Automatic Tuning of Probe Noise for Continuous Acoustic Feedback Cancelation in Hearing Aids

Muhammad Tahir AKHTAR* and Akinori NISHIHARA†

*Department of Electrical Engineering, COMSATS Institute of Information Technology, Park Road, Chak Shahzad, Islamabad 44000, Pakistan. Emails: tahir.akhtar@comsats.edu.pk, akhtar@ieee.org
†Department of Human System Science, Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Oakayama, Meguro-ku, Tokyo, JAPAN. Email:aki@cradle.titech.ac.jp

Abstract—In this paper, we propose a probe signal-based method for continuous acoustic adaptive feedback cancelation (AFC) in digital hearing aids. The main idea is to incorporate a time varying gain with the probe signal, such that a high level probe noise is injected during the transient state, and a low level probe signal is used after the system has converged. The proposed method is essentially based on two adaptive filters working in tandem. The weights between these two adaptive filters are exchanged by an efficient weight-transfer strategy, such that both adaptive filters give a good estimate of the acoustic feedback path. Simulation results demonstrate that the proposed method achieves good modeling accuracy, preserves good speech quality, and maintains high output SNR at the steady-state.

Keywords—Hearing aids, acoustic feedback, NLMS algorithm, probe noise.

I. INTRODUCTION

Acoustic feedback is a major problem in the hearing aids, limiting the maximum gain available to the user, and making the hearing aids oscillate at higher gain thus producing annoying sounds of whistling, screeching or howling. A generic digital hearing aid system is shown in Fig. 1, where \( G(z) \) represents the forward path of the hearing aid and comprises all signal processing for noise reduction and signal amplification, and \( F(z) \) denotes the acoustic feedback path. Here \( s(n) \) is the desired input signal to be processed by \( G(z) \), and \( y_f(n) \) is the feedback due to the output signal \( y(n) \). Assuming that the components for the adaptive feedback cancelation (AFC) (shown in a dashed box) are not present, and hence, \( u(n) = x(n) \); the closed loop transfer function between \( y(n) \) and \( s(n) \) is given as:

\[
H(z) = \frac{G(z)}{1 - G(z)F(z)},
\]

which shows that due to the acoustic feedback the hearing aid will be unstable, if \( G(z) \) is large enough so that \( G(z)F(z) = 1 \) at some frequency.

A number of approaches have been proposed to solve the problem of acoustic feedback [11] – [6]. The most successful approach is based on adaptive filtering as shown in Fig. 1., where \( W(z) \) is employed to model \( F(z) \). The most famous adaptive algorithm to adapt \( W(z) \) is the normalized least mean square (NLMS) algorithm [7], which is due to its simplicity, and ease of implementation. The update equation for NLMS algorithm is given as:

\[
w(n+1) = w(n) + \mu \frac{y_f(n) y(n) + \delta}{y_f(n) y(n)} e(n),
\]

where \( \mu \) is the step-size parameter, \( w(n) \) is the coefficient vector for \( W(z) \), \( e(n) = x(n) - y_w(n) \) is the error signal, \( y(n) \) is the vector for the output signal \( y(n) \), and \( \delta \) is a small positive constant to avoid division by zero. Ideally, \( W(z) \) is expected to generate a replica of \( y_f(n) \), so that \( u(n) = e(n) \approx s(n) \). However, the input and the desired-response signals of \( W(z) \), \( y(n) \) and \( x(n) \), respectively, are correlated with each other and would result in a biased convergence [8]. A simple approach to perform decorrelation is to use an appropriate delay either in the cancelation path [1], or in the forward path [9]; however, it degrades the speech quality. Another solution is to filter the error and/or input signal of \( W(z) \), through appropriate decorrelation filters, before being used in the update equation of the NLMS algorithm [10], resulting in the so-called Filtered-x adaptive algorithm [11]. A filter bank-based frequency-domain technique has been investigated for AFC [12]. The frequency-domain techniques, however, result in an increased computational load and require a lot of battery power [2]. For time-domain continuous AFC, a dual microphone-based solution has been proposed where two microphones are used to pick the input signal and dual adaptive filters are employed to perform AFC [13], [14]. Those techniques have obvious physical and computational limitations.

Another solution is a noncontinuous adaptation, or an open-loop algorithm in which the hearing aid forward path is broken and a probe signal is injected during particular intervals, for example, when howling is detected by an appropriate oscillation detector [15]. The ON/OFF switching of the probe signal produces annoying effects to the hearing aid user. A continuous injection of probe noise has been considered, however, either the level of the probe noise must be kept low to have an appreciable signal-to-noise ratio (SNR) [16], or an appropriate masking filter be introduced to perceptually mask the probe signal [17]. Previously, we have proposed a method comprising two adaptive filters \( W_1(z) \) and \( W_2(z) \) [18], where \( W_1(z) \) is the same as in the conventional approach (Fig. 1), and \( W_2(z) \) is excited by a probe signal. A delay is inserted in the path for probe signal, which allows implementing delay-based adaptive filtering [19] to track convergence-status of \( W_2(z) \). The level of probe noise must be kept low, which affects the convergence speed. Furthermore, there is no check on the biased convergence of the first adaptive filter. In this paper, we attempt to solve these problems. Essentially: 1) we consider a two adaptive filter-based structure as in the previous method [18], 2) an appropriate delay is inserted in the forward path of the hearing aid, and hence delay-based adaptive filtering is used to adapt the two adaptive filters, 3) an efficient strategy is developed to transfer the weights between the two adaptive filters such that both adaptive filters give a good estimate of \( F(z) \), 4) the problem of biased convergence is mitigated by freezing the adaptation once a good solution is obtained, and finally, 5) a time-varying gain is
proposed to control the level of added probe noise: a large value is used at the start-up for a fast convergence, and gain is reduced to a small value as the system converges thus achieving appreciable SNR at the steady-state. Up to the best knowledge of authors, automatic tuning of the probe noise has not been considered in the existing literature on continuous AFC in hearing aids.

Notation: For sake of convenience, we have used mixed notation in the block diagram, where transfer function is expressed in $z$-domain, and input-output signals are in discrete-time domain. For example, $y_{w_1}(n) = W_1(z)y(n)$ would represent filtering of $y(n)$ via $W_1(z)$. In the algorithm description, on the other hand, filtering has been represented by inner product as $y_{w_1}(n) = \mathbf{w}_1^T(n)y(n)$, where $\mathbf{w}_1(n)$ represents the coefficient vector of $W_1(z)$, and $y(n)$ denotes the corresponding signal vector.

II. PROPOSED METHOD

The block diagram of the proposed method is shown in Fig. 2, where AFC is achieved by two adaptive filters, $W_1(z)$ and $W_2(z)$, working in tandem. The adaptive filter $W_1(z)$ is excited by $y(n)$, and is expected to provide a neutralization signal for $y_f(n)$. The second adaptive filter $W_2(z)$ is excited by the probe signal $v(n)$, and is expected to provide a neutralization signal for the feedback component $v_f(n)$ due to the added probe signal $v(n)$. The probe signal $v(n)$ is generated from a white Gaussian noise $v_0(n)$ with a time-varying gain $\alpha(n)$. It is assumed that $v_0(n)$, and hence $v(n)$, is uncorrelated with the input signal $s(n)$ and hence with the output signal $y(n)$. The signal picked up by the input microphone, $x(n)$, is now given as:

$$x(n) = s(n) + y_f(n) + v_f(n),$$  

(3)

where $y_f(n) = f(n) + y(n) - D$ and $v_f(n) = f(n) + v(n) - D$ are the acoustic feedback components, and where $D$ is an appropriately selected delay and $*$ denotes the linear convolution. The error signal for $W_1(z)$, $g(n)$, is computed as:

$$g(n) = x(n) - y_{w_1}(n) = s(n) + [y_f(n) - y_{w_1}(n)] + v_f(n),$$  

(4)

which is used as the desired response for $W_2(z)$, and hence the error signal for $W_2(z)$, $e(n)$, is given as:

$$e(n) = g(n) - y_{w_2}(n) = s(n) + [y_f(n) - y_{w_1}(n)] + [v_f(n) - y_{w_2}(n)].$$  

(5)

A. Weight-Transfer Strategy

A delay-based technique has largely been applied in the field of acoustic echo cancelation [19]: An ‘appropriate’ delay $D$ is inserted in the signal flow path, and an ‘extended filter’ is used for system identification. The part of filter employed for modeling the delay is sure to be converged to zero. Since the NLMS algorithm spreads the error among the filter coefficients: the norm of extension coefficients can be used as an estimate for the filter mismatch. For hearing aids, traditionally such type of delay is used to solve the correlation problem in the AFC filter [9]. In our approach, the objective of the appended delay is two fold: 1) to provide (some) decorrelation, as well as 2) to help designing an efficient strategy for weight transfer between $W_1(z)$ and $W_2(z)$, as explained below.

Insertion of delay at the output of the hearing aid increases the effective path to be identified by the AFC filters $W_1(z)$ and $W_2(z)$.

1. It is important to note that setting $D$ too low would yield a poor estimator; however, the extension of the adaptive filter implies increased memory and complexity requirements: thus there is a tradeoff situation.

Thus both $W_1(z)$ and $W_2(z)$ are considered with extended-length coefficient vectors being given as:

$$\mathbf{w}_1(n) = \begin{bmatrix} \mathbf{w}_{11}(n) \\ \mathbf{w}_{1F}(n) \end{bmatrix}, \quad \text{and} \quad \mathbf{w}_2(n) = \begin{bmatrix} \mathbf{w}_{21}(n) \\ \mathbf{w}_{2F}(n) \end{bmatrix},$$  

(6)

where $\mathbf{w}_{11}(n) = [w_{11,0}, w_{11,1}, \cdots, w_{11,D-1}]^T$ and $\mathbf{w}_{21}(n) = [w_{21,0}, w_{21,1}, \cdots, w_{21,D-1}]^T$ represent the part used to model the delay, and both $\mathbf{w}_{1F}(n)$ and $\mathbf{w}_{2F}(n)$ model $F(z)$. The two adaptive filters are adapted using delay-based adaptive algorithms [19] as summarized in Table 1. At the start-up, convergence of the two adaptive filters $W_1(z)$ and $W_2(z)$ can be monitored on the basis of norm of extension coefficients modeling the appended delay as:

$$\rho_1(n) = \|\mathbf{w}_{11}(n)\|^2, \quad \text{and} \quad \rho_2(n) = \|\mathbf{w}_{21}(n)\|^2,$$  

(7)

where $\| \cdot \|$ denotes Euclidean norm. Both $\mathbf{w}_{11}(n)$ and $\mathbf{w}_{21}(n)$ are initialized with all 1’s and $\mathbf{w}_{1F}(n)$ and $\mathbf{w}_{2F}(n)$ may be initialized by null vectors of appropriate orders. The convergence of $W_1(z)$ is faster than $W_2(z)$ and initially $\rho_1(n) < \rho_2(n)$, and hence weights from $W_1(z)$ are copied to $W_2(z)$ as $\mathbf{w}_{1F}(n) \rightarrow \mathbf{w}_{2F}(n)$. As soon as $\rho_1(n)$ approaches a certain threshold $T_1$, and $W_2(z)$ starts converging too (which may be checked by condition $\rho_2(n) < 1$), the adaptation of $W_1(z)$ is frozen, and an estimate of $F(z)$ is obtained as:

$$\hat{f}(n) = \mathbf{w}_{1F}(n).$$  

(8)

Now the normalized squared deviation (NSD) for $W_2(z)$ may be computed as:

$$\Delta W_2(n) = 10 \log \left\{ \frac{\|\hat{f}(n) - \mathbf{w}_{2F}(n)\|^2}{\|\hat{f}(n)\|^2} \right\}.$$  

(9)

The following strategy is proposed to avoid a biased convergence problem, once an estimate of $F(z)$ is available:

1. If $W_1(z)$ is kept fixed and only $W_2(z)$ is adapted.
2. If $\rho_1(n)$ and $\rho_2(n)$ are computed as in (7) and if $\rho_1(n) < \rho_2(n)$ then weights from $W_1(z)$ are copied to $W_2(z)$ as $\mathbf{w}_{1F}(n) \rightarrow \mathbf{w}_{2F}(n)$ or if $\rho_2(n) < \rho_1(n)$ then $\mathbf{w}_{2F}(n) \rightarrow \mathbf{w}_{1F}(n)$ (thus $W_1(z)$ is in fact treated as a piece-wise fixed filter), and
3. If NSD $\Delta W_2(n)$ increases beyond certain predefined threshold $T_2$, i.e., $\Delta W_2(n) > T_2$, and the error signal $e(n)$ has also diverged (e.g., $P_e(n) > T_3$) then plant has changed significantly, and thus the two adaptive filters must be re-initialized to find a new estimate of the feedback path. Here $P_e(n)$ denotes power of the error signal $e(n)$, which

![Fig. 2. Block diagram of the proposed method for continuous AFC.](image-url)
Based on this observation, we propose to compute the time-varying gain for probe signal as:

\[ e(n) = g(n) - y_w(n) \]

\[ P_e(n) = \lambda P_e(n-1) + (1 - \lambda) e^2(n) \]

\[ \hat{N}_D(n) = \lambda \hat{N}_D(n-1) + (1 - \lambda) \frac{\mathbf{w}_2^T(n)\mathbf{w}_2(n)}{D} \]

\[ \mu_1(n) = \begin{cases} \frac{\hat{N}_D(n)}{P_e(n)} > \mu_{1\text{min}} \\ \mu_{1\text{min}} \end{cases} \]

\[ \mathbf{w}_2(n+1) = \mathbf{w}_2(n) + \frac{\mu_2(n)}{\mathbf{v}^T(n)\mathbf{v}(n) + \delta} e(n)\mathbf{v}(n) \]

where \( \lambda \) is the forgetting factor (0.9 < \( \lambda \) < 1).

### B. Time-Varying Gain for Probe Signal

In the proposed method (see Fig. 2), \( \alpha(n) \) is the time-varying gain for the probe signal \( v(n) \). Intuitively, we would like to use a high-level probe noise at the start-up (or when there is a change in the acoustic feedback path) so that convergence of \( W_2(z) \) is fast. After \( W_2(z) \) has converged, the gain for the probe signal \( v(n) \) must be reduced to have good output SNR at the steady-state. In (7), we have found a parameter \( \rho_2(n) \) which can be used to monitor the convergence status of \( W_2(z) \) as: from a large value at the start-up it convergence to a small value \( \mathbf{w}_2(n) \to \mathbf{0} \) (null vector) as \( n \to \infty \).

Based on this observation, we propose to compute the time-varying gain \( \alpha(n) \) as:

\[ \alpha(n) = \frac{\rho_2(n)}{\rho_2(n) + C} \]

where \( C \) is a positive constant, and finally the probe signal is computed as:

\[ v(n) = \alpha(n)v_0(n) \]

When \( \rho_2(n) \) is large, then \( v(n) \) tends to \( v_0(n) \). On the other hand, when \( \rho_2(n) \) is small, the gain \( \alpha(n) \) and hence the probe signal \( v(n) \) is small.

### C. A Few Remarks

1) Considering the situation when the two adaptive filters \( W_1(z) \) and \( W_2(z) \) give good estimate of \( F(z) \): \( y_{w1}(n) \to y_f(n), \ y_{w2}(n) \to y_f(n) \), and hence \( g(n) \to s(n) + v_f(n) \to e(n) \to s(n) \). Thus, \( u(n) = e(n) \) can be used as an input to the hearing aid \( G(z) \).

2) \( \mathbf{w}_2(n) \to \mathbf{0} \) (null vector) and hence \( T_3 \) can be selected as a small number close to zero. 3) Our experience shows that modeling accuracy of 20~30 dB is good enough for a stable operation of the hearing aid, and hence \( T_3 \) can be selected accordingly. 4) \( T_3 \) is a threshold for power of the error signal \( e(n) \). Since \( e(n) \) is used as an estimate for the reconstructed speech signal and is input to the hearing aid, the value for the threshold \( T_3 \) can be estimated from past values of \( P_e(n) \) computed in (10).

### III. SIMULATION RESULTS

In this section, we present results of the computer simulations. Fig. 3 shows the characteristics of the feedback path, \( F(z) \). All adaptive filters are assumed to be FIR filters of tap-weight length \( 64 \). The sampling frequency is \( F_s = 16 \) kHz. The forward path representing the hearing-aid processing unit, is assumed to be given as \( G(z) = Kz^{-\Delta} \), where \( K \) and \( \Delta \), respectively, represent the gain and delay of the system. The following methods are considered in this simulation study (where the simulation parameters are determined experimentally for fast and stable convergence). 1) NLMS-algorithm based conventional method (\( \mu_1 = 1 \times 10^{-3}, \ \delta_1 = 1 \times 10^{-4} \), 2) Previous method [18] (\( \mu_1 = 1 \times 10^{-3}, \ \mu_2 = 1 \times 10^{-4}, \ \delta_1 = 1 \times 10^{-4}, \ \lambda = 0.97, \ \text{SNR}_{\text{probe}} = \sigma^2_s/\sigma^2_e = -15 \) dB, \( D = 64, \ \mu_{2\text{min}} = 1 \times 10^{-6}, \ T_1 = 1 \times 10^{-3}, \ T_2 = -20 \) dB), and 3) Proposed method (\( D = 8, \ \mu_{1\text{min}} = 1 \times 10^{-6}, \ T_3 = 10, \ \text{SNR}_{\text{probe}} = \sigma^2_s/\sigma^2_e = 0 \) dB. The empirical constant \( C \) is set to 1.5, and the rest of the parameters are set to the same value as in Previous method.

Fig. 4 shows signals used in the computer simulations, where S1 to S8 denote speech signals. The signals denoted as S9 to S15 are with strong tonal characteristics, and are hereby used to evaluate the performance of hearing aid for entrainment. Entrainment is typically described as feedback after cessation of the sound, additional tones, warbling, or echoes [20]. It is a common artifact in hearing aids which occurs when the feedback cancelation algorithm erroneously attempts to cancel a tonal input to the hearing aid.

The simulation results for hearing aid model with \( K = 10 \) for signals S1 to S8 are presented in Fig. 5(a)~(d), where Fig. 5(a) shows error in reconstruction of the desired signal at the input of the hearing aid for the speech signal S1, being computed as \( \Delta S(n) = |s(n) - u(n)| \). It is obvious that for a perfect reconstruction, we must have \( \Delta S(n) \to 0 \). From Fig. 5(a), we see that the proposed method gives a very fast convergence speed in reproducing the desired signal at the input of hearing aid processing unit. In fact, we ‘hear’ some musical noise in the case of conventional and previous methods, whereas the proposed method produces no such musical noise. The variation of
We observe that the proposed method do no update signals S1 to S8 to compute the mean NSD. We observe that the comparison with the input probe signal $v(n)$ for MSG (in dBs). This avoids any fluctuations, due to the non-stationarity of the speech signal for example, which we do observe in the previous and obtained. This avoids any fluctuations, due to the non-stationarity of speech for the experiment with S1, which shows that the level of probe signal is reduced to a very low level as the system converges; this in turn must improve the SNR at the output of the hearing aid (as explained later). The corresponding curves for NSD of filter $W_1(z)$, $\Delta W_1(n)$, being defined as

$$
\Delta W_1(n) = 10 \log \left\{ \frac{\|f(n) - W_1F(n)\|^2}{\|f(n)\|^2} \right\},
$$

are shown in Fig. 5(c), where each curve is averaged for all speech signals S1 to S8 to compute the mean NSD. We observe that the proposed method do no update $W_1(z)$ once a good solution is obtained. This avoids any fluctuations, due to the non-stationarity of speech signal for example, which we do observe in the previous and conventional methods. Fig. 5(d) shows curves for maximum stable gain (MSG) as computed as [14]:

$$
\text{MSG} = 20 \log \left\{ \max_{\omega} \|F(\omega) - W_1F(\omega)\|^2 \right\},
$$

which is determined by the frequency where the mismatch between the actual and the estimated path is greatest. However, the system will only be unstable when the phase at that frequency equals a multiple of $2\pi$ [14]. From Fig. 5(d), we observe that the proposed method gives largest MSG as compared with the other methods considered in this paper.

The simulation results for signals S9 to S15 are presented in Fig. 6(a)-(d), where Fig. 6(a) shows error in reconstruction of the desired signal at the input of the hearing aid for the input signal S13, and Fig. 6(b) shows the corresponding variation in the probe signal in the proposed method. The mean NSD and MSG curves, averaged over all speech signals S9 to S15, are shown in Fig. 5(c) and (d), respectively. We have observed that the conventional and previous methods produces a lot of ringing and suffers from the entrainment artifact. Furthermore, the proposed method gives best NSD and MSG performance among the methods considered in this paper.

For qualitative assessment, the following performance measures are employed.

Perceptual Evaluation of Speech Quality (PESQ): It is an ITU-T standard to evaluate quality of speech signals [21]. The maximum score of 4.5 is for clean signal with no degradation.

Signal to Distortion Ratio (SDR): It is based on the Hilbert transform and measure levels of the nonlinear distortion in the processed signal in comparison with the original signal [22]–[24]. In the simulation model (see Figs. 1 and 2), the input signal $s(n)$ and the reconstructed signal $u(n)$ are taken as reference and test signals, respectively.

Mutual Information (MI): It is a non-parametric measure of relevance between two random variables $z_1$ and $z_2$, and can be interpreted using Kullback-Leibler divergence as [25]

$$
\text{MI} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(z_1, z_2) \log \left( \frac{f(z_1, z_2)}{f(z_1)f(z_2)} \right) dz_1 dz_2,
$$

where $f(z_1, z_2)$ denotes joint probability distribution function of $z_1$ and $z_2$, and $f(z_1)$ and $f(z_2)$ denote their respective marginal probability distribution functions (PDFs). MI is always non-negative and zero if and only if the two random variables are strictly independent. We compute MI between input $s(n)$ and the reconstructed signal $u(n)$, with the understanding that (MI $> 0$) indicates close relevance between these signals, and larger the MI better $u(n)$ resembles the input $s(n)$.

Normalized Mean Squared Error (NMSE): It is computed (in dB) as

$$
\text{NMSE} = 10 \log \left\{ \frac{1}{J} \sum_{j=1}^{J} \frac{\|s(j) - u(j)\|^2}{\sigma_y^2} \right\},
$$

where $J$ denotes the total number of samples used in the computation. Being ratio of two similar quantities, NMSE is a unit-less quantity, and $\text{NMSE} \to -\infty$ shows that the corresponding signal is reconstructed with the minimum error.

Averaged MSG: This is obtained by averaging the steady-state value of MSG across all speech signals considered in the experiment. Output SNR: For the previous and proposed methods, the output SNR may be computed as $	ext{SNR}_{\text{out}} = 10 \log \left\{ \frac{1}{J} \sum_{j=1}^{J} \frac{\|s(j) - \sigma_y^2\|^2}{\sigma_y^2} \right\}$, where $\sigma_y^2$ and $\sigma_v^2$ denote variances of signals output signal $y(n)$ and probe noise $v(n)$.

Table 2 summarizes the corresponding results averaged over all speech signals (from mid sample to the last value). The speech quality is severely degraded for both the NLMS and previous algorithms, as depicted by the low SDR values. The proposed method gives the best performance as compared with the rest of methods considered in this paper. Especially, the values for PESQ, MSG, and output SNR are substantially improved as compared with the other methods.

IV. CONCLUDING REMARKS AND FUTURE WORK

In this paper, the main idea is to vary the level of the probe signal so that a fast convergence as well as a good output SNR may be achieved. The simulation results show excellent performance.
of the proposed method, and it appears as a promising choice for practical hearing aids. It is worth mentioning that the proposed method, being comprising two adaptive filters, has an increased computational complexity as compared with the conventional method. This increased computational complexity is the price paid for an improved performance. A detailed computational complexity analysis is omitted for the sake of space. A theoretical analysis of the proposed method is a task for the future work.

**REFERENCES**


