34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Improving Misalignment for Feedback Path Estimation in Hearing Aid by Multiple Short-Time Noise Injections

Soudeh A. Khoubrouy, Issa M.S. Panahi, *Senior Member, IEEE*

Abstract— **Adaptive Feedback Cancellation (AFC) methods are used to find an FIR filter to cancel the negative effect of acoustic feedback between the loudspeaker and microphone of the hearing aid. Finding the AFC filter of appropriate order/length directly affects the performance and complexity of the system. In this paper, we use noise injection method to find the AFC filter estimating the feedback path model. We show that the optimum length which guarantees a good compromise between the quality and the complexity of the system may be smaller than the length of the actual feedback path model. However, in order to improve the performance of the system in terms of Misalignment criterion, we propose using multiple short-time noise injections and averaging method to find the best filter estimate of appropriate length.**

I. INTRODUCTION

The small size of current hearing aid devices allows some signal leakage from the loudspeaker to the microphone. This phenomenon, called acoustic feedback, degrades the performance of the hearing aids and causes some howling and whistling [1]. In order to cancel this irritating effect, a feedback compensator or an Adaptive Feedback Canceller (AFC) can be used. Fig. 1 is a case in which the feedback compensator, i.e. an FIR filter, sits in parallel with the feedback path. According to this figure, a zero-mean White Gaussian noise is injected to the loudspeaker for a short time and an estimate of the feedback path is calculated using auto-correlation of the loudspeaker signal and cross-correlation between the loudspeaker and microphone signals. The number of the data samples used (thus, the noise injection duration) defines the accuracy of the computed correlation lags, especially for the large lag values. This consequently affects the estimated AFC filter and its order selection and thus, the system performance. The larger the number of the data samples, the more accurate the correlation samples, and the higher filter order can be estimated. The system performance is commonly measured by the Misalignment (MISA) and the Perceptual Evaluation of Speech Quality (PESQ) criteria. However, injecting the noise for a long period of time would corrupt the intelligibility of the speech signal during the filter estimation. To avoid injecting the noise for a long duration, we propose using several short-duration noise injections which would decrease the errors in high lags of correlation by averaging while the intelligibility of speech signal is preserved during the estimation of AFC filter. More accurate correlation samples enable us to determine AFC filter of higher order improving the performance criteria. The averaging of the correlation samples in our approach is indirect. That is, after each short-time noise injection we find a new estimate of the feedback path model and average of this new model and the previous one is used as the feedback compensating FIR filter. Applying the proposed method to several speech files has verified the improvement of system performance in terms of MISA and PESQ criteria. The MISA and PESQ performance criteria are evaluated during the filter estimation process and after finding the final filter as the feedback estimator. Furthermore, we show that in cases where the quality of the signal, which is calculated by PESQ, is the most concern, even one 20-msec of noise injection can be enough. Otherwise, if we require better Misalignment we may use several short-time noise injections.

The paper is organized as follows. Section II describes the noise injection method used in this paper. Section III explains the idea of multiple noise injections and averaging procedure. Experimental results and conclusions are provided in Sections IV and V, respectively.

II. NOISE INJECTION METHOD

Noise injection methods are categorized into two major groups, i.e. continuous noise injection [2] and non-continuous noise injection methods [3]-[6]. The applied method in this paper is classified in the second group. Zero-mean White Gaussian noise is injected into the loudspeaker instead of the microphone signal for a while. As shown in Fig. 1 we have:

$$y[n] = s[n] + \hat{u}[n] * f[n] \tag{1}$$

where $f[n]$ is the actual impulse response of the feedback path; $\hat{u}[n]$ is the loudspeaker input where $\hat{u}[n] = r[n]$ during noise injection period. $r[n]$ is the injected zero-mean White noise. Using Eq. (1), cross-correlation between the microphone and loudspeaker signals is calculated by:

$$r_{y\hat{u}}[l] = r_{s\hat{u}}[l] + r_{\hat{u}\hat{u}}[l] * f[l] \tag{2}$$

Variable l represents the correlation lag and "*" denotes linear convolution operation. $r_{s\hat{u}}[l]$ is negligible since the white noise is uncorrelated with speech signal.

Considering different lags for the cross-correlation we have a new equation in matrix form:

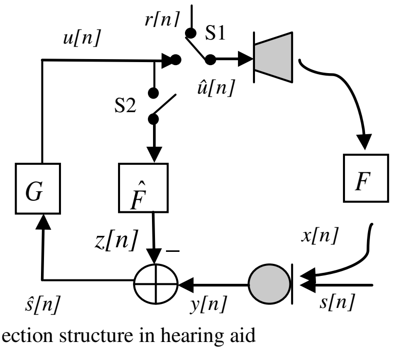

Fig. 1. Noise injection structure in hearing aid

---

* This research work was supported in part by the Department of Electrical Engineering, the University of Texas at Dallas.

$$\boldsymbol{r}_{y\hat{u}} = \ \boldsymbol{R}_{\hat{u}\hat{u}} \cdot \hat{\boldsymbol{f}} \tag{3}$$

$$\boldsymbol{r}_{y\hat{u}} = [r_{y\hat{u}}[0] \ \ r_{y\hat{u}}[1] \dots r_{y\hat{u}}[L-1]]^T_{1\times L} \tag{4}$$

$$\hat{\boldsymbol{f}} = [\hat{f}_0 \ \ \hat{f}_1 \cdots \hat{f}_{M-1} \ ]^T_{1\times M} \tag{5}$$

$$\boldsymbol{R}_{\hat{u}\hat{u}} = \begin{bmatrix} r_{\hat{u}\hat{u}}[0] & r_{\hat{u}\hat{u}}[-1] & \dots & r_{\hat{u}\hat{u}}[-M+1] \\ r_{\hat{u}\hat{u}}[1] & r_{\hat{u}\hat{u}}[0] & \dots & r_{\hat{u}\hat{u}}[-M+2] \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ r_{\hat{u}\hat{u}}[L-1] & r_{\hat{u}\hat{u}}[L-2] & \dots & r_{\hat{u}\hat{u}}[-M+L] \end{bmatrix}_{L\times M} \tag{6}$$

The length of the AFC filter $\hat{\boldsymbol{f}}$ is represented by M (order = M-1) and the maximum lag of correlation is L. Using a short duration of injection corresponds to windowed version of the injected noise resulting in non-zero off diagonal entries of matrix in (6). Hence, non-diagonal $\boldsymbol{R}_{\hat{u}\hat{u}}$ is used in our calculations instead of a diagonal matrix with $r_{\hat{u}\hat{u}}[0]$ on the main diagonal. $\hat{\boldsymbol{f}}$ can be found by (least squared estimation) when $L > M$ [7]:

$$\hat{\boldsymbol{f}} = \ (\boldsymbol{R}_{\hat{u}\hat{u}}{}^T \boldsymbol{R}_{\hat{u}\hat{u}})^{-1} \boldsymbol{R}_{\hat{u}\hat{u}}{}^T \boldsymbol{r}_{y\hat{u}} \tag{7}$$

We assume L=M in this paper. Thus, the FIR filter coefficients are given by:

$$\hat{\boldsymbol{f}} = \ (\boldsymbol{R}_{\hat{u}\hat{u}})^{-1} \boldsymbol{r}_{y\hat{u}} \tag{8}$$

The estimated model derived from Eq. (8) is put in parallel with the feedback path to cancel the feedback signal, i.e. the canceller path is connected to the other parts of the system by switch S2 (Fig. 1).

Using noise injection method solves the biased estimation problem [1] because of having uncorrelated signals as the loudspeaker output and microphone input. However, the accuracy is dependent on the number of samples used for computing the correlation.

III. MULTIPLE NOISE INJECTIONS

One important step in design of an AFC is selecting a proper length for the filter $\hat{\boldsymbol{f}}$. A filter with an improperly small length does not have adequate efficiency. Choosing a length higher than what is required increases the computational complexity of the system; while, it does not improve the performance compared to the performance gained by the proper length. Method in [8] has used two criteria, i.e. PESQ and MISA to monitor the performance and find the optimum filter length. The two criteria are defined as follows:

A. *Perceptual Evaluation of Speech Quality (PESQ)*

PESQ is a test methodology for automated assessment of speech quality. The PESQ value is in the range of 0.5 (bad) to 4.5 (excellent) and its computation is depicted in Fig. 2 [9].

B. *Misalignment (MISA)*

MISA is the normalized energy of the error between the actual feedback model and the estimation:

$$MISA = 10log_{10}\left(\frac{\int_0^{\pi} \left|F(e^{j\omega}) - \hat{F}(e^{j\omega})\right|^2 d\omega}{\int_0^{\pi} \left|F(e^{j\omega})\right|^2 d\omega}\right) \tag{9}$$

According to this formula smaller MISA represents better performance.

To explain the idea proposed in [8], we present Table I which is the results of noise injection with different settings. For analysis and simulation purposes, an actual feedback path model measured in the laboratory is used whose impulse response and transfer function are plotted in Figs. 3 and 4, respectively. It is an FIR filter of length 88.

Table I summarizes some experimental results. Three different durations of noise injection are considered in this table, i.e. 100, 60, and 20 msec. For each of them three different filter lengths are examined for the AFC filter $\hat{\boldsymbol{f}}$, i.e. 60, 88, and 100. Last two columns of the table represent PESQ and MISA values. For 100-msec noise injection, the best MISA is related to the length 88 which is the length of actual feedback model. For lengths below or above this, MISA is obviously deteriorated. However, the best PESQ corresponds to the length 60 which is smaller than the actual length. Comparing the PESQ values for all three orders shows that it is not required to increase the length of $\hat{\boldsymbol{f}}$ to 88 if our main concern is the quality of the signal heard by the patient. Checking the results provided by 60-msec noise injection verifies the previous conclusions. For 20-msec noise injection the best PESQ and MISA are obtained for the filter of length 60. These results are not completely compatible with the 60-msec and 100-msec noise injection cases. The incompatibility is due to the existence of error in correlation lags. In 20-msec noise injection at sampling frequency of 16 KHz, 320 samples of noise are injected to the loudspeaker. Since the actual feedback path model has length of 88, the first 88~100 samples of the input noise and the output of the feedback path are discarded as the transient parts of the input and output data. Hence, the number of recorded data samples that are actually used for calculating the correlation lags is at most 232. As a rule of thumb, the maximum lag which can be calculated for correlation without significant error is around 1/3 of the number of data samples, i.e. maximum lag of 77 for 20-msec noise injection. Hence, for 20-msec noise injection, correlation samples for the filter lengths of 88 and 100 contain errors causing inconsistent results especially for MISA. However, the results obtained for the filter length of 60 is acceptable.

Table I. Performance evaluation for different noise injection durations and filter lengths

Injection Duration (msec)	Filter Length	PESQ	MISA
100	60	**4.1238**	-17.9024
100	88	4.1233	**-24.4570**
100	100	4.1221	-23.5346
60	60	**4.1429**	-17.8192
60	88	4.1425	**-23.8443**
60	100	4.1422	-23.7637
20	60	**4.1669**	**-12.6040**
20	88	4.1498	-12.4835
20	100	4.1488	-11.8656

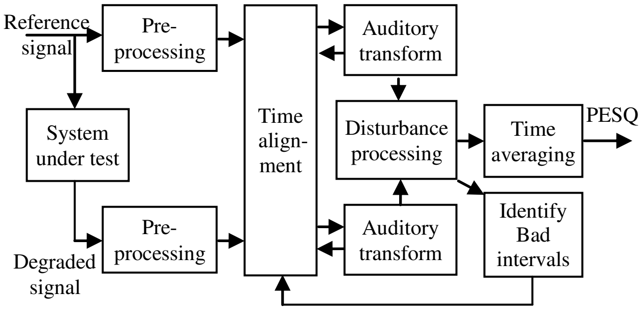

Fig. 2. Block diagram of the PESQ measure computation

According to the results shown in Table I, for the 60-msec and 100-msec noise injections, if the quality (PESQ) of the signal heard by the patient is the main concern then using a filter length around 60 should suffice. We note that in this case, we can even use only 20-msec noise injection. The 20-msec noise injection makes the PESQ even better than that of the same filter length but for 60-msec and 100-msec noise injections as shown in Table I. However, in a case where MISA is more important, the filter length 60 is not enough.

Improving the MISA requires higher order AFC filter which in turn calls for more data via longer duration of noise injection. Long duration of noise injection degrades the intelligibility of the speech signal and causes some irritation to the patient. To alleviate these drawbacks, we propose the following approach: (a) let AFC filter have high length, e.g. 88 or higher, (b) apply multiple short-time (20-msec) noise injections, each sufficiently apart in time, (c) find an AFC filter for each noise injection, (d) find weighted average of this new filter and the previous filter, (e) use the filter obtained by averaging as the estimator of the feedback path model. The averaging process is given in Eq. (10).

$$\hat{\boldsymbol{f}}_m = \propto \hat{\boldsymbol{f}}_{new} + (1-\propto)\hat{\boldsymbol{f}}_{m-1} \qquad (10)$$

$\hat{\boldsymbol{f}}_{new}$ is the new filter estimate found from Eq. (8) using samples of correlation of m^{th} noise injection. $\hat{\boldsymbol{f}}_m$ is the averaged filter estimate used for the AFC filter after m^{th} noise injection. $\propto$ is a positive weighting factor chosen less than 0.5 . Clearly $\propto < 0.5$ (e.g. $\propto = 1/m$ for $m \geq 2$) gives the priority to the average of previously estimated filter and decreases the effect of the new estimate. This situation is suitable when the new filter estimate has some error, as discussed in the next section.

IV. EXPERIMENTAL RESULTS

In this section the proposed multiple noise injection method is implemented and analyzed. The actual feedback path model is the one shown in Figs. 3 and 4. Table II shows the result for 5 injections of zero-mean White Gaussian noise each with 20 msec duration. Recall that for each such injection only about 232 data samples are available for computing the correlation lags required by Eq. (8). The period of noise injection is 500 msec. Table II represents the average of 50 executions of the algorithm for each speech file. 15 different speech files are used as $s[n]$ in Fig. 1. The first result column of Table II shows the MISA values corresponding to $\hat{\boldsymbol{f}}_{new}$ for each noise injection. Second column contains MISA values corresponding to $\hat{\boldsymbol{f}}_m$ after m^{th} noise injection. That is, the MISA of the system after m 20-msec noise injections. Comparing different rows of this column verifies the improvement of MISA using multiple noise injections. Four right columns of Table II show the values for PESQs computed differently for the system. These values are obtained according to Fig. 5 as follows. The entire length of the experiment is ≈160,000 msec of which the first 2,020 msec is used for noise injection and finding the filter estimate. The Section PESQ$_m$ column shows the value of PESQ for 500 msec period of the signal which consists of the m^{th} 20-msec noise injection part. The ratio of $\frac{Noise\ injection\ duration}{File\ duration} = \frac{20\ msec}{500\ msec}$ is much higher than what we would have using only the one-time noise injection $\frac{20\ msec}{\approx 160000 msec}$. Thus, the values in the Section PESQ$_m$ column do not seem good enough compared to what we had for the system with one-time noise injection. The values shown in the column under Sub-Section PESQ$_m$ show the values of the PESQ for only the 480 msec of the signal between the m^{th} and the $(m+1)^{th}$ noise injections. The Total PESQ shown in the table is the value of the PESQ for the whole duration of signal which consists of 5 noise injections. The Final PESQ value shows the quality of the signal heard by the patient after finishing the noise injections for finding the FIR filter. The Final PESQ is better than what we would get by injecting only a single 20 msec of noise.

Table III shows the results for the proposed method with 10 noise injections, each of 20 msec long, with repetition period of 1000 msec. The values in Table III show some improvements over those in Table II. The final value of MISA is improved compared to that in Table II because of having the average of 10 noise injections instead of 5. Also all the PESQ values are slightly better than those shown in Table II mainly due to the longer repetition period.

In our experiments $\propto = 1/m,\ m \geq 1$ was used in Eq. (10). That is, as m increases we put more emphasis on the filter estimate obtained by averaging the m-1 previous filters than the new filter estimate obtained for only the m^{th} noise injection. The accuracy of the estimation is not completely independent of the speech signal $s[n]$. When the short-time noise injection coincides with the speech signal having a large energy, i.e. when a vowel is being uttered (Region (a) in Fig. 6), then the resulting filter estimate is not as accurate as that obtained when the speech signal has low energy, i.e. silence part or consonant uttering part (region (b) in Fig. 6).

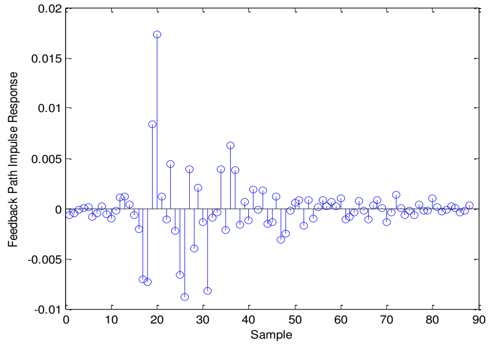

Fig. 3. Impulse Response of the measured/true feedback path

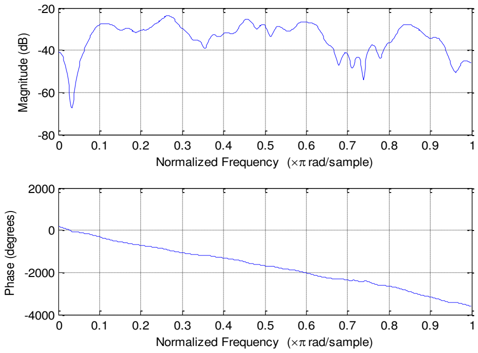

Fig. 4. Frequency Spectrum of the measured/true feedback path

The multiple noise injection and averaging method proposed in this paper alleviates the impact of the speech energy on the AFC filter estimation.

Table II Results for five 20-msec noise injections with period of 500 msec

m	MISA for $\hat{f}_{new}$	MISA for $\hat{f}_m$	Section PESQ$_m$	Sub-Section PESQ$_m$	Total PESQ	Final PESQ
1	-12.1963	-12.1963	3.4193	4.1446		
2	-11.5042	-13.9588	3.6005	4.1649		
3	-11.1724	-14.8535	3.2477	4.2641		
4	-11.1299	-15.4257	3.2213	4.1961		
5	-11.4764	**-15.9692**			3.2212	**4.4371**

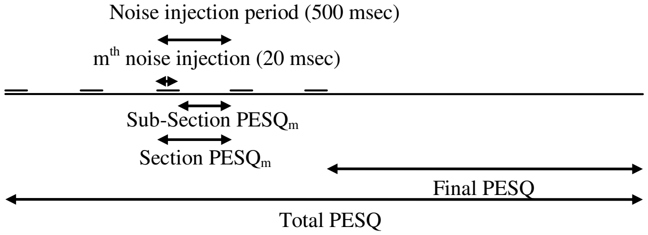

Fig. 5. Different types of PESQs

Table III Results for ten 20-msec noise injections with period of 1000 msec

m	MISA for $\hat{f}_{new}$	MISA for $\hat{f}_m$	Section PESQ$_m$	Sub-Section PESQ$_m$	Total PESQ	Final PESQ
1	-12.2338	-12.2338	3.7682	4.2232		
2	-11.7812	-14.1274	3.6351	4.3220		
3	-11.2166	-14.9700	3.5704	4.2803		
4	-12.2515	-15.8006	3.6878	4.3338		
5	-10.8213	-16.1445	3.4016	4.3226		
6	-11.1965	-16.4203	3.5885	4.3202		
7	-11.0911	-16.6833	3.5144	4.3648		
8	-11.1332	-16.9003	3.3987	4.3682		
9	-10.9751	-17.0604	3.5698	4.3957		
10	-10.8122	**-17.1962**			2.6642	**4.4644**

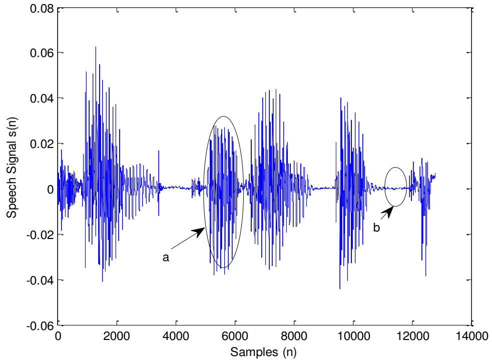

Fig. 6. A section of speech file. Region (a) corresponds to uttering a vowel. Region (b) corresponds to the silence or uttering a consonant

V. CONCLUSION

We have used noise injection method in order to estimate the acoustic feedback path model in hearing aid. The estimated model, as an AFC FIR filter, is used to cancel the irritating effect of the acoustic feedback. Typically, a filter length smaller than the length of the actual feedback path model is enough to obtain a good quality and low computational complexity in the system design. However, if instead of quality represented by the PESQ, the amount of Misalignment criterion is important to us, we need longer duration of noise injection which degrades the intelligibility of the speech signal. To preserve the intelligibility of speech signal and obtain high order AFC filter, we proposed a new method which would (1) apply multiple short-time noise injections instead of one long-duration injection, (2) find a filter estimate for most recent noise injection, (3) find a weighted average of the recently found filter and those obtained previously to use as the AFC FIR filter. Experimental results verified the improvement achieved by the proposed method.

REFERENCES

[1] T.V. Waterschoot and M. Moonen, "Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges," *Journal, Proceedings of the IEEE*, Vol. 99, No 2, pp. 288-327, Feb. 2011.

[2] C. P. Janse and C. C. Tchang, "Acoustic Feedback Suppression," *WIPO Patent Appl*. WO/2005/079 109, Aug. 2005.

[3] S. Ibaraki, H. Furukawa, and H. Naono, "Howling Canceller (Patent style)," U.S. Patent 4 747 132, May 1988.

[4] J. M. Kates, "Feedback Cancellation in Hearing Aids", *IEEE, ICASSP*,Vol. 2, pp. 1125-1128, 1990.

[5] J. A. Maxwell and P. M. Zurek, "Reducing Acoustic Feedback in Hearing Aids," *IEEE Transactions on Speech and Audio Processing*, Vol. 3, pp. 304-313, Jul. 1995.

[6] D. Graupe, J. Grosspietsch, S.P. Basseas, "Method of and Means for Adaptively Filtering Screeching Noise Caused by Acoustic Feedback (Patent style)," US Patent 4783818, Nov. 1988.

[7] S.A.Khoubrouy, I.M.S. Panahi," Modeling feedback path in hearing aid by single injection of short-time noise," *Submitted to IEEE Transaction on Audio, Speech, and Language Processing*, Dec. 2011.

[8] S.A. Khoubrouy, I.M.S. Panahi, V. Montazeri, "Order Determination for the Feedback Path Estimator in Hearing Aid", *Inter-noise 2012.*

[9] P. Loizou, *Speech Enhancement: Theory and Practice. CRC Press, 2007*, pp. 514–525.