Active Noise Control of Speech in Headphones using Linear Prediction

Mikkel Krogh Simonsen, Kasper Kiis Jensen, Maxime Démurger,
Oliver Palmhøj Jokumsen, Christian Claumarch
Aalborg University, Acoustics and Audio Technology - 1st Semester

Active Noise Control (ANC) is a widely used technique. However most solutions are targeted at attenuating static noise with limited bandwidth. This paper focusses on using prediction to improve performance for non stationary signals. The solution proposed uses an FIR-filter adapted by an Filtered-x Least Mean Square (FXLMS)-algorithm combined with Linear Prediction (LP) to attenuate a speech signal. This solution uses the quasi-periodic properties of speech to estimate the upcoming content of a speech signal. Thereby decreasing the sample delay in the system which increases bandwidth, yielding better attenuation.

1 Introduction

Active Noise Control (ANC) is a used method in a wide variety of applications. Research is being done in creating silents zones[1], improving room acoustics using multiple loudspeakers[2] and is already available in various headphones.

Noise canceling headphones are used in various environments which have different noise characteristics. These characteristics vary from periodic low frequency noise (X-Y Hz), e.g. from machinery and plane turbines[3], mid-range frequency noise (Y-Z Hz), e.g. speech[4], to high frequency noise (Z-Q Hz), e.g. X[5]. The high frequency noise are naturally attenuated by using a closed headphone cup[6] and static low frequency noise are attenuated by the present consumer ANC headphones[7]. Speech can be shown in figure 1, not to be attenuated by active noise cancellation in present market ANC headphones.

Here is going to be a graph of different headphones tested with non stationary signals (Speech)

Figure 1: It is clearly seen that speech is not attenuated by current headphones.

This paper will be focusing on attenuation of speech using an ANC-system. Speech will hence be seen as the primary noise source. The solution will be based on a digital feedforward system using the Filtered-X Least Mean Squares (FXLMS). The solution is chosen due to it being the optimal system for non stationary signals[8]. The general solutions in ANC are described thoroughly by Hansen et al. [9].

The problem of a feedforward system, shown in figure 2, is the dependency of noise having a larger propagation delay from the noise is measured (1) to a headphone loudspeaker(2) can produce a counter-phase signal in the desired point(3).

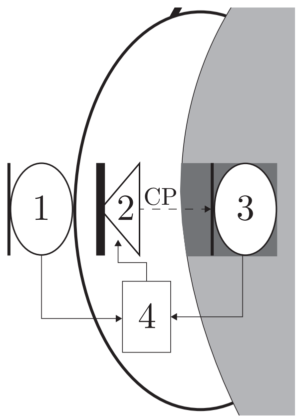

Figure 2: Simplified circumaural ANC Headphone fitted on ear. Showing the reference microphone (1), a headphone loudspeaker (2), an error microphone (3) and a DSP (4).

When implemented in a real-time system delay will occur, introduced by sampling and processing of the signal. For instance, measured on a cost-wise $\Sigma\Delta$-converter a sampling/conversion delay in the range of 220 to 880 μs is introduced, depending on a sampling frequency of 192 kHz to 48 kHz respectively. Assuming a 0° angle of incident, which is the ideal case, the reference microphone must be placed 75.5

mm to 302 mm further out from the error microphone, to compensate for the delay. In order to compensate for the introduced delay a Linear Prediction (LP)-scheme is proposed as a solution. In order to do prediction, certain signal characteristics must be known. Prediction of speech is described by Wai C. Chu [10].

The paper is split into two parts. The first part describes the feedforward FXLMS system and how to predict speech using linear prediction. The second part describes simulations and the implementation. Performance of the feedforward FXLMS system with and without prediction are determined and compared.

2 Methods

Using Figure 2 as the outline, the DSP(4) is expanded into Figure 3.

2.1 Feedforward ANC using FXLMS

This consists of converters and a control filter which is a FIR-filter where the coefficients are adapted by the FXLMS-algorithm.

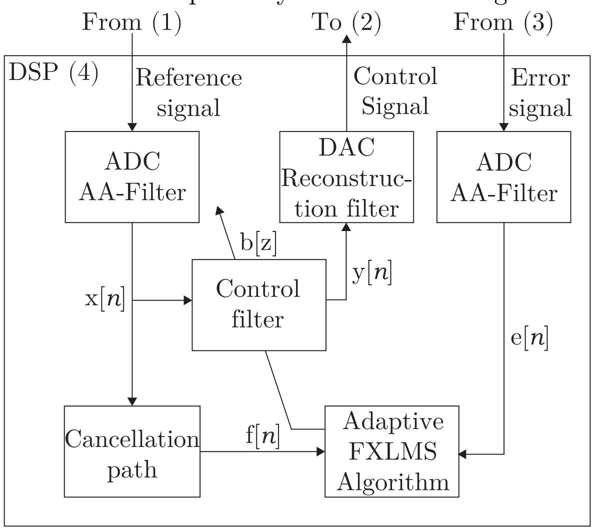

Figure 3: Adaptive feedforward ANC system.

Control Filter is the filter representing the transfer function from the reference microphone to the headphone loudspeaker. The filter, shown in equation 1, is initialized with the inverse of the measured impulse response of the represented transfer function. An order of 256 taps is chosen based on subjective testing in simulation.

$$y[n] = \sum_{n=0}^{L-1} b_j[n]x[n-j] \tag{1}$$

Where:

$b_j[n]$ is the weight coefficients written as $b[n] = [b_0[n], b_1[n], \cdots, b_{L-1}[n]]^T$

FXLMS is the optimization algorithm which updates the control filter coefficients using the FXLMS method shown in Equation 2 and derived in **??**.

$$b_j[n+1] = b_j[n] - 2\mu e[n]f[n-j] \tag{2}$$

Where:

μ is the convergence factor

$e[n]$ is the error

$f[n]$ is the reference convolved with the Cancellation Path

Cancellation Path (CP) is the transfer-function from the headphone loudspeaker to the error microphone. In the literature [9] the CP is adaptively adjusted, but it is assumed constant in this setup because the headphone position is fixed.

When implementing the system, delays exist due to the anti-aliasing and reconstruction filters. The delays of the system exceeds the propagation time of sound from the reference microphone to the headphone loudspeaker resulting in poor performance. Therefore an LP-algorithm is proposed to predict future samples in order to decrease the effect of the time delays.

2.2 Characteristic of Speech

Speech can be split into two main classes, voiced and unvoiced. Voiced sounds are characterized by a strong periodicity, with the fundamental frequency referred to as the pitch frequency (50 - 500 Hz). Unvoiced sounds are characterized as random. Speech is a non stationary signal and can only be assumed wide sense stationary (WSS) for periods of 20 - 30 ms[11].

2.3 Linear Prediction of Speech

In order to predict future samples the auto correlation function (ACF), which is used in LP, of speech must be estimated in frames. The outlines of the prediction system is shown in figure 4.

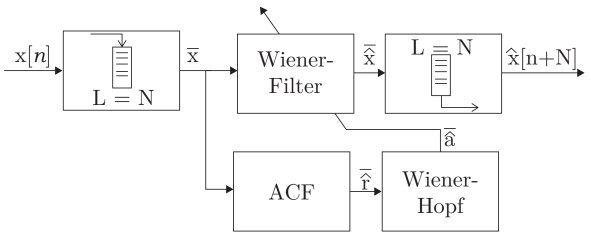

Figure 4: Linear prediction overview

Utilizing a nonrecursive estimation of the ACF, shown in Equation 3, it is possible to determine the linear prediction coefficients (LPCs) which are used for predicting future samples [12]. When estimating the ACF of the input (x) a Hamming window (w) is applied.

$$r_x[l] = \sum_{n=|l|}^{N} x_l[n] w_l[N-n] \tag{3}$$

Where:
$x_l[n] \ = x[n]x[n-l]$

$w_l[n] \ = w[n]w[n+l]$

l is the lag

m is the frame index

N is the frame size

The LPCs are determined using Equation 4, known as the Wiener-Hopf equation.

$$R \cdot \bar{a} = -\bar{r_x} \tag{4}$$

Where:
R is the covariance matrix C_{xx}

$\bar{a}$ is the LPCs, $\bar{a} = [a_1, a_2, \cdots, a_N]^T$

$\bar{r_x}$ is the ACF, $\bar{r_x} = [r_x[1], r_x[2], \cdots, r_x[N]]^T$

Equation 4 can be rewritten as shown in Equation 5 yielding the LPCs directly.

$$\bar{a} = -R^{-1} \cdot \bar{r_x} \tag{5}$$

Calculating R^{-1} is computationally heavy on a DSP. To estimate the LPCs the Levinson-Durbin method is used [13]. Prediction using wiener filtering, shown in equation 6, can then be applied to the current frame for prediction of the next frame.

$$\hat{x}[n+1] = -\sum_{i=1}^{N} a_i x[n-i] \tag{6}$$

Using equation 6 in cascade where $\hat{x}[n+2]$ is estimated using $\hat{x}[n+1]$ and $x[n]$. The predicted frame is then used as input for the ANC system.

Results

#404 Error..

- Prediction gain
- comparison of ref and predict ref
- comparison of ref and predict ref
- comparison of frequency response between:
  - no cancel
  - ref cancel
  - ref predict cancel

The system

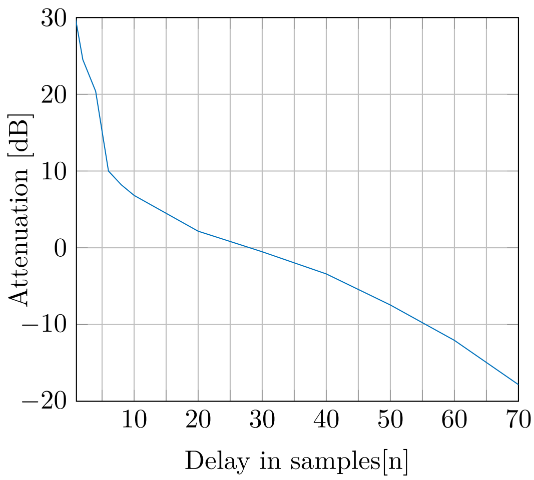

Figure 5: Attenuation achieved by the system for different system delays.

3 Dicussion

4 Conclusion

Bibliography

[1] I. T. A. . W. H. Abdulla, “Adaptive signal processing algorithms for creating spatial zones of quiet,” 2014.

[2] S. B. N. . A. Celestinos, “Active low frequency sound field control in a listening room using controlled acoustics bass system will also reduce the sound transmitted to neighbour rooms,” 2012.

[3] S. dude, “Source name,” at some time.

[4] S. dude, “Source name,” at some time.

[5] S. dude, “Source name,” at some time.

[6] S. dude, “Source name,” at some time.

[7] S. dude, “Source name,” at some time.

[8] C. Hansen, *Understanding Active Noise Cancellation.* Spon Press, 1st ed., 2003.

[9] D. M. . S. S. . X. Q. . L. B. . C. Hansen, *Active Control of Noise and Vibration.* CRC Press, 2nd ed., 2012.

[10] W. C. Chu, *Speech Coding Alorithms.* Wiley, 1st ed., 2013.

[11] S. dude, “Source name,” at some time.

[12] S. dude, “Source name,” at some time.

[13] S. dude, “Source name,” at some time.