operations in multiple frequency bands at a reduced sampling rate, where shorter adaptive filter lengths are employed. Furthermore, a subband structure allows the frequency domain nature of loudspeaker distortion to be exploited. As shown in Figures 2 and 3 the loudspeaker distortion of typical hands-free devices is confined mainly to specific frequency regions, thus applying adaptive Volterra filtering in only the correspondingly affected subbands would help to further reduce complexity.

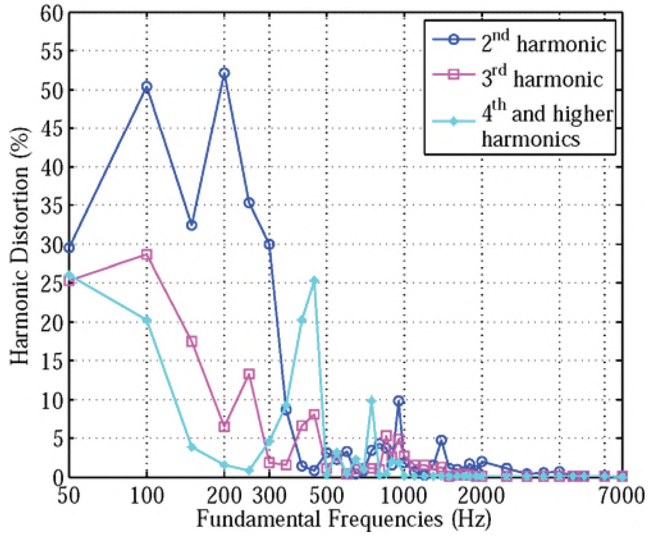

Figure 2 – Harmonic distortion for the telephone set.

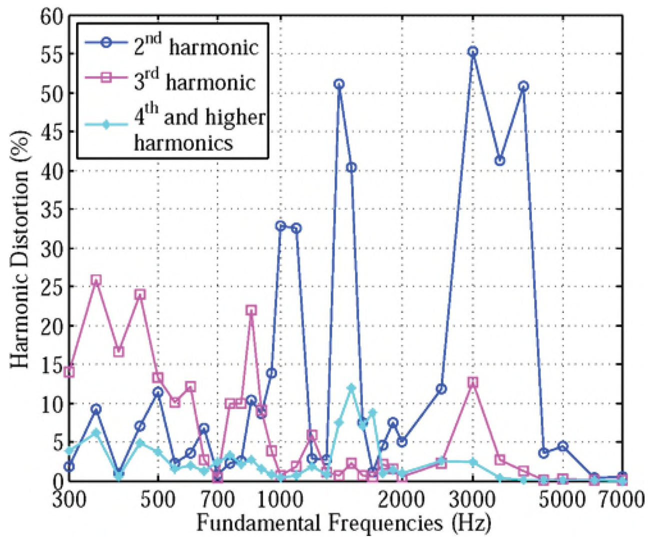

Figure 3 – Harmonic distortion for the cellular telephone.

4. CONCLUSIONS

This paper presented a distortion analysis for the loudspeakers of typical hands-free devices under high volume conditions. The results revealed that the distortion is confined to specific frequency regions, which provides insight into designing reduced complexity nonlinear ECs based on subband adaptive Volterra filters.

REFERENCES

[1] A. N. Birkett and R. Goubran, "Nonlinear loudspeaker compensation for hands free acoustic echo cancellation," *IEE Electronics Letters*, vol. 32, no. 12, pp. 1063–1064, Jun. 1996.

[2] A. Stenger, L. Trautmann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive Volterra filters," in Proc. IEEE ICASSP, vol. 2, Mar. 1999, pp. 877–880.

[3] V. J. Mathews and G. L. Sicuranza, *Polynomial Signal Processing*. New York, NY: Wiley, 2000.