Acoustic Feedback Cancellation For Public Address Systems

Experiments using a personal computer
to implement an adaptive filter

Edgar Berdahl
8/25/05

EE373
Prof. Widrow and Dr. Aaron Eliasib Flores

Table of Contents

1.0 Problem Description

People use public address (PA) systems to amplify their voices in various situations (see Figure 1). In many cases, acoustic paths exist between the speaker device and the person, who will from here onward be referred to as the *talker* to avoid confusion with the *speaker* device. The sum of the acoustic paths is modeled by the dotted line in Figure 1.

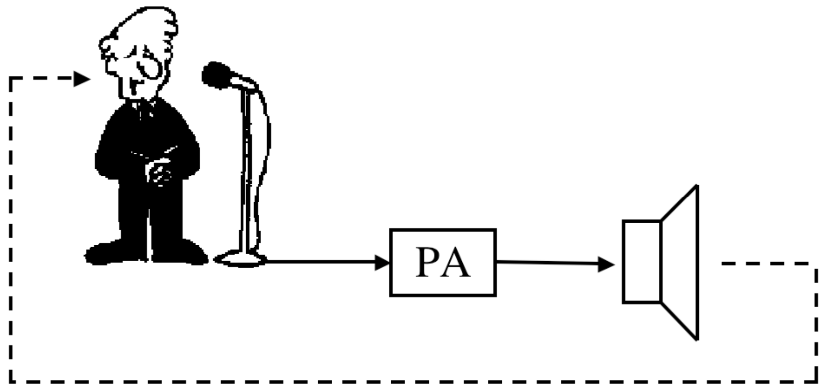


Figure 1. Public address (PA) system with the acoustic feedback path (dotted line)

The block diagram presented in Figure 2 describes the feedback loop inherent in the PA system more explicitly. Assuming that the elements are not being overdriven, the system is linear. $H_{PA}(f)$ represents the transfer function due to the amplifier, while $H_f(f)$ represents the cascade of the speaker, room, and microphone transfer functions.

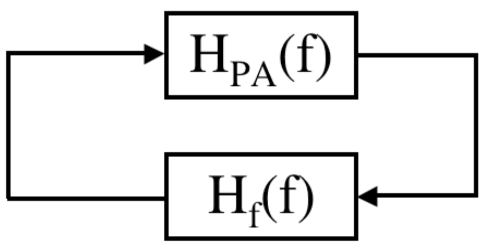


Figure 2. Block diagram of the feedback loop

PA systems are often plagued by acoustic feedback, which is also sometimes referred to as howling. This occurs when the signal in the feedback loop grows unboundedly, or at least until elements of the system begin clipping. Since howling is usually very loud, it is unpleasant. Here systems for canceling acoustic feedback are investigated.

The Nyquist criterion for stability requires that $|H_{PA}(f)H_f(f)| < 1$ for all f such that $\angle(H_{PA}(f)H_f(f)) = n360°$ where n is any integer. However, note that the talker might move the microphone slightly. Furthermore, the talker's own movements could cause phase changes in $H_f(f)$ that are difficult to predict. Thus, to be sure that the feedback system is stable, the phase requirement is no longer particularly helpful. That is, to be sure that the system is stable, we need the following at all frequencies f: $|H_{PA}(f)H_f(f)| < 1$

Proper sound design can help alleviate howling. Often the speaker is placed in front of the microphone, and since both the microphone and speaker are directionally-dependent, this configuration can significantly reduce $|H_f(f)|$. Well-versed sound technicians will also often test the system by increasing the gain until feedback occurs and then apply notch filters corresponding to the peaks in the loop transfer function. By flattening out the spectrum, these sound technicians are actually increasing the amount of power that can be reproduced by the sound system before howling sets in.

However, in some situations, too many notch filters would be required to avoid significantly changing the signal, and the loop gain of the system will simply need to be decreased. This solution can be undesirable in many instances, and PA's that could automatically adjust themselves to their surroundings would be easier to use. Furthermore, by taking advantage of the increase in the allowable loop gain, the distance between the talker and the microphone could be increased, which would help provide an unobstructed view of the talker.

It is important to note that howling tends to occur at one particular frequency. Before howling occurs, the loop transfer function has magnitude less than 1 at all frequencies. Then, as howling begins to set in, the transfer function's magnitude will exceed 1, possibly at several different frequencies. Generally, the frequency for which the loop transfer function has the highest magnitude will grow faster than the others and "win the race." This can be considered the dominant eigenfrequency of the loop transfer function.

Various techniques for overcoming acoustic feedback cancellation will be studied in this report. Most of them rely on using adaptive filters that are trained using the least-mean squares (LMS) algorithm to cancel the feedback.

2.0 Previous Work

2.1 Adaptive Periodic Noise Canceller (APNC)

The earliest work of which we are aware was carried out by J.B. Foley in 1989. It relies on the fact that howling is periodic, while speech is not when viewed over a long enough time scale. Foley used an adaptive periodic noise canceller (APNC), as depicted in Figure 3, to eliminate sinusoidal components from the speech signal[1]. The delay *D* was chosen to be large enough such that the speech was largely uncorrelated with itself, while *D* was small enough such that the howling could be cancelled before it grew too much in magnitude. We chose *D* corresponding to 2ms for our experiments although the exact value was not particularly crucial.

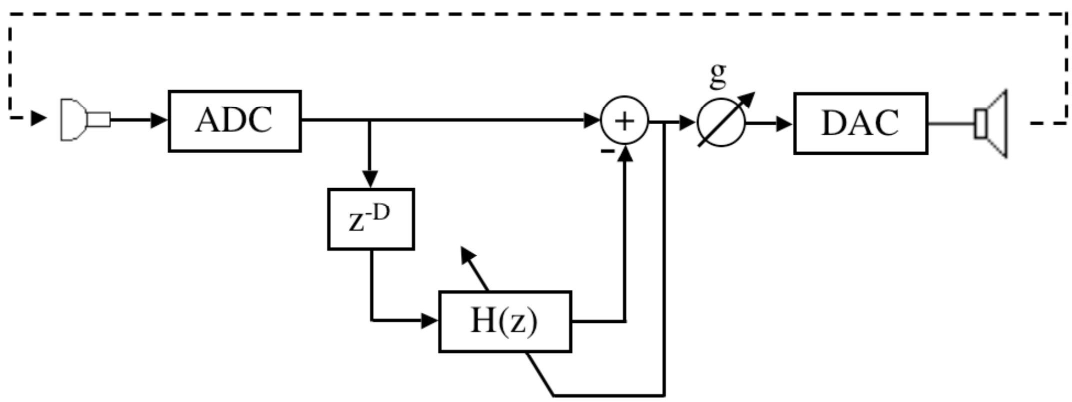


Figure 3. Block diagram of the adaptive periodic noise canceller (APNC)

2.2 Feedback Noise Canceller (FBNC)

The feedback noise canceller (FBNC) is a more common topology used in hearing aids and is shown in Figure 4. Again a delay *D* is required to reduce the correlations in the speech[2]. By making *D* correspond to the same delay imposed by the cascade of the ADC and DAC, the adaptive filter can be made to converge to a transfer function that models the transfer function of the cascade of the DAC, speaker, room, microphone, and ADC. This configuration uses the transfer function estimate to subtract out an estimate of the feedback signal at the microphone. Note that *D* may be chosen much larger than in the case of the APNC since the ADC and DAC delays can be quite long.

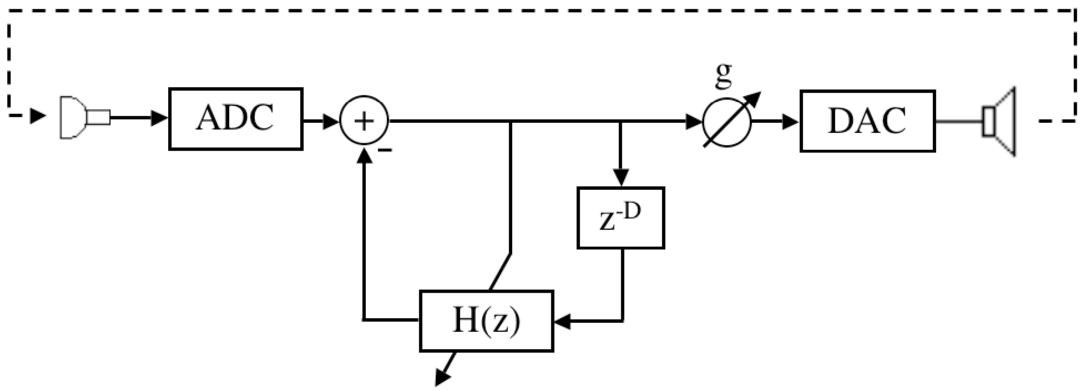


Figure 4. Block diagram of the feedback noise canceller (FBNC)

This topology is used in many commercial hearing aids, and so a great deal of research has been carried out with regards to this configuration. Some hybrid configurations also exist including one where band-pass filters are used to help a filtered-X adaptive filter converge more quickly in a particular frequency band of interest[3].

Note that the presence of any feedback path can introduce the local minima for the LMS algorithm. Thus, under conditions involving large amounts of acoustic feedback, the APNC algorithm could theoretically suffer due to multiple minima, meaning that the adaptive filter might settle on a locally optimal solution rather than the globally optimal solution. In contrast, in the case of the FBNC, the feedback path is hard-wired into the signal processing block, so the FBNC is more likely to suffer from locally-optimal solutions.

2.3 Block Methods

Since any given room's impulse response can last up to a few seconds, and since the FBNC models the impulse response of the loop transfer function, one might imagine the need for adaptive FIR filters implementing tens of thousands of adaptive coefficients. To implement any filters of this length in real-time on standard hardware, one would need to use block processing methods. Thomas Fillon and Jacques Prado applied a particular block algorithm known as the generalized multi-delay filter algorithm (GMDα) to acoustic feedback cancellation in hearing aids and achieved some success in a computer simulation[4]. For a more detailed description of GMDα, see papers written by Eric Moulines et al[5] and Soo and Pang[6]. GMDα has also been applied to related topics in acoustic echo cancellation[7].

Note that because the talker might move quickly, we need the adaptive filter to adapt quickly. However, the excess MSE of the LMS algorithm increases with the number of filter taps if the rate of adaptation is held constant[8]. This implies that very long filters should have convergence problems, and so we chose not to use block methods for our investigation. Perhaps GMDα is better suited to acoustic echo cancellation where the transfer function is less likely to quickly change drastically.

2.4 Other Methods

2.4.1 Quasi-Proportional Frequency Shifting

Another intriguing idea is introducing a nonlinear block into the loop to eliminate the effects of the positive feedback. For instance, one might imagine a transformation where the frequencies of all of the components of the signal are increased or decreased by a small amount. Written in terms of the power density spectrum S, we could impose that $S_{out}(f)=S_{in}(\beta f)$ for some $\beta\approx 1$. This transformation is particularly convenient because it preserves harmonic relationships and so the perceived distortion of speech should be minimal. Alango Technology in fact sells a system that implements the relation $S_{out}(f)\approx S_{in}(\beta f)$. It seems likely that their system uses the Fast Fourier Transform to this end, and they claim that the approximation is supposedly not noticeable due to the limited frequency resolution of the human ear. However, their system must not work in all configurations because they also apply a "subband howling blocker" for situations where the system becomes unstable despite the nonlinear transformation (probably they set FFT bins that are escalating out of control to zero). Furthermore, their system cannot be particularly efficient as it imposes a processing delay of 5ms and requires up to 5 million instructions per second on a DSP.

2.4.2 Automatic Adjustment Of Notch Filters

Behringer GmbH sells a feedback elimination device, which they call the "Feedback Destroyer"[9]. It operates in a manner similar to a very quick sound engineer. It waits until howling occurs and then places a notch filter at the offending frequency. Behringer has released a number of products relying on this technology, so the system must work at least fairly well.

3.0 Implementation On A Personal Computer

3.1 Software

Past systems have used DSPs to implement the adaptive filters, but in our case we used a personal computer due to its ease of programming and debugging. However, the main disadvantage of using a personal computer was that common operating systems are not optimized for high-performance audio signal processing. Instead, most operating systems allow a large number of tasks to run concurrently. As a result, many different kinds of interrupts can occur at essentially unpredictable times, and so to ensure that no samples are dropped as they arrive from the ADC or are sent to the DAC, large audio vectors must be used. All applications using the audio drivers then require only one interrupt per audio vector. These vectors result in total system delays known as latencies. For our experiments, we chose to use Linux, as it allows small system latencies on the order of 2ms.

We implemented the signal processing part of the system in Miller Puckette's open source environment known as Pure Data (`Pd`), which allows various audio signal processing blocks to be connected together in a graphical user interface. However, there was no adaptive filter block, so we had to write one called `adaptfilt~` in C++. As a caviat for any others who attempt this method, use the `delread~` and `delwrite~` objects for implementing the delay of D samples carefully. For simplicity of implementing the signal processing chain, no delay shorter than the audio vector length can be created; however, delays of all other numbers of samples can be created, even though they are specified in terms of milliseconds. There is no round-off error.

To help monitor the performance of the system, we ran the program `freqtweak` concurrently. It can be used to display the spectrum of the signals in the ADC and DAC in real-time. This helps one visualize the growth and death of various eigenfrequencies and shows in which frequency bands they tend occur.

In particular, a Pentium IV processor running at 2GHz, allows about 250 taps for an adaptive filter at a sampling rate of 22.05kHz before occasional audio drop outs occur. However, since the LMS algorithm for adaptive filters is so robust, about 500 taps can be used before the system stops converging to a reasonable solution. Probably this could be improved some by optimizing the graphical user interface of Pd or by using more stable audio drivers. (We found that the Jack audio drivers for Linux were not particularly stable, but we used them anyway because they allowed us to run `freqtweak` and `Pd` at the same time.)

3.2 α-LMS

Consider the situation where an adaptive filter is trying to converge to an optimal solution to cancel acoustic feedback. At a certain point in time, a sinusoid at a given eigenfrequency of the system is increasing in magnitude, and the adaptive filter is trying to catch it before the system becomes nonlinear. If the system were to become nonlinear, then the adaptive filter would be less able to take nonlinearities into account and would likely loose control of the system. As the sinusoid increases in magnitude, the standard LMS algorithm will adapt more quickly. This is undesirable because if the LMS algorithm adapts too quickly, then due to its increased excess MSE, the system will become unstable and explode. As a result, α-LMS, which is also known as normalized LMS, is required to control the rate of adaptation μ. Instead of the user specifying μ, the user specifies the misadjustment M.

The adaptive filter then adapts according to $\mu = M / P / L$, where L is the length of the adaptive filter and P is an estimate for the power of the input signal, which is computed using an envelope follower. Since P changes with every input sample, so does μ, and thus the system can also adapt quickly to inputs of various amplitudes. Note finally that applying the constraint $\mu_{eff} = \mathrm{MAX}(\mu, \text{constant})$ is a detail required to limit the speed of adaptation when the talker is not talking.

Applying α-LMS to the feedback cancellation problem is reminiscent of Joseph A. Maxwell's results. When the input amplitude was small, he injected white noise into the system to keep the system adapting and help track changes in the feedback transfer function $H_f(z)$ [10]. In another experiment, James M. Kates also enabled adaptation that was dependent on the amplitude of the input signal[11].

3.3 2mic Feedback Canceller

We further experimented with a novel feedback cancellation configuration. It requires an additional microphone, yet the optimal solution that the adaptive filter is seeking is usually slightly shorter in length because it is not the entire loop transfer function, but rather the transfer function between the two microphones. The system requires no additional significant computation, and is shown in Figure 5. The first microphone is used by the talker and is referred to as the *talker mic*, while the second microphone is placed directly in front of the speaker and is referred to as the *speaker mic.* An additional advantage of this configuration is that the system latency delay need not be identified. Rather, D represents only the delay between microphones two and one. One would expect behavior similar to the FBNC.

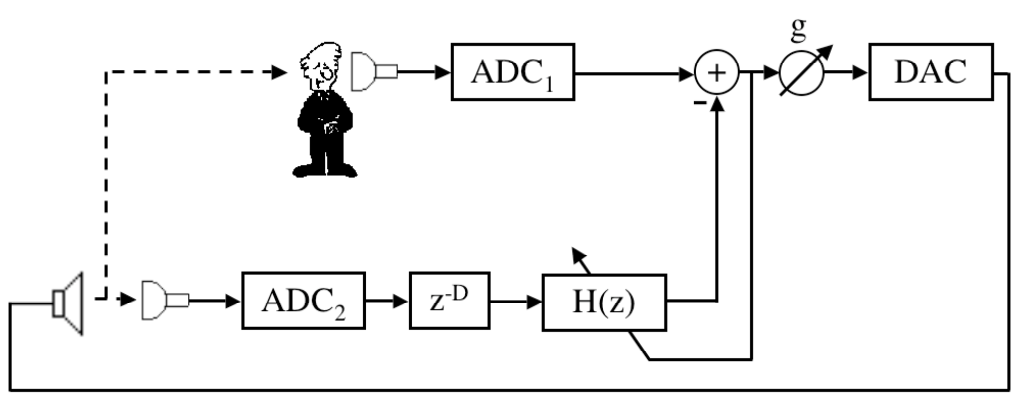


Figure 5. 2mic feedback canceller

4.0 Experimental Results

4.1 Setup

A picture of the measurement setup is shown in Figure 6. Since acoustic feedback can become quite loud, we implemented the speaker with a Fender Princeton 112 42W guitar amplifier due to its robustness. We used karaoke microphones for input and simulated the talker using white noise output by an additional speaker. Of course, as soon as the noise left the speaker and was detected by a microphone it was no longer white but colored somewhat.

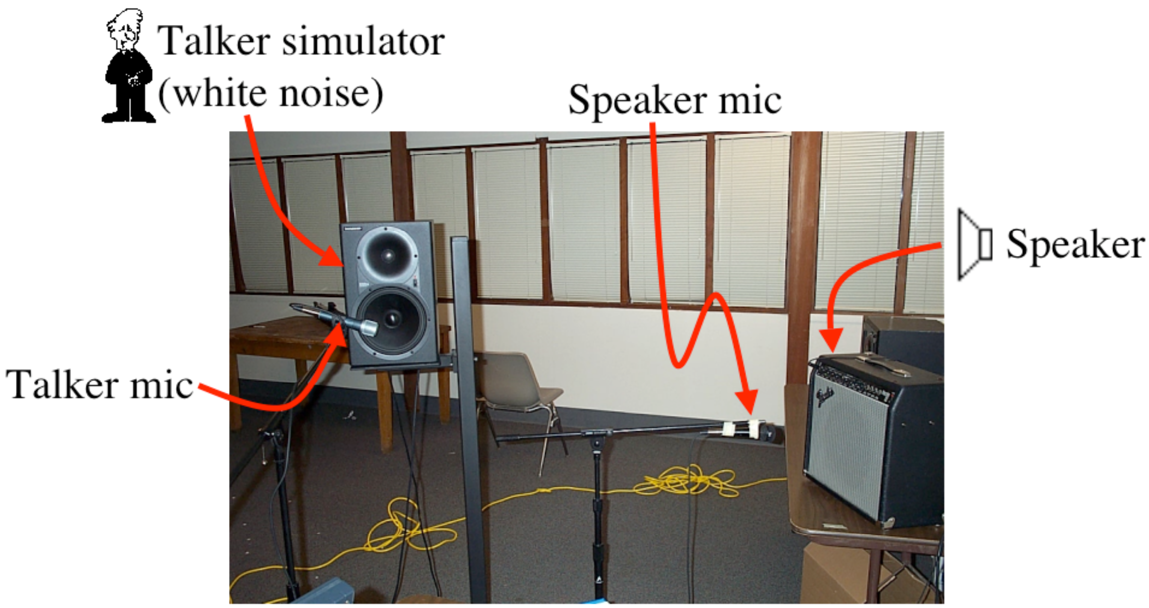


Figure 6. Measurement setup

4.2 Loop Transfer Functions

The transfer function magnitudes of particular loop systems from the DAC to the ADC's (inclusive) via the talker and speaker microphones are shown in Figure 7. Here it can be seen that the talker loop transfer function is "bumpier" or "more complicated" than that of the speaker loop. This results from the talker microphone's increased distance from the speaker (see Figure 6). This gives one a rough idea of to what degree the room's transfer function affects the talker loop more than the speaker loop. The difference is especially notable at higher frequencies.

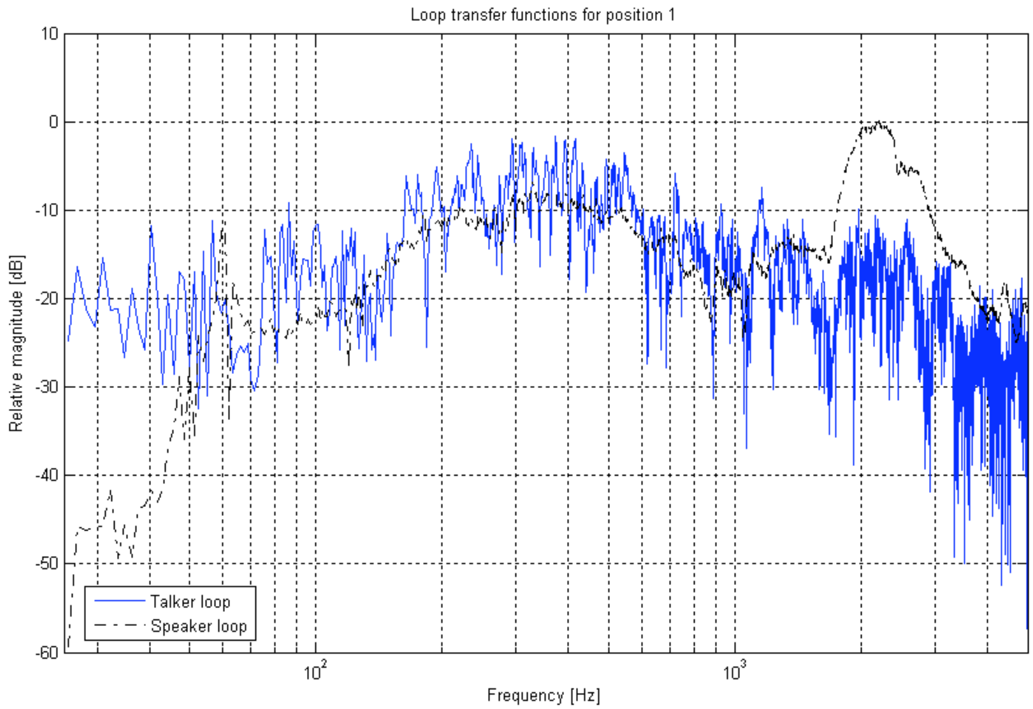


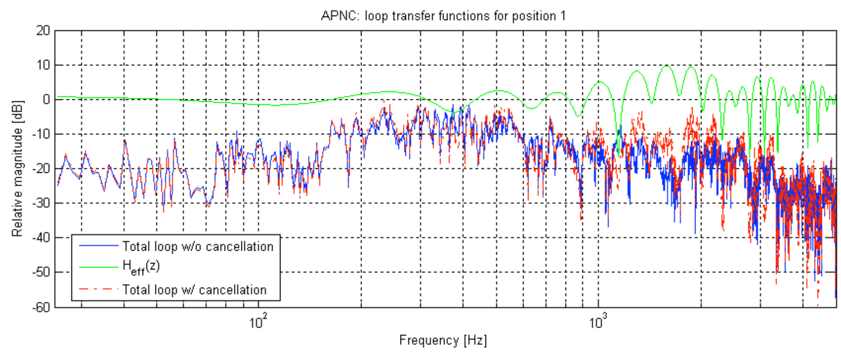
Figure 7. Loop transfer functions

4.3 APNC

Theoretically speaking, since howling primarily consists of one sinusoid, only 2 taps should be required for the adaptive FIR filter; however, generally about 20 worked much better in our case since the howling frequencies were much lower than the Nyquist rate and since generally some competitor sinusoids at lower amplitudes were also present at any given time.

To test the quality of the acoustic feedback cancellation, we measured the increase in the maximum stable gain of the system (MSG increase). That is, we first removed the adaptive filter from the system and measured the minimum loop gain g at which howling would set in. Then we decreased the gain, enabled the adaptive filter, let it begin adapting, and determined how much further g could be increased before the system would loose control in comparison with the case without the adaptive filter. We then repeated the procedure for a number of various locations in the room. In the case of the APNC, we achieved an increase in MSG of 1.5dB to 2dB, depending on the particular orientation of the microphones, speaker, and other objects in the room.

As the system ran in time, a sinusoid at an eigenfrequency would start increasing in amplitude until the frequency canceller adapted and eliminated it. In general, several eigenfrequencies were competing at a given point in time, and the system was in a marginally-stable state. Figure 8 shows "snapshots" of the relevant transfer function magnitudes at a particular point in time. The blue transfer function magnitude represents the system without the influence of the adaptive filter. The red transfer function magnitude represents the change in the system due to the adaptive filter. Thus, Figure 8 shows how the adaptive filter decreases the amplitude of the part of the transfer function with the largest magnitude. (The green transfer function depicts the portion of the loop transfer function magnitude due to the adaptive filter *H(z)*. Thus, pointwise multiplying the blue transfer function by the green transfer function results in the red transfer function.)



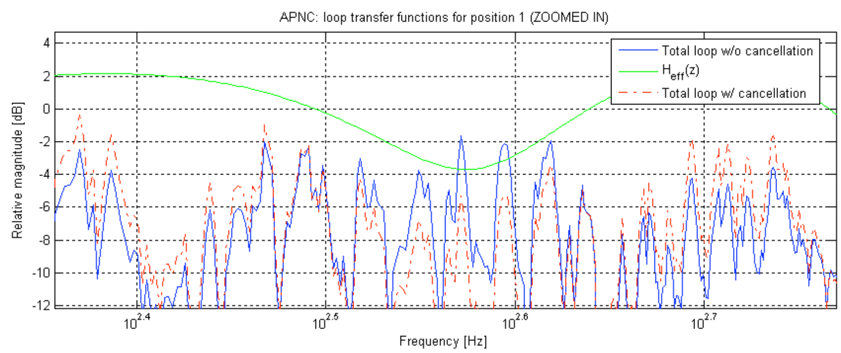


Figure 8. Relevant transfer function snapshots for the APNC, shown zoomed out (a) and zoomed in (b) on the frequency region where the total loop transfer function had the largest magnitude

4.4 FBNC and 2mic

A good choice for the delay *D* turned out to be the delay such that the largest reflection due to the loop transfer function of the system arrived about 1/4 of the way through the time modeled by the adaptive filter. This gave the adaptive filter a little bit of time to prepare for the largest reflection, and then more time to manage the tail of the (room's) impulse response. If the delay *D* was set improperly, or if the adaptive filter was not given enough coefficients to effectively approximate the loop transfer function, then the FBNC and 2mic configurations performed roughly the same as the APNC, although sometimes several dB better in certain positions. In addition, there was no evidence of a particularly optimal solution. That is to say, the solution chosen by the adaptive filter depended heavily on the statistics of the particular given input and changed over time.

However, given the proper system delay *D* and 110 adaptive filter coeffiecients, adaptive filters in the FBNC and 2mic configurations converged to estimates of the transfer function. This resulted in MSG increases in the range of **6dB** to **20dB**. Such a large allowable increase in gain would make the system useful in speech applications. However, the distortion to the voice might be unacceptable for musical applications due to the effects of the misadjustment of the adaptive filter, which was on the order of 0.6% to 2%.

The best results were obtained from configurations where the microphones were pointing directly at the speaker. This reduced the effects of the room reverberation so that the adaptive filter could converge quickly to a shorter approximate solution. The training signals used for the adaptive filter were thus also whiter because they were less colored by the room. Notably, if the talker made long "sssss…" sounds, then the adaptive filter would converge to the approximate loop transfer function of the system, while if the talker produced a signal with less higher-frequency content, then the adaptive filter would converge to some related solution which that did not contain much higher-frequency content but relied, of course, on the same system delay D.

In comparison with our results, Kates obtained MSG increases between 7dB and 30dB[12]. He used fewer adaptive FIR filter coefficients, but he artificially imposed further whiteness on the adaptive filter loop system by cascading a recursive filter with the adaptive filter. He pre-calculated the recursive filter coefficients using measurements made at eight very closely-spaced locations in the room, which were near the locations where he performed the feedback cancellation experiments. That is, he performed his experiments over a small area in the room, whereas we performed our experiments over many different locations in the room that were spaced multiple meters apart. This explains how Kates obtained slightly higher MSG increases, albeit in a less realistic configuration from a practical standpoint for PA systems.

4.5 Effects of OS Latency

Most of the experiments previously performed used DSPs to carry out the signal processing; however, our system was significantly different due to the larger delays imposed by the personal computer's ADC and DAC blocks. One could imagine that systems with lower latencies would perform better because they could catch sinusoids of increasing magnitude more quickly. Thus, we needed to verify that our results were comparable by verifying that our results were not particularly dependent on the system latencies.

For example, a particular configuration similar to the one shown in Figure 6 was chosen; however, the microphones were not both pointing directly at the speaker, which explains why the increases in MSG were not as large as 20dB. We applied the 2mic acoustic feedback cancellation algorithm using a 0.6% misadjustment and an adaptive FIR filter with 110 coefficients. The delay between the two microphones was 5ms. Table 1 shows the various increases in MSG that resulted from different OS latencies, each of which consisted of the sum of the respective ADC and DAC latencies.

OS Latency [ms]	2mic [dB]
5.8	9.2
11.6	9.4
23.2	9.3
46.4	10.5
92.9	9.9

Table 1. Various increases in MSG measured for different OS latencies

Note that the increases in MSG are all almost the same, and their differences can be attributed to the phase differences in the loop transfer functions and measurement error. This suggests that our implementation is likely comparable with previous the DSP implementations by Kates and others.

5.0 Sound Examples

5.1 APNC

We created some recordings of the input picked up by the talker microphone to help determine the practical limitations of the algorithms. The first example is of the APNC implementing 0dB of gain increase. That is, without the acoustic feedback canceller, the system would have been just barely unstable. This is evidenced by the ringing tones heard in the sound file. However, the adaptive filter is able to adapt quickly enough to prevent the signals in the loop from escalating out of control.

http://ccrma.stanford.edu/~eberdahl/Projects/FBNC/APNC.wav

The speech sounds quite clear, the larger dynamic range of the claps does not cause the feedback canceller system to loose control, and the sung glissando is only lightly distorted. This small amount of distortion is due to the fact that the misadjustment of the adaptive filter is only 0.06%. In fact, the distortion is small enough that the system could be used for speech applications; however, the MSG increase might not be large enough to justify the cost of implementation.

5.2 FBNC and 2mic

Qualitatively speaking, the recordings from the FBNC sound just like the recordings from the 2mic configuration, so only the recordings of the 2mic configuration are presented. The first example demonstrates 3dB of gain increase with again a misadjustment of only 0.06%. It sounds roughly the same as in the case of the APNC, but supports an additional 3dB of gain increase, which could make it worth implementing in some situations.

http://ccrma.stanford.edu/~eberdahl/Projects/FBNC/2micLoMisadj.wav

Finally, increasing the misadjustment to 0.6% made it possible to obtain a higher gain increase of 6dB for the recording. This time the intelligibility of the speech seems to be just as good as before while the sung tone sounds more distorted than before. This is because sung tones are more highly correlated at larger time delays than speech due to the quasi-periodicity of sung tones. These results suggest that this method for canceling feedback would indeed work well for speech, while it would be less than ideal in musical situations unless used as a sound effect.

http://ccrma.stanford.edu/~eberdahl/Projects/FBNC/2micHiMisadj.wav

(Note: Here the MSG increases are slightly lower than those cited in section 4.4 because at the time of the recordings, we had not yet realized that pointing the microphones directly at the speaker greatly increased the amount of MSG increase possible due to the increased flatness of the loop transfer functions.)

5.3 Convergence

One final test demonstrates the difficulty in predicting the behavior of various configurations when the feedback cancellation algorithm is pushed to its limits. This results from the fact that predicting the changes in the loop transfer function due to reorienting a microphone are difficult to predict. In this case, the 2mic feedback canceller was applied to the configuration shown in Figure 9 using a 0.06% misadjustment. The microphone closest to the speaker is the speaker microphone, and the two microphones placed further away are talker microphones. After every additional second in time, the system switches between using one of the talker microphones versus the other. This simulates the very fast movement of the talker microphone by the talker.

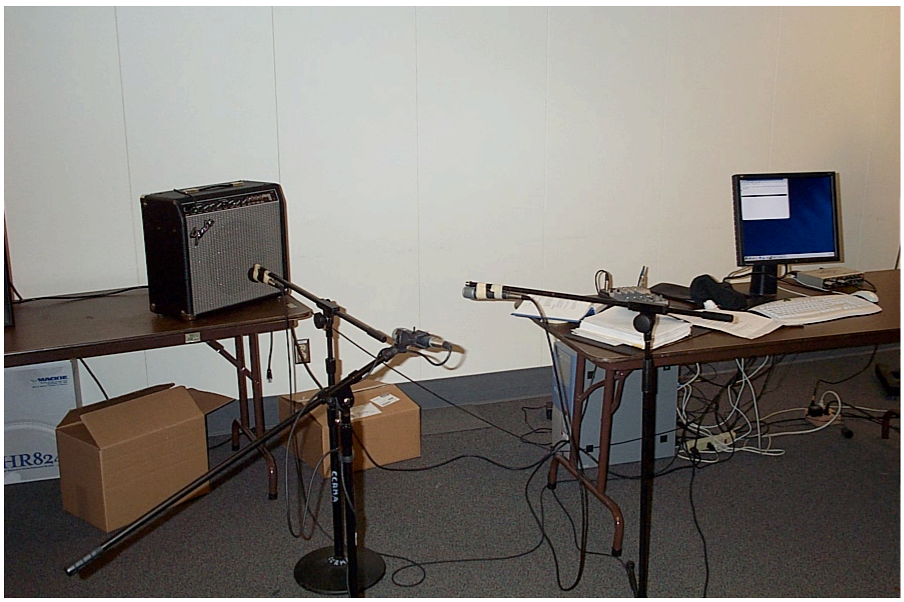


Figure 9. Microphone and speaker orientation for the convergence measurement

In this case, the positive feedback inherent in the first system cannot be subtracted out sufficiently for the adaptive filter to converge to a good solution, while the second configuration can be compensated for. The following sound file contains the results measured by the speaker microphone.

http://ccrma.stanford.edu/~eberdahl/Projects/FBNC/Convergence.wav

Figure 10 shows a spectrogram for the first four seconds of the results. It can be clearly seen how for the first configuration, a group of sinusoids starts increasing in power, is then subtracted out by the adaptive filter, but then under the new loop transfer function a new group of sinusoids starts increasing in power, which cannot be cancelled by the feedback canceller. In the second configuration, the adaptive filter simply adjusts quickly and accurately enough to gain control over the system again. (The low frequency energy shown in the spectrogram is measurement noise.)

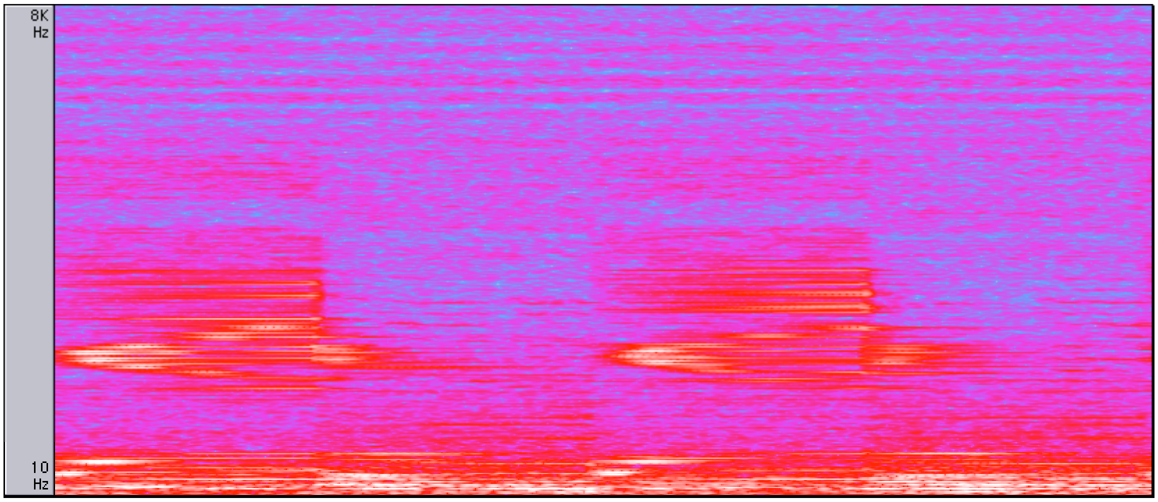


Figure 10. Spectrogram of the first four seconds of `Convergence.wav`

6.0 References

[1] Foley, J.B. *Adaptive Periodic Noise Cancellation for the Control of Acoustic Howling*, IEE Colloquium on Adaptive Filters, March 22, 1989, London, UK. Vol. 7, p. 1-4.

[2] Siqueira, Marcio G., and Abeer Alwan. *Steady-State Analysis of Continuous Adaptation in Acoustic Feedback Reduction Systems for Hearing-Aids*, IEEE Transactions on Speech and Audio Processing, Vol. 8, No. 4, July 2000.

[3] Chi, Hsiang-Feng, Shawn X. Gao, Sigfrid D. Soli and Abeer Alwan. *Band-limited feedback cancellation with a modified filtered-X LMS algorithm for hearing aids,* Speech Communication, January 2003, Vol. 39, No. 1-2, pp. 147-61.

[4] Fillon, Thomas and Jacques Prado. *Acoustic Feedback Cancellation for Hearing-Aids*, Using Multi-Delay Filter, 5th Nordic Signal Processing Symposium, October 4-7, 2002, On board Hurtigruten. Also available online: http://www.tsi.enst.fr/~fillon/.

[5] Moulines, Eric, Omar Ait Amrane and Yves Grenier. *The Generalized Multidelay Adaptive Filter: Structure and Convergence Analysis*, IEEE Transactions on Signal Processing, Vol. 43, No. 1, January 1995.

[6] Soo, Jia-Sien and Khee K. Pang. *Multidelay Block Frequency Domain Adaptive Filter*, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 38, No. 2, Feb. 1990.

[7] Jacques Prado and Eric Moulines. *Frequency-domain adaptive filtering with applications to acoustic echo cancellation*, Annales des Télécommunications, 49(7–8):pp. 414-428, 1994.

[8] Bloland, F. M. and J. B. Foley. *Filter Length Effects in the Design of Digital Adaptive Filters for Speech Applications,* Digital Processing of Signals in Communications Conference, April 22-26, 1985, Loughborough, UK, p. 9-13.

[9] Behringer Specialized Studio Technologies Website, June 19, 2005. http://www.behringer.com/DSP1124P/index.cfm?lang=ENG.

[10] Maxwell, Joseph A. and Patrick M. Zurek, *Reducing Acoustic Feedback in Hearing Aids*, IEEE Transactions on Speech and Audio Processing, Vol. 3, No. 4, July 1995.

[11] Kates, James M. *Feedback Cancellation in Hearing Aids: Results from a Computer Simulation*, IEEE Transactions on Signal Processing, Vol. 39, No. 3, March 1991.

[12] Kates, James M. *Room reverberation effects in hearing aid feedback cancellation*, Journal of the Acoustical Society of America, Vol. 109 (1), January 2001, pp. 367-378.