IMPROVING ADAPTIVE FEEDBACK CANCELLATION IN HEARING AIDS USING AN AFFINE COMBINATION OF FILTERS

Henning Schepker¹, Linh T. T. Tran², Sven Nordholm², and Simon Doclo¹

¹ Signal Processing Group, Department of Medical Physics and Acoustics and Cluster of Excellence "Hearing4All", University of Oldenburg, Oldenburg, Germany
{henning.schepker,simon.doclo}@uni-oldenburg.de
² Faculty of Science and Engineering, Curtin University, Perth, Australia
t.tran57@postgrad.curtin.edu.au,S.Nordholm@curtin.edu.au

ABSTRACT

In adaptive feedback cancellation an adaptive filter is used to model the acoustic feedback path between the hearing aid loudspeaker and the microphone. An important parameter for adaptive filters is the step-size, providing a trade-off between fast convergence and low steady-state misalignment. In order to achieve both fast convergence as well as low steady-state misalignment, it has been proposed to use an affine combination scheme of two filters operating with different step-sizes. In this paper we apply such an affine combination scheme to the acoustic feedback cancellation problem in hearing aids. We show that for speech signals a time-domain affine combination scheme yields a biased solution. To reduce this bias we propose to use a partitioned-block frequency-domain affine combination scheme. Experimental results using measured acoustic feedback paths show that in terms of misalignment and added stable gain the proposed adaptive feedback cancellation system outperforms a system that only uses a single adaptive filter with either of the fixed step-sizes used for the affine combination scheme.

Index Terms— feedback cancellation, hearing aids, affine combination, PBFDFAF

1. INTRODUCTION

In recent years the number of hearing impaired persons supplied with open-fitting hearing aids has been steadily increasing. While largely alleviating problems related to the occlusion effect (e.g., the perception of one’s own voice), open-fitting hearing aids are especially prone to acoustic feedback most often perceived as howling. This requires robust and fast-adapting feedback cancellation algorithms.

Several solutions exist to reduce acoustic feedback (see, e.g., [1, 2] and references therein), where adaptive feedback cancellation (AFC) seems to be the one of the most promising approaches as it theoretically allows for perfect cancellation of the feedback signal. In AFC the acoustic feedback path, i.e., the impulse response (IR) between the hearing aid loudspeaker and the hearing aid microphone, is approximated using an adaptive filter. Commonly, the least mean squares (LMS) algorithm or the normalized LMS (NLMS) algorithm is used to estimate the IR of the acoustic feedback path. However, the estimated IR will usually be biased due to the correlation between the incoming signal and the feedback signal [3, 4, 5].

To reduce this bias, for speech signals the so-called prediction-error method (PEM) can be used [1, 4, 6]. In the PEM it is assumed that the speech signal can be modeled as a white gaussian noise sequence which is filtered with a time-varying vocal tract filter. The goal is then to simultaneously obtain an estimate of the time-varying vocal tract filter by means of linear prediction and use the prediction error to obtain an unbiased feedback path estimate.

Assuming that the conditions for an unbiased estimation in the closed-loop system are fulfilled [7], the choice of the step-size in the LMS and NLMS algorithm usually is a trade-off between slow convergence but low steady-state misalignment and fast convergence but a higher steady-state error [8, 9]. In order to achieve both fast convergence as well as low steady-state misalignment, several solutions have been proposed that use either variable step-sizes [10, 11] or adaptively combine the outputs of two filters with different step-sizes [12, 13, 14, 15, 16, 17]. Similarly, as variable step-size algorithms, the combination of two adaptive filters can be intuitively interpreted as changing the global step-size controlled by the output of the two filters. While both approaches have been successfully applied to acoustic echo cancellation, their application to acoustic feedback cancellation is more challenging due to the correlation between the loudspeaker signal and the incoming signal. To this end, mostly the use of variable step-size algorithms has been considered for AFC in hearing aids [18, 19, 20]. In this paper we propose to apply the combination of two adaptive filters to the problem of acoustic feedback cancellation in hearing aids, more specifically the affine combination as proposed in [14, 16]. We show that in case of correlation between the loudspeaker signal and the incoming signal the combination scheme will in general adapt to a biased solution. In order to reduce the influence of this bias, we propose to use the PEM together with the partitioned block frequency-domain adaptive filter (PBFDFAF) [6]. Simulation results using measured acoustic feedback paths show that the proposed AFC system outperforms a system that only uses either of the individual adaptive filters in terms of misalignment and added stable gain.

2. AFC FRAMEWORK

Consider the single-loudspeaker single-microphone AFC system depicted in Figure 1. The microphone signal $y[k]$ is the addition of the incoming signal $x[k]$ and the feedback signal $f[k]$, i.e.,

$$y[k] = f[k] + x[k].$$

The feedback signal $f[k]$ is the convolution of the acoustic feedback path $h[k]$ and the loudspeaker signal $u[k]$, i.e., $f[k] = h[k] * u[k]$. Assuming that $h[k]$ is a finite impulse response (FIR) filter of length...
$L_h$, $f[k]$ can be expressed as

$$f[k] = \mathbf{h}^T[k]\mathbf{u}[k]$$  \hspace{1cm} (2)

with $\mathbf{h}[k] = [h_0[k], h_1[k], \ldots, h_{L_h-1}[k]]^T$ and $\mathbf{u}[k] = [u[k], u[k-1], \ldots, u[k-L_h+1]]^T$. The filter $\mathbf{h}[k]$ can be represented as a polynomial transfer function in $q$, i.e., $H(q, k) = \mathbf{h}^T[k]q$ with $q = \left[q^{-1}, \ldots, q^{-L_h+1}\right]^T$. Hence, $f[k]$ can be represented as

$$f[k] = H(q, k)u[k].$$  \hspace{1cm} (3)

The so-called error signal $e[k]$ corresponds to an estimate of the incoming signal $x[k]$ and is computed as

$$e[k] = y[k] - \hat{H}(q, k)u[k]$$  \hspace{1cm} (4)

where $\hat{H}(q, k)$ is an estimate of $H(q, k)$. This estimate can be obtained, e.g., using the NLMS algorithm (cf. Section 3.1). The loudspeaker signal $u[k]$ is then computed by amplifying $e[k]$ using the (possibly time-varying) forward path gain $G(q, k)$, i.e.,

$$u[k] = G(q, k)e[k],$$  \hspace{1cm} (5)

where it is typically assumed that $G(q, k)$ contains a delay $d_G \geq 1$ [1, 4].

3. PROPOSED ADAPTIVE FEEDBACK CANCELLATION SYSTEM

An overview of the proposed novel AFC system is depicted in Fig. 2. In order to achieve both fast convergence as well as low steady-state misalignment, the system comprises two adaptive filters $\hat{H}_1(q, k)$ and $\hat{H}_2(q, k)$, operating on the same input signal $\mathbf{u}[k]$, where the first adaptive filter uses a large step-size and the second adaptive filter uses a small step-size. The affine combination aims at combining the estimated feedback signals $\hat{e}_1[k]$ and $\hat{e}_2[k]$ such that the squared error signal $\hat{e}^2[k]$ is minimized, theoretically showing universal potential, i.e., the affine combination always performs at least as good as the best single filter [14, 16]. In order to reduce the bias of the filter estimation, adaptive pre-whitening is performed [4] using the filter $\hat{A}(q, k)$, which is estimated from the error signal $e[k]$.

In the following we first present the time-domain implementation and theoretically show that the affine combination parameter is unbiased when $u[k]$ and $x[k]$ cannot be perfectly decorrelated. Second, we introduce the PBFDAF implementation that makes additional use of transform-domain processing to reduce the estimation bias.

3.1. Time-domain implementation

In this implementation both adaptive filters are updated in the time-domain, where each adaptive filter estimates the acoustic feedback path $H(q, k)$ using the modified NLMS algorithm [21], i.e.,

$$\hat{H}_i[k+1] = \hat{H}_i[k] + \frac{\eta_i}{p_i[k]}\hat{e}_i[k], \quad i = 1, 2$$  \hspace{1cm} (6)

with $\hat{e}_i[k] = \hat{e}[k] - \hat{f}_i[k]$ the error signal of the $i$th filter, $p_i[k] = \alpha_h p_i[k-1] + (1 - \alpha_h) (\hat{e}_i^2[k] + \hat{e}^2[k])$ a power normalization factor with $\alpha_h$ a smoothing constant, and $\hat{H}_i[k] = [\hat{h}_{i,0}[k], \hat{h}_{i,1}[k], \ldots, \hat{h}_{i,L_h-1}[k]]^T$ the estimated filter vector of length $L_h$.

The outputs of both adaptive filters $\hat{f}_i[k] = \hat{H}_i^T[k]\hat{u}[k]$ are then combined using an affine combination scheme [16] to obtain the estimated feedback signal $\hat{f}[k]$, i.e.,

$$\hat{f}[k] = \eta[k]\hat{f}_1[k] + (1 - \eta[k])\hat{f}_2[k]$$  \hspace{1cm} (7)

$$= \eta[k]\hat{H}_1^T[k]\hat{u}[k] + (1 - \eta[k])\hat{H}_2^T[k]\hat{u}[k],$$  \hspace{1cm} (8)

where $\eta[k]$ us a (real-valued) adaptive combination parameter and $\hat{H}_i[k]$ is the estimate of $H_i[k]$ obtained by combining both adaptive filters. Aiming to minimize the squared error signal

$$E\{\hat{e}^2[k]\} = E\{(\hat{e}[k] - \hat{f}[k])^2\},$$  \hspace{1cm} (9)

with $E\{\cdot\}$ denoting expectations and $\hat{f}[k]$ defined in (7), the combination parameter $\eta[k]$ can be updated using a gradient-descent rule, e.g., an LMS-based update rule

$$\eta[k+1] = \eta[k] + \mu_e (\hat{f}_1[k] - \hat{f}_2[k])\hat{e}[k],$$  \hspace{1cm} (10)

with $\mu_e$ a positive step-size parameter. In [16] it has been shown that an improved performance for the affine combination can be achieved when an NLMS-based update rule or a sign-regressor LMS (SR-LMS)-based update rule is used. Therefore, in the proposed AFC system we use the SR-LMS algorithm to update $\eta[k]$, i.e.,

$$\eta[k+1] = \eta[k] + \mu_s \text{sgn}((\hat{f}_1[k] - \hat{f}_2[k]))\hat{e}[k]$$  \hspace{1cm} (11)

In order to avoid instability and following [14], $\eta[k+1]$ is restricted to be smaller than or equal to 1. The optimal solution $\eta^{\text{opt}}[k]$ is obtained by setting the gradient of $\hat{e}^2[k]$ in (9) with respect to $\eta[k]$ to zero, yielding

$$\eta^{\text{opt}}[k] = \frac{(f_1[k] - f_2[k])\Delta f_2[k]}{(\hat{e}_1[k] - \hat{e}_2[k])^2} + \frac{(f_1[k] - f_2[k])\hat{e}[k]}{(\hat{e}_1[k] - \hat{e}_2[k])^2},$$  \hspace{1cm} (12)
where we define $\Delta \tilde{f}_{2}[k] = \tilde{f}[k] - \tilde{f}_{2}[k]$ with $\tilde{f}[k] = \hat{A}(q,k)f[k]$ the prefiltered version of $f[k]$ defined in (3) which depend on the input signal $\tilde{u}[k]$ and used $\hat{e}[k] = \eta[k](\tilde{e}_1[k] - \tilde{e}_2[k]) + \tilde{e}_2[k]$, which can be obtained from (7). Hence, in general, for the considered application of AFC the optimal solution of the combination parameter $\eta[k]$ will be biased if the PEM does not perfectly decorrelate $\tilde{x}[k]$ and $\tilde{u}[k]$, e.g., for speech signals.

3.2. Partitioned Block Frequency-Domain implementation

As will be shown in the experimental evaluation (cf. Section 4), when using speech signals, the solution of the combination parameter $\eta[k]$ in the time-domain implementation is still biased even if the PEM is applied since for speech signals the PEM is not able to perfectly decorrelate the loudspeaker signal from the incoming signal. Therefore, in this section we present an PBFDAF-based implementation, which makes use of transform-domain processing to decorrelate the loudspeaker signal from the incoming signal in addition to the PEM. While the affine combination of filters has already been derived for block and partitioned block filters [15] and block frequency-domain filter [17], here we extend this approach to the PBFDAF framework.

In PBFDAF [1, 6, 22] the $i$th adaptive filter $\hat{H}_i(q,k)$ is partitioned into $L_i/P$ partitions of length $P$ each, i.e., $\hat{H}_{i,p}[k] = [\hat{h}_{i,p}[k] \hat{h}_{i,p+1}[k] \ldots \hat{h}_{i,(p+1)-1}[k]]^T$, $p = 0, \ldots, L_i/P - 1$, and transformed to the frequency-domain using an $M$-point DFT matrix $\mathcal{F}$, i.e.,

$$\hat{H}_{i,p}[k] = \mathcal{F} \begin{bmatrix} \hat{h}_{i,p}[k] \\ \vdots \\ \hat{h}_{i,(p+1)-1}[k] \end{bmatrix}. \quad (13)$$

For each partition the $i$th adaptive filter is then updated as

$$\hat{H}_{i,p}[l + 1] = \hat{H}_{i,p}[l] + \mathcal{F} C \mathcal{F}^{-1} \Delta_i[l] \tilde{U}_{i,p}[l] \tilde{E}_i[l]. \quad (14)$$

The partitioned filter input $\tilde{U}_{i,p}[l]$ is computed as

$$\tilde{U}_{i,p}[l] = \text{diag} \begin{bmatrix} u_i[l(l+1)L - pP - M + 1] \\ \vdots \\ u_i[l(l+1)L - pP] \end{bmatrix}. \quad (15)$$

with $L$ the block length. The error signal $\tilde{E}_i[l]$ in (14) is computed as

$$\tilde{E}_i[l] = \mathcal{F} \begin{bmatrix} 0 \\ 1 \end{bmatrix} (\tilde{y}[l] - \tilde{f}_i[l]), \quad (16)$$

with $\tilde{y}[l] = [\tilde{y}_i(L + 1) \ldots \tilde{y}_i(L + 1)L]^T$ and $\tilde{f}_i[l] = [\tilde{f}_i(L + 1) \ldots \tilde{f}_i(L + 1)L]^T$, computed as

$$\tilde{f}_i[l] = [0 1 \mathcal{F}^{-1} \sum_{p=0}^{L_i/P-1} \tilde{U}_{i,p}[l] \hat{H}_{i,p}[l]]. \quad (17)$$

The frequency-dependent step-size matrix $\Delta_i[l]$ in (14) is equal to

$$\Delta_i[l] = \text{diag} \begin{bmatrix} \mu_{i,0}[l] \\ \vdots \\ \mu_{i,M-1}[l] \end{bmatrix}. \quad (18)$$

with

$$\mu_{i,m}[l] = \frac{\mu_i}{|\tilde{E}_{i,m}[l]|^2 + \sum_{p=0}^{L_i/P-1} |\tilde{U}_{i,p,m}[l]|^2 + \delta} \quad (19)$$

and $\delta$ is a small positive constant. In order to avoid circular convolution effects, the matrix $C$ is used in (14) to constrain the gradient [6, 22].

Using a partition- and frequency-dependent combination parameter $\eta_p[l]$, the affine combination of the filters for the $p$-th partition is equal to

$$\hat{F}_p[l] = \eta_p[l] \hat{F}_{1,p}[l] + (1 - \eta_p[l]) \hat{F}_{2,p}[l] \quad (20)$$

with $\eta_p[l] = \text{diag} \{ \eta_{p,0}[l], \ldots, \eta_{p,M-1}[l] \}$. The time-domain representation of $\hat{F}[l]$ is then computed as

$$\hat{f}[l] = [0 1 \mathcal{F}^{-1} \sum_{p=0}^{L_i/P-1} \hat{F}_p[l]]. \quad (21)$$

The error signal is then computed using the combined filter output and the microphone signal, i.e.,

$$\tilde{E}[l] = \mathcal{F} \begin{bmatrix} 0 \\ 1 \end{bmatrix} (\tilde{y}[l] - \hat{f}[l]), \quad (22)$$

Assuming that PBFDAF-based processing provides sufficient independence between frequency-bands, we use a frequency- and partition-dependent update rule to compute the combination parameter. Similarly as for the time-domain implementation, we use an SR-LMS based update rule and restrict the combination parameter to be real-valued, i.e.,

$$\eta_{p,m}[l + 1] = \eta_{p,m}[l] + \mu_p \text{sgn} \{ \Re \{ \hat{F}_{1,p,m}[l] - \hat{F}_{2,p,m}[l] \} \} \Re \{ \tilde{E}[l] \}. \quad (23)$$

where $\Re \{ \cdot \}$ denotes the real value of a complex number and $m = 0, \ldots, M - 1$ denotes the frequency index. Similar to the time-domain implementation $\eta_{p,m}[l + 1]$ is restricted to be smaller or equal to 1.

4. EVALUATION

In this section the time- and the frequency-domain implementations of the proposed AFC system using the affine combination of two adaptive filters is evaluated. Acoustic feedback paths were measured on a dummy head with adjustable ear canals [23] using a two-microphone behind-the-ear hearing aid and open-fitting ear molds.
b) with PEM

-20
-10
0  
ǫ / dB

a) without PEM b) with PEM µ1
µ2
comb.
0 10 20 30 40 50 60 70 80
-0.5 
-0.25
0  

that the combined filter is able to outperform each individual
filter in terms of misalignment and added stable gain.

The performance was evaluated for two different incoming sig-
als \(x[k]\): 1) a stationary speech-shaped noise (SSSN) and 2) a
speech signal consisting of female and male speech used in [23].
These signals allow to evaluate the proposed AFC system under
the following conditions: 1) the incoming signal and the loudspeaker
signal can be perfectly decorrelated by the PEM, i.e., for SSSN, and
2) the signals can only be approximately decorrelated by the PEM,
for speech. All signals were 80s long and an instantaneous
change of the acoustic feedback path was simulated after 40s by
switching from the IR measured in free-field to the IR measured
with the telephone receiver.

As instrumental measures, the normalized misalignment \(\varepsilon\)
and the added stable gain \(ASG\) were used. The normalized misalign-
ment is defined as

\[
\varepsilon = 10 \log_{10} \frac{\| h - \hat{h}\|_2}{\| h \|_2},
\]

while the added stable gain is defined as [4, 26]

\[
ASG = 10 \log_{10} \frac{1}{\max_\Omega |H(e^{j\Omega}) - \hat{H}(e^{j\Omega})|^2} - 10 \log_{10} \frac{1}{\max_\Omega |\hat{H}(e^{j\Omega})|^2},
\]

with \(H(e^{j\Omega})\) and \(\hat{H}(e^{j\Omega})\) the frequency responses of the measured
and the estimated acoustic feedback paths at normalized frequency
\(\Omega\), respectively.

The following settings were used in all simulations. The forward
path gain of the hearing aid was set to \(G(q, k) = 10^{-d_{\Omega}}\) with \(d_{\Omega}\)
corresponding to a delay of 6 ms. For the time-domain implementa-
tion we used \(L_h = 64\), \(\alpha_h = 0.992\) and \(\mu_h = 1\) and for the SSSN
we chose \(\mu_1 = 0.02\) and \(\mu_2 = 0.0004\), while for the speech signal
we chose \(\mu_1 = 0.002\) and \(\mu_2 = 0.0004\). For the frequency-domain implementa-
tion we used \(L_h = 64\), \(L = 32\), \(P = 32\), \(M = 128\),
\(\mu_1 = 2\), \(\mu_1 = 0.015\) and \(\mu_2 = 0.001\). For both approaches the
prediction-error filter \(\hat{A}(q, k)\) was of order 20 and was updated ev-
ery 10 ms using the Levinson-Durbin recursion.

Figure 4 shows the results for the SSSN using the time-domain
implementation. The left column depicts the normalized misalign-
ment and the affine combination parameter \(\eta[k]\), when the PEM is
not used. As expected from (12), the time-domain implementa-
tion is not able to track the best filter in case of correlation between \(x[k]\)
and \(u[k]\), i.e., it follows only the fast filter and \(\eta[k] \approx 1\) most of the
time. However, if the PEM is used (right column) the affine com-
bination scheme is well able to track the best filter (i.e., initially the
fast filter and after a while the slow filter) and even outperforms
the best filter in some time instances.

Figure 5 shows the results for the speech signal using both the
time-domain implementation (left column) and the PBFDASF imple-
mentation (right column) both using the PEM for both instrumen-
tal measures. While for the time-domain implementation the affine
combination is not able to track the best filter, for the PBFDASF im-
plementation the affine combination is able to track the best filter
and even outperforms the fast filter when the slow filter has not yet con-
verged. This is especially visible between 30-40s, where the ASG
(cf. Figure 5d) can be increased by about 3dB for the affine combi-
nation. This indicates that the additional decorrelation achieved by
the transform-domain processing allows the affine combination to
track the best filter. Additionally, a less fluctuating ASG over time
is achieved for the affine combination compared to the fast filter. These
results show the benefit of using the affine combination of two adap-
tive filters compared to using only a single adaptive filter with a fixed
step-size.

5. CONCLUSION

In this paper we have proposed a novel AFC system that uses the
affine combination of two adaptive filters with different step-sizes
in order to yield a fast convergence and low steady-state misalign-
ment of the combined filter. We have theoretically shown, that for
adaptive feedback cancellation in hearing aids the affine combina-
tion is biased when no decorrelation is applied to the loudspeaker
and the incoming signals. Simulation results using measured acoustic
feedback paths show that for speech signals the time-domain imple-
mentation of the proposed AFC system is not able to track the best
filter even when the PEM is used to decorrelate the signals. How-
ever, using the PBFDASF implementation to additionally benefit from
the decorrelating properties of transform-domain processing we have
shown that the combined filter is able to outperform each individual
filter in terms of misalignment and added stable gain.

234
6. REFERENCES


