International Workshop on Acoustic Echo and Noise Control (IWAENC2003), Sept. 2003, Kyoto, Japan

EFFECTS OF HARMONIC COMPONENTS GENERATED BY POLYNOMIAL PREPROCESSOR IN ACOUSTIC ECHO CONTROL

Tuomo Mäkelä and Riitta Niemistö

Institute of Signal Processing, Tampere University of Technology
P.O. Box 553, 33101 Tampere, Finland
e-mail: {tuomo.makela, riitta.niemisto}@tut.fi

ABSTRACT

The aliasing problem in adaptive polynomial filtering is described and two different approaches to overcome it are proposed. The polynomial model considered is the separable homogeneous Volterra filter, which is a cascade of polynomial preprocessor and a linear filter. The polynomial preprocessor produces higher order harmonics to the output and they are aliased to lower frequencies. It is shown that if the aliasing is taken into account the performance of acoustic echo cancelers can be improved by 1–2 dB.

1. INTRODUCTION

In this paper we focus on a problem related to polynomial model used for compensating acoustic distortion produced by an inexpensive loudspeaker in hands free set.

Traditionally, acoustic echo path is modeled with linear FIR filter

$$\hat{y}(n) = \sum_{m=0}^{M-1} w_m x(n-m) \tag{1}$$

where $x(n)$ is input (far-end signal), $\hat{y}(n)$ output (replica of the echo) and M is the number of filter parameters $w_0, w_1, \ldots, w_{M-1}$. The performance of this linear model suffers when the echo path is nonlinear, e.g. when small inexpensive loudspeakers are used. We have successfully used polynomial models for compensating acoustic distortion originating from low-powered amplifier and small loudspeaker on high volumes [1, 2].

Polynomial preprocessor [3, 4] that is shown in Figure 1 was first introduced to model acoustic distortion produced by low quality amplifiers. The model for the echo path becomes

$$\hat{y}(n) = \sum_{m=0}^{M-1} w_m \sum_{q=1}^{Q} a_q x^q(n-m), \tag{2}$$

THIS WORK IS A PART OF A RESEARCH PROJECT FUNDED BY THE NATIONAL TECHNOLOGY AGENCY (TEKES) AND NOKIA RESEARCH CENTER

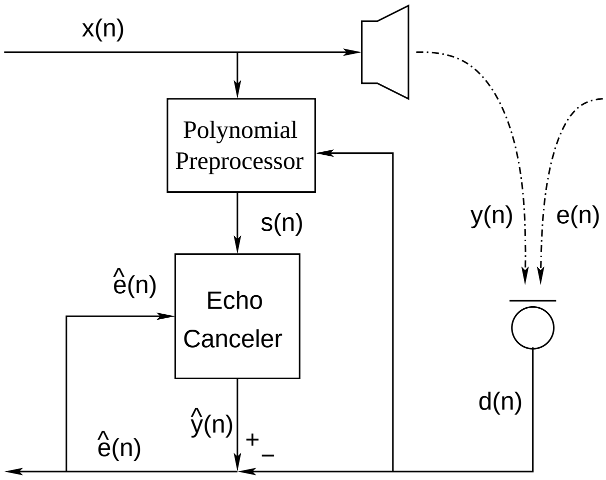

Figure 1: Block diagram of the echo canceller system with signals: $x(n)$ far-end, $y(n)$ far-end echo,$e(n)$ near-end, $d(n)$ near-end microphone, $s(n)$ output of the preprocessor, $\hat{y}(n)$ the replica of far-end echo and $\hat{e}(n)$ the estimation error

where Q is the order of the nonlinearity. The polynomial part (3) of the model (2) generates higher harmonic frequencies to the polynomial preprocessor output

$$s(n) = \sum_{q=1}^{Q} a_q x^q(n). \tag{3}$$

that are aliased to lower frequencies unless sampling frequency is sufficiently high. In this paper we propose two straight forward methods to overcome this problem and discuss their limitations.

This paper is organized as follows. In Section 2 we introduce the problem related to polynomial preprocessor and aliasing and give two straight forward solutions for the problem. Simulation arrangements and experimental results are presented in Section 3 and some final conclusions are drawn in Section 4.

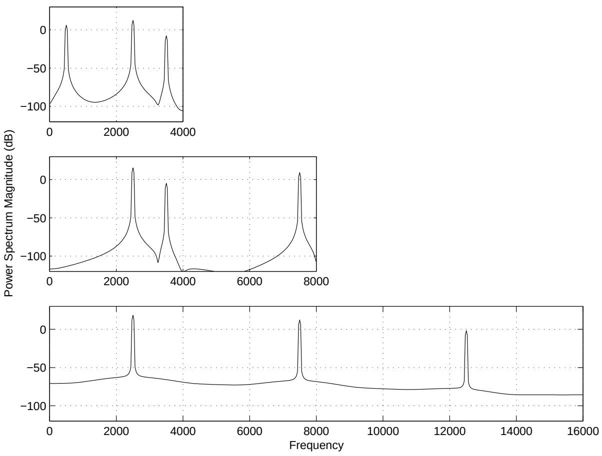

Figure 2: Aliasing effect: Spectrum of x^5 where $x = \sin(2\pi n 2500/Fs)$, Fs = 8000Hz, 16000Hz and 32000Hz

2. POLYNOMIAL MODEL AND ALIASING

Polynomial preprocessor (3), shown also in Figure 1, generates harmonic frequencies to the polynomial preprocessor output $s(n)$. If the sampling frequency is too low those harmonic frequencies are aliased to the lower frequencies.

The harmonic frequencies appear when computing powers of input in the polynomial preprocessor ([5], pp. 44, 54–55). When the the signal is preprocessesed by (3), such frequencies are generated to the output $s(n)$ that are not presented in the input $x(n)$.

As an example, let us consider sine signal

$$x(n) = \sin(n\omega) \tag{4}$$

and use it as input for polynomial preprocessor of order $Q = 2$. The output results as

$$s(n) = a_1 \sin(n\omega) + a_2 \sin^2(n\omega) \tag{5}$$

$$s(n) = a_1 \sin(n\omega) + \frac{a_2}{2}(1 - \cos(n2\omega)) \tag{6}$$

There is now two frequencies ω and 2ω present in the output signal $s(n)$, while only the base frequency ω was present in the input signal $x(n)$. If the order of the polynomial preprocessor is higher then there exists more and higher harmonic frequencies also in the output. If the frequency ω is higher than the half of the Nyquist frequency then all the higher frequencies in the output are aliased. This aliasing may decrease the advantage achieved by the adaptive polynomial preprocessor in acoustic echo control.

The aliasing phenomena is illustrated in Figure 2 where 2.5 kHz sinusoidal signal is raised to the 5th power. Three different sampling frequencies 8 kHz, 16 kHz and 32 kHz are used. As shown in the figure, the harmonics at 5 kHz and 7.5 kHz are aliased unless the sampling frequency is sufficiently high.

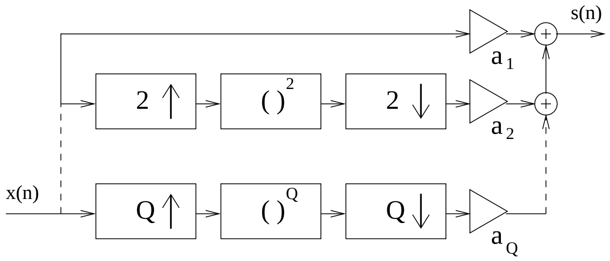

Figure 3: Block diagram of the upsampling and downsampling solution,

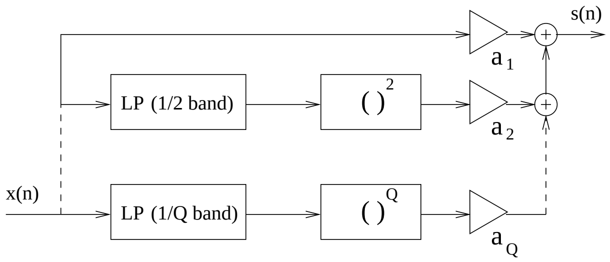

Figure 4: Block diagram of the low-pass filtering solution.

The most straight-forward method to overcome this aliasing problem is to upsample the signal by the factor of q before raising it to the qth power, and downsample signal afterwards, as illustraded in Figure 3. This approach is computationally demanding because the polynomial order Q may become very high in polynomial filtering, for example, order $Q = 13$ is used in [3].

As another approach to the aliasing problem, we have also considered an approach where the input signal is filtered with the low-pass filter with stop-band edge at the $1/q$ of the Nyquist frequency and only then raised to the qth power. The block diagram of the preprocessor of this solution is shown in Figure 4.

3. SIMULATIONS

In the following simulations the linear filter is adapted using normalized LMS algorithm and the polynomial preprocessor is adapted using the QR–RLS algorithm.

The data we use is measured in a halted car in garage using a hands-free set that has an inexpensive distorting loudspeaker. Test signal consist of four Finnish sentences grouped in two groups and said first by a male speaker and then by a female speaker. All simulation signals are recorded in single-talk situation where there is no other near-end excitation than echoes from the far-end.

The performance of echo cancellers is measured in terms of Echo Return Loss Enhancement (ERLE) defined by

$$\mathrm{ERLE} = 10 \log_{10} \frac{\mathrm{E}(y^2(n))}{\mathrm{E}(\hat{e}^2(n))}, \tag{7}$$

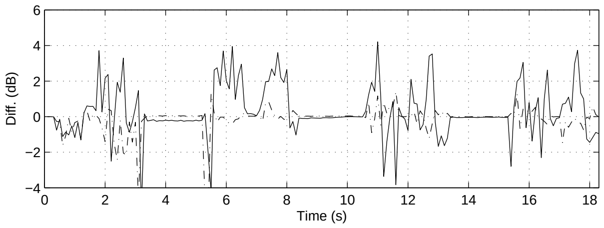

Figure 5: Adaptive linear filter with polynomial preprocessor compared to the ordinary adaptive linear filter in intermediate (dashdot) and high (solid) volume level.

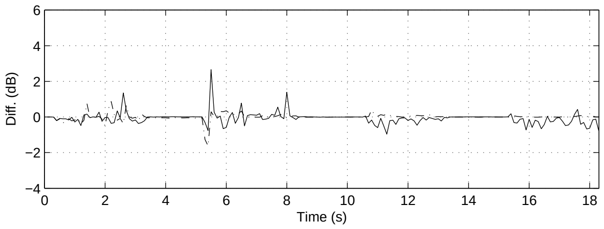

Figure 6: Difference between the upsampling-downsampling approach and the linear filter with ordinary polynomial preprocessor in intermediate (dashdot) and high (solid) volume level.

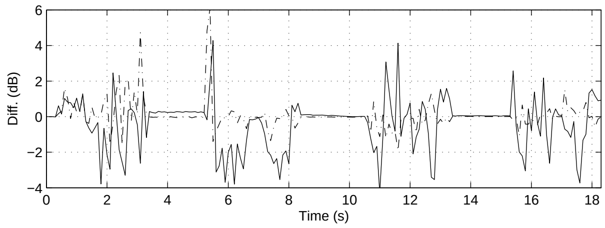

Figure 7: Difference between the lowpass filtering approach and the linear filter with ordinary polynomial preprocessor in intermediate (dashdot) and high (solid) volume level.

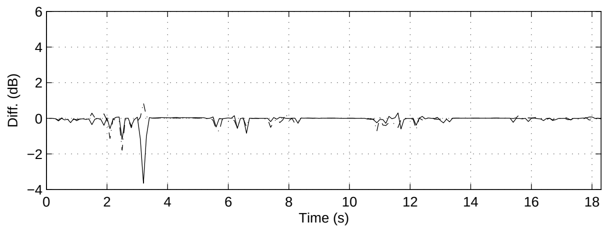

Figure 8: Difference between adaptive linear filtering and lowpass filtering approach in intermediate (dashdot) and high (solid) volume level.

which is estimated from data in 100 ms windows. Far-end echo signal $y(n)$ is estimated with microphone signal $d(n)$.

First we notice that the linear adaptive filter with adaptive polynomial preprocessor has better performance than the ordinary linear adaptive filter, as shown in the Figure 5.

When we used the upsampling-downsampling approach we get at the most 1–2 dB enhancement in ERLE compared to the normal adaptive filter with adaptive polynomial preprocessor, as shown in the Figure 6.

The improvement is the most significant in the beginning of speech activity that indicates that sensitivity to speech activity detection, shown in Figure 5 and observed in [1, 2], is possibly caused by aliasing.

We also tested lowpass filtering approach. As shown in the Figure 7, this model causes significant performance loss compared to the normal polynomial preprocessor and is no better than the linear model, as shown in Figure 8.

Finally, we tested these two different models to overcome aliasing effects with a non-adaptive model, where near-optimal values for both preprocessor parameters $a_1, a_2, \ldots, a_Q$ and filter parameters $w_0, w_1, \ldots, w_{M-1}$ in (2) are computed in turns, so that first the parameters of preprocessor are kept constant and system is solved in least squares sense respect to filter parameters, then filter parameters are kept constant and system is solved respect to preprocessor parameters. The process is depicted in Table 1. We noticed that difference between the parameters from different iterations $\Delta\mathbf{w} = \mathbf{w}(k+1) - \mathbf{w}(k)$ and $\Delta\mathbf{a} = \mathbf{a}(k+1) - \mathbf{a}(k)$, where k is iteration, diminishes quite rapidly.

One should note, that the whole signal is used for computing parameters $w_0, w_1, \ldots, w_{M-1}$ and $a_1, a_2, \ldots, a_Q$ in each iteration, so that resulting pre-

Table 1: Computation of the parameters of the non-adaptive model

Initialization:

1. $w_m = 0$ for $m = 0, 1, \cdots, M-1$
2. $a_1 = 1$, $a_q = 0$, $q = 2, 3, \ldots, Q$

Iterate:

1. Compute $s(n)$ for $n = 1, 2, \ldots, N$
2. Solve $w_0, w_1, \ldots, w_{M-1}$ so that $\sum_{n=1}^{N}(d(n) - \sum_{m=0}^{M-1} w_m s(n-m))^2$ is minimized.
3. Compute $u(n) = \sum_{m=0}^{M-1} w_m x(n-m)$
4. Solve $a_1, a_2, \ldots, a_Q$ so that $\sum_{n=1}^{N}(d(n) - \sum_{q=1}^{Q} a_q u^q(n))^2$ is minimized

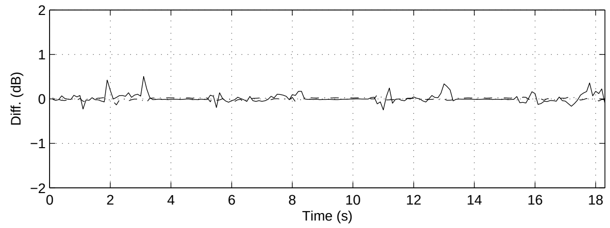

Figure 9: Difference between model where aliasing effects are not considered and the upsampling-downsampling approach in intermediate (dashdot) and high (solid) volume level.

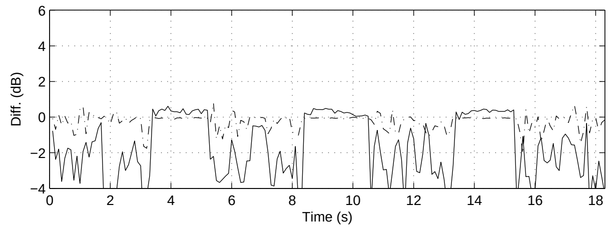

Figure 10: Difference between a static model where aliasing effects are not compensated and the lowpass filtering approach in intermediate (dashdot) and high (solid) volume level.

processor and filter parameters are one kind of time averages over the whole signal. However, in our measurements echo path is quite time invariant and therefore this kind of approach is justified.

In Figure 9 we see that with this nonadaptive model the advantage achievieved with up- and downsampling is very small.

If we use low-pass filtering to cut away frequencies that are aliased to the lower frequencies the result is similar to the case with adaptive algorithms as we can see from Figure 10. The performance loss is significant compared to the model where no compensation is made, although the performance is still better than with non-adaptive linear filter, as shown in Figure 11.

4. CONCLUSIONS

In this paper we studied the effect of aliased higher harmonic frequencies to simulations if they are not compensated. The harmonic frequencies are generated by a polynomial preprocessor. In our simulations we could achieve at best 1 – 2 dB improvement in ERLE, while computational complexity increased due to upsampling and downsampling. Thus the aliasing effect does not severely harm the performance of polynomial filters in this application and therefore increase in computational complexity might not be justified. However, the aliasing effects is something to be considered when trying to reach the best possible echo attenuation.

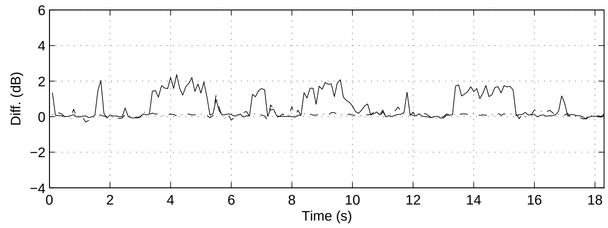

Figure 11: Difference between a fixed linear model and lowpass filtering approach in intermediate (dashdot) and high (solid) volume level.

We also tried to simply cut such frequencies that are aliased instead of computationally consuming upsampling and downsampling, but this approach did not show any improvement compared to linear filters in terms of echo attenuation. Thus, finding a method that is computationally less demanding can be a topic of future study.

5. REFERENCES

[1] R. Niemistö and T. Mäkelä, “Robust adaptive polynomial filters for acoustic echo cancellation,” in *5th IEEE Nordic Signal Processing Symposium, NORSIG*, October 2002.

[2] R. Niemistö, T. Mäkelä, and V. Myllylä, “Robust fast affine projection algorithm for nonlinear acoustic echo cancellation,” in *XI European Signal Processing Conference, EUSIPCO*, September 2002.

[3] Alexander Stenger and Rudolf Rabenstein, “An acoustic echo canceller with compensation of nonlinearities,” in *IX European Signal Processing Conference, EUSIPCO-98*, September 1998.

[4] Alexander Stenger and Walter Kellermann, “Adaptation of a memoryless preprocessor for nonlinear acoustic echo cancelling,” *Signal Processing*, vol. 80, pp. 1747–1760, 2000.

[5] V. John Mathews and Giovanni L. Sicuranza, *Polynomial Signal Processing*, John Wiley & Sons, Inc., 605 Third Avenue, New York, 2000.