Prediction Filter Design for Active Noise Cancellation Headphones

Markus Guldenschuh

Abstract—Digital active noise control (ANC) for headphones usually has to predict the noise because of the latency of AD-conversion. In adaptive feedback ANC, the prediction is based on the noise that entered the headphone. This noise is low pass filtered due to the physical barrier of the ear cups. In this paper, this low pass characteristic is exploited to define a prediction filter which does not require real-time updates. For broadband noises, the prediction filter performs better than adaptive prediction methods like the least mean squares (LMS) algorithm or iterated one-step predictions.

I. Introduction

Active noise control (ANC) systems sense, process and play-back noise such that it interferes destructively with the ambient noise. In headphones, ANC is a very efficient and powerful protection against loud sound levels. It cancels ambient noises directly at the user’s ear and reduces the need for loud music play-back.

There are analogue and digital realizations of ANC headphones, whereas digital applications have the advantage that the inherent filters can be adapted to changing conditions [1], [2]. However, the latency of conventional audio converters severely limits the ANC performance [3]. In these cases, the adaptive filter has to predict the noise to compensate for the delay. In the feedback ANC approach, the prediction is based on noise that actually entered the headphone. This has two advantages. Firstly, ANC is independent from the direction of incident noise and also works in diffuse sound-fields [4], [5]. And secondly, the upper frequencies of the entered noise are damped by the ear cup. This low-pass characteristic is advantageous when it comes to signal prediction [6].

In literature, prediction is mostly done by different kinds of the least mean squares (LMS) algorithm [7], [8], [9] or by iterated one-step-ahead predictions [10]. Both algorithms are based on sequential updates of the prediction filter and account for changes in the noise-signal characteristics. Since most environmental noises such as traffic noises are broadband, the signal characteristic of the penetrated noise mainly depends on the passive damping of the ear cups.

In this paper, we therefore suggest the design of a prediction filter that is only based on this passive damping characteristic. The proposed filter does not require runtime coefficient updates which makes its application simple, economical and robust, and still it yields equal - if not better - ANC results than the adaptive methods.

The structure of digital feedback ANC is reviewed in the following section and the prediction filter design is outlined in section II. Section IV shows simulation results and compares the proposed prediction with LMS and iterated one-step-ahead predictions before the conclusion and an outlook are given in the last section.

II. Digital Feedback ANC

Headphones with feedback ANC, as depicted in Fig.1 only require one microphone inside each ear cup. This microphone measures the residual error signal \( e(t) \) which is the superposition of the entered noise \( x(t) \) and the played-back anti-noise \( y(t) \). With omitted time dependency,

\[ e = x + y. \]

![Diagram of Digital Feedback ANC](image)

Figure 1: Digital feedback ANC: An estimate of the noise inside the ear cup \( \hat{x}(n) \) is used as input for the prediction unit. The inverted output is played back to cancel the entered noise \( x(t) \).

Since \( y \) should equal \( x \) with an inverted phase, an estimate of \( x \) is required. This estimate \( \hat{x} \) can be obtained by subtracting the played-back anti-noise from the sensed residual error, \( \hat{x} = e - \hat{y} \). The actual anti-noise \( y \) is digitally unavailable. This is because the digital noise cancellation signal is analogue converted and modified by the loudspeaker transfer-function \( S(z) \), also called secondary path. Thus, an estimate \( \hat{y} \) is generated by means of a secondary-path measure \( \hat{S}(z) \).

The resulting noise estimate \( \hat{x} \) is delayed by the digital to analogue converter (ADC) and the final anti-noise \( y \)
will be delayed once again by a DAC and the groupdelay of $S(s)$. To compensate for these delays, the prediction unit tries to predict future samples. If ADC and DAC each have a latency of $N$ samples, at least $2N$ samples have to be predicted. The group delay of $S(s)$ usually stays below 50 µs for most of the used bandwidth and can be neglected compared to the latency of conventional audio codecs.

### III. Noise Signal Prediction

Assuming that a correct estimate of $x$ is given, $\hat{x}(n)$ equals $x(n-N)$. With this, the prediction unit builds a weighted sum of the available past samples to predict the future noise sample

$$x(n+N) = \sum_{i=0}^{L-1} w_i x(n-N-i),$$  \hspace{1cm} (1)

with $L$ being the prediction order and $w_i$ the coefficients of a linear prediction filter. The residual error $\epsilon$ follows as

$$\epsilon(n) = x(n) - w^T S x(n-D),$$  \hspace{1cm} (2)

where $w$ is the vector of filter coefficients, $S$ is a convolution matrix of the secondary-path impulse response and $x$ is a signal vector starting from $D = 2N$ samples in the past. The minimum of the expected squared error leads to the well known Wiener-Hopf equation [11] and to the solution of the optimal prediction filter

$$w_{opt} = (SR_z S^T)^{-1} r,$$  \hspace{1cm} (3)

where $R_z$ is the autocorrelation matrix of the latest available noise samples

$$R_z = \begin{bmatrix} r_0 & r_1 & \ldots & r_{L-1} \\ r_1 & r_0 & \ldots & r_{L-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{L-1} & r_{L-2} & \ldots & r_0 \end{bmatrix}$$

and $r$ is a vector of autocorrelation elements starting from lag $D$

$$r = \begin{bmatrix} r_0 \\ r_{D+1} \\ \vdots \\ r_{D+L-1} \end{bmatrix}$$

The matrix inversion in eq. (3) can be avoided with the Levinson-Durbin algorithm [12], but only if $D = 1$. The resulting one-step-ahead predictor can still be used for a delay of $2N$ samples when the linear prediction of eq. (1) is iterated $2N$ times. However, since the prediction filter is stable, the recursive one-step-ahead prediction converges to zero. Thus, the filter outputs have to be amplified to get a reasonable multi-sample prediction.

To avoid this problem, gradient search algorithms are often used for prediction problems with more than one sample delay. Gradient search algorithms like the normalized LMS use the noise estimate and the error signal to recursively calculate the prediction filter

$$w(n+1) = w(n) + \mu \frac{\epsilon(n) \hat{x}}{\hat{x}^T \hat{x}},$$  \hspace{1cm} (4)

where $\mu$ is the step size parameter which determines the speed of convergence. In the LMS algorithm, the convergence speed has to be held rather low to keep the recursion stable.

However, the steady update of the coefficients is unnecessary if the main noise characteristic always stays the same. And in our ANC application, this is the case because the upper frequencies of the outside noise $u$ (as in Fig. 1) will always be damped by the physical barrier $P(s)$ of the ear cup. This passive attenuation can be written down as a convolution operation

$$x = Pu,$$

where $P$ is a convolution matrix of a low-pass impulse response that simulates the passive attenuation. With this, the autocorrelation matrix $R_x$ follows to

$$R_x = PR_u P^T.$$  \hspace{1cm} (5)

When $u$ has a flat spectrum, its autocorrelation matrix $R_u$ reduces to an identity matrix and $R_x$ solely depends on $P$. Thus for broadband noises, the calculation of the optimal prediction filter (eq. (3)) simplifies to

$$w_{opt} = (SPP^T S^T)^{-1} p,$$  \hspace{1cm} (6)

where $p$ is column number $D$ of the low-pass autocorrelation matrix $PP^T$. This equation allows for an a priori filter design where no real-time calculation of the filter coefficients is needed because all required data ($S(\omega)$ and $P(\omega)$) can be measured and designed in advance. The prediction filter that follows from this a priori calculation can be used as a fixed prediction filter in the ANC headphone.

### IV. Simulation Results

The following simulations are based on measurement data from a prototype headphone. Its passive damping and the low-pass filter that is used to derive the prediction filter are displayed in Fig. 2. The pass-band of the low-pass filter is chosen to be narrower than the one of the ear-cup damping. On the one hand, this increases the prediction error for the upper frequencies, but on the other hand it leads to better prediction in the low frequency band.

In order to reduce computational load, the DSP might be sampled down to 24kHz. Lower sampling frequencies are not advisable because they would reduce the spectral information of the passive attenuation. Common audio codecs have a talk-through latency of about 170µs at 192kHz. At 24kHz this corresponds to a total delay of approximately 4 samples.

In the first simulation, the prediction filter is used for aeroplane noise and its ANC performance is compared
with adaptive methods like LMS and iterated one-step linear prediction. Fig. 3 shows the spectrum of the aeroplane noise and the spectra of the residual noises after ANC.

Above 2000 Hz, the LMS algorithm and the iterated one-step-ahead prediction produce a lower error than the proposed prediction filter, but the ANC performance of the proposed filter between 100 and 1000 Hz is superior to the other two methods. As the passive noise attenuation above 1000 Hz is already very pronounced, the ANC of the fixed predictor is preferable.

In the second simulation (Fig. 4), the noise of an accelerating engine is used for the ANC comparison. The fixed prediction filter still leads to a slightly better performance, although it is designed a priori and does not change during runtime.

For this narrowband noise, an update of the prediction filter would be required to improve the ANC. This update can be done by inverting the autocorrelation matrix as in eq. (3). For the update in Fig. 5 the first 150 available noise samples are taken to perform an autocorrelation. From this autocorrelation, only four samples (beginning form time lag 0) are taken for the autocorrelation matrix $R_x$ and four further samples, starting from lag $2D = 4$, are taken for $r$. Thus an inversion of a $4 \times 4$ matrix is required to calculate the prediction filter coefficients. This updated $4^{th}$ order prediction filter again performs better.
than the iterated one-step-ahead prediction and the LMS prediction.

If narrowband signals are expected, sparse updates via matrix inversion are therefore a strategy that does not consume too much processing power and which leads to better results than LMS or iterated one-step-ahead predictions.

V. Conclusion and Outlook

This paper presents a new predictive solution for the biggest problem in digital ANC headphones: The unavoidable delay of the secondary path and the AD/DA conversion. We show that a main part of the noise characteristic is determined by the passive attenuation due to the ear cups. With this information, we design an a priori prediction filter that does not need any real-time coefficient updates and is therefore very economical. To the knowledge of the authors, this elegant prediction solution, elaborated by [6, Jain & Ranganath], has never been suggested for the feedback ANC applications. Simulations show that this filter reaches better results than adaptive prediction methods like the LMS or an iterated one-step-ahead linear prediction for usual environmental noises such as aeroplane or engine noises. The proposed prediction filter might also be used in hybrid ANC systems as proposed in [13], [5], [14], where it is expected to lead to improved ANC results too.

References