Noise Cancellation using Fixed Beamforming

Jacek Dmochowski, Rafik Goubran
Dept. of Systems and Computer Engineering, Carleton University
jacek@sce.carleton.ca , goubran@sce.carleton.ca

Abstract

This paper presents a noise cancellation structure with a fixed beamformer front end. Simulation results show a noise reduction of 21 dB with a speech source-noise source separation of 1m. Moreover, experimental results (real data acquired in a reverberant conference room) show a complex noise reduction pattern with a maximum noise reduction of 17 dB.

1. Background and Introduction

The adaptive noise cancellation (ANC) process [1-3] entails a scheme in which noise is subtracted from a received signal in an intelligent fashion to achieve a greater signal-to-noise ratio [4]. The operation of the classical adaptive noise canceller is depicted in Figure 1.

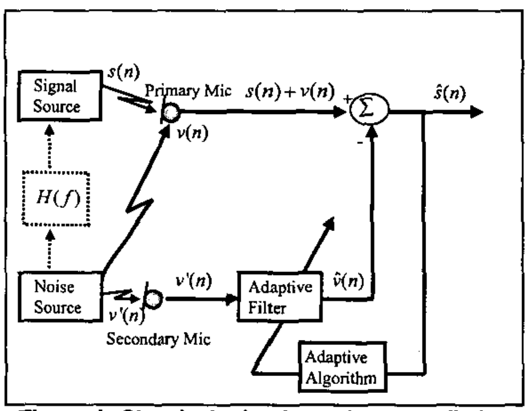

Figure 1. Classical adaptive noise cancellation

A primary sensor is located in the vicinity of the desired signal, while a secondary sensor is positioned near the origin of the unwanted noise. An adaptive algorithm drives the tap-weights of the adaptive filter towards the transfer function between the noise source and the primary sensor, such that the performed subtraction results in a close estimate of the desired signal.

The placement of the primary and secondary sensors is critical to the proper operation of the ANC. In some applications, it is not possible to place the secondary sensor near the noise source. To circumvent this problem, a modified implementation of ANC, involving beamforming via a microphone array, is proposed. The proposed scheme no longer requires the installment of a secondary sensor.

2. Proposed Approach

The proposed scheme exploits the fact that the desired signal and unwanted noise originate from disparate spatial locations, thus enabling the employment of spatial filtering. It is proposed to perform the spatial filtering using beamforming [5] (data-independent, or adaptive) by means of a microphone array. In this paper, delay-and-sum beamforming has been selected. The microphone array thus forms two "beams": a beam "pointing" in the direction of the desired signal, and another beam "pointing" in the direction of the unwanted noise. The outputs of the respective beams now feed the two inputs of the standard adaptive noise canceller. The inclusion of the beamformer has effectively eliminated the need for the installment of a noise-alone sensor. The simplicity of delay-and-sum beamforming [6] has been coupled with the proven and long-standing effectiveness of the classical ANC scheme to yield an enticing noise reduction scheme. With the cascading of two relatively simple sub-systems, the delay-and-sum beamformer and the classical adaptive noise canceller, a system with a complex beampattern has been implemented. The proposed scheme is shown in Figure 2.

It is assumed that the desired signal is speech. It is proposed to run the adaptive algorithm during

0-7803-8817-8/04/$20.00

silent periods (i.e., speech source off), with the adaptation being "frozen" as soon as the speech is resumed. Moreover, we assume that the physical dimensions of the environment permit the far-field plane wave propagation model.

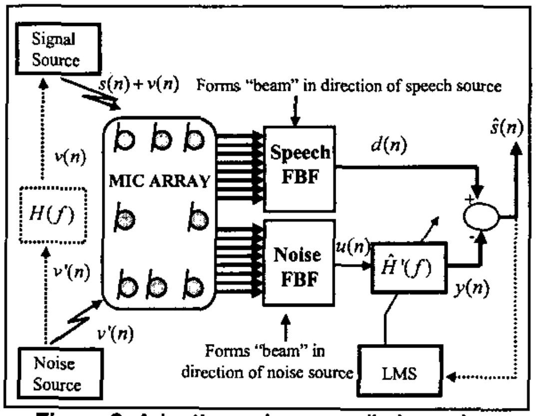

Figure 2. Adaptive noise cancellation using fixed beamforming

3. Simulation and Experimental Environment

In the discussion that follows, assume the acoustic environment shown in Figure 3. Furthermore, d_i^s denotes the distance between the speech source and element i in the microphone array. Similarly, d_i^v denotes the distance between noise source and element i in the microphone array. It then follows that τ_i^s (τ_i^v) denotes the direct path propagation delay from speech source (noise source) to microphone array element i. Assume that one can ignore any reflections on the signal path from desired signal source to the microphone array, and make the same assumption for the signal path from noise source to the microphone array.

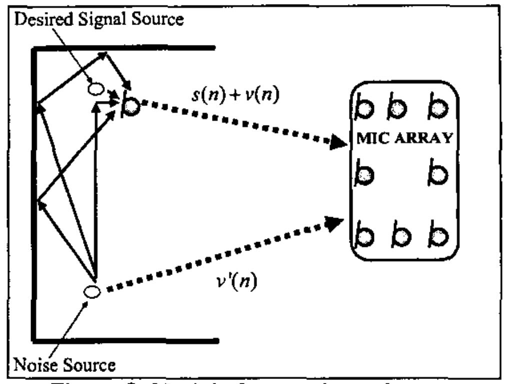

Figure 3. Model of acoustic environment

At any time, the signal received by each element in the microphone array is given by:

$$r_i(n) = \frac{1}{d_i^s}\{s(n-\tau_i^s)+v(n-\tau_i^s)\}+\frac{1}{d_i^v}v'(n-\tau_i^v) \quad (1)$$

In analyzing the error produced by the system, one must properly define what is meant by the "error". In general, the error of the system at every sample is the discrepancy between the overall system output and the corresponding original speech sample. Note that due to the inherent delaying nature of the beamforming process, a lag is introduced between the speech input and speech output. Moreover, the two modes of operation of the system, namely the silent period (speech source off) and active period (speech source on), may be examined separately.

Consider first the operation of the proposed scheme during the silent mode of operation. Delay-and-sum beamforming is accomplished by time-aligning and summing all N received signals; In particular, we note the received signal (or corresponding microphone) with the maximum propagation delay (i.e., the greatest source-microphone distance), and align the remaining the signals with it. Note that the simulations assume accurate location information (by the beamformer) pertaining to both the speech and noise sources.

Expressions for the error produced by the LMS algorithm at any time (either convergence or steady-state) are well known and given in many texts such as [7]. If the silent period is long enough to allow the adaptive filter to converge to the optimal Wiener solution, the mean-squared-error of our system at the end of the silent period is simply given by the MMSE, whose value can easily be computed from the statistics of adaptive filter input, *u(n)*, and desired signal, $d(n)$. Expressions for these two signals during the silent period can be easily determined from the definition of delay-and-sum beamforming:

$$d(n) = \frac{1}{N}\left(\sum_{i=1}^{N}\frac{1}{d_i^s}\right)v(n-\tau_{max}^s) + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_i^v}v'(n-\tau_{max}^s-\tau_i^v+\tau_i^s) \quad (2)$$

$$u(n) = \frac{1}{N}\left(\sum_{i=1}^{N}\frac{1}{d_i^v}\right)v'(n-\tau_{max}^v) + \frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_i^s}v(n-\tau_{max}^v-\tau_i^s+\tau_i^v) \quad (3)$$

Consider now the error of the proposed scheme during the active (source-on, adaptation frozen) mode of operation. During this phase, the desired signal (output of the speech beamformer) now contains a speech component. More importantly perhaps, the output of the noise beamformer now includes incoherently added speech. This can be envisioned as "leakage" of the speech signal into the noise-beam. In the best-case scenario, the positioning of the speech source will lead to this *incoherent summation* resulting in a low-power speech component in $u(n)$.

The incoherently summed speech component that leaks into the noise beam during active periods is filtered by the adaptive filter (although the filter's taps are frozen). The convolution of the filter taps with this leakage speech component is added to the error of the system. Note that the error that is present during the end of the silent period (whose variance is given by the MMSE) is also present during the active period.

An expression for the extra error term (denoted $e_{leakage}$) that appears when the speech source becomes active can be easily derived from the vector definition of convolution and definition of delay-and-sum beamforming. Letting the M filter coefficients at the time of adaptation freeze be $\mathbf{w}(N_f)$, and denoting $\tau_{max}^s = \max_i\{\tau_i^s\}, \tau_{max}^v = \max_i\{\tau_i^v\}$,

$$e_{leakage}(n) = \mathbf{w}^T(N_f)\mathbf{r}(n) \quad (4)$$

$$\mathbf{r}(n) = [r(n)\ r(n-1)\ \ldots\ r(n-M+1)]^T \quad (5)$$

$$r(n) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_i^s}s(n-\tau_{max}^v+\tau_i^v-\tau_i^s) \quad (6)$$

A point of interest is that the adaptive filter that models the transfer function between the two inputs of the adaptive noise canceller must now encompass the actions of the beamformer. Note that the adaptive filter attempts to model the impulse response between the output of the noise beam and the output of the speech beam with a finite-length (linear) FIR filter. Since fixed beamforming is comprised of two linear operations (i.e., delay, sum), it stands to reason that the beamformer will not introduce nonlinearities in the transfer function between the outputs of the two beams in the proposed scheme. However, a non-zero MMSE can result from under-modeling and any nonlinearities present in the signal environment.

4. Simulation Results

The noise reduction capability of the fixed-beamforming implementation of ANC has been assessed quantitatively by means of a computer simulation model. The model consists of an 8-element microphone array with an inter-element spacing of 10 cm and the acoustic environment shown in Figure 3.

Figure 4 depicts the performance of the proposed scheme under the parameters given above in terms of the noise reduction achieved throughout the adaptation. Note that once the speech source resumes activity, the noise is expected to rise slightly due to the leakage error. The value of this leakage error is expected to be constant throughout the active period due to (6), and thus there is no need to include the active period in the plot.

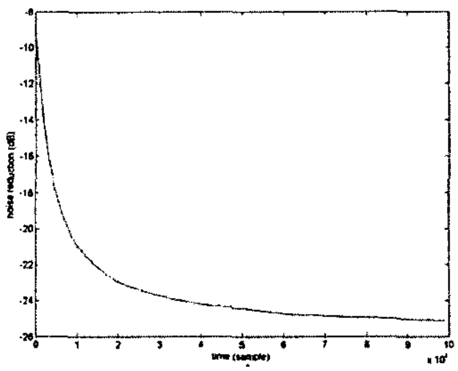

Figure 4. Noise reduction across adaptation phase of simulation

Figure 5 exhibits the performance of the fixed beamforming-ANC scheme in terms of the overall noise reduction achieved by the structure. To clarify, let P_d be the power of the signal outputted by the speech beamformer, *d(n)*, during the silent period. Let P_{output} be the power of the signal outputted by the overall system, also during the silent period. The noise reduction is then given by

$$NR = 10\log_{10}\left(\frac{P_d}{P_{output}}\right). \qquad (7)$$

This noise reduction is shown for various spatial separations between speech and noise sources.

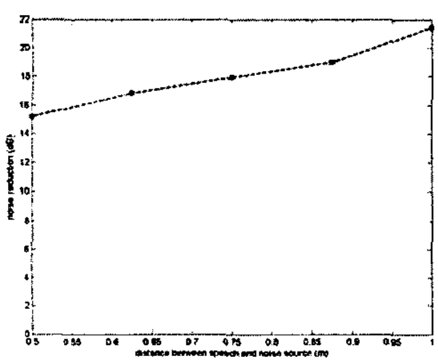

Figure 5. Noise reduction versus spatial separation between speech and noise source

As expected, the noise reduction achieved by the proposed scheme increases with increased distance between speech and noise source. Although not depicted in Figure 5, at virtually no separation between the signal sources, no noise reduction is present. With such a configuration, the output of the speech beam will be virtually identical to the output of the noise beam. Therefore, the impulse response between the two beams is given by a unit impulse function. Consequently, the output of the system will be an all-zero signal. Note that during adaptation, the desired signal is indeed the all-zero signal.

5. Experimental Results

In order to assess the noise reduction capability of the proposed scheme employing real (experimental, as opposed to simulated) data, the following procedure was followed. A white noise sequence was recorded with a single microphone – the energy of the observed noise over a fixed time interval was computed and comprises the reference value. The noise sequence was then recorded with a 6-element microphone array, with the received signals being processed according to the proposed fixed beamforming/ANC scheme. To be specific, delay-and-sum beamforming was performed twice on the received signal vector: once with respect to the position of the white noise source, and again with respect to a set location. The outputs of these two beams were then fed into the adaptive noise canceller, and the energy of the ANC output sequence was computed. The reduction in noise energy was then determined.

The noise reduction varies with the spatial co-ordinates of the set location to which the second beam is formed. By computing the noise reduction over several such locations, a noise reduction pattern was plotted. The acoustic environment in which the recordings were made is a reverberant conference room. The noise reduction pattern is shown over a 10m-by-10m rectangular slice of 3-dimensional space. The center of the circular, 6-element microphone array was positioned at the origin, while the white noise source's spatial co-ordinates are given by *(1.79m, 0.595m, 0.09m)*.

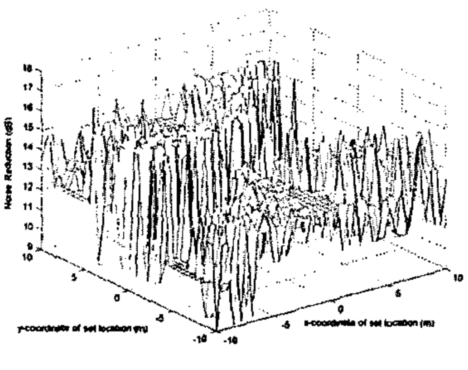

Figure 6. Noise reduction pattern over rectangular area in reverberant conference room

Note that the noise reduction is relatively minimal when our second beam is pointed to a location adjacent to the white noise source. Similarly, the noise reduction is low for locations near the actual microphone array. The intricacy of the noise reduction pattern points to the complexity of the beampattern of the proposed scheme.

6. Conclusion

The proposed ANC structure (fixed beamforming front-end) offers the benefit of not having to employ a secondary sensor, while maintaining significant noise reducing capability. Moreover, the structure avoids the drawbacks of complex adaptive beamforming structures, and yet has the potential to provide an intricate beampattern.

7. Acknowledgments

The authors would like to gratefully acknowledge the financial support from Communications and Information Technology Ontario (CITO), and Ontario Graduate Scholarships (OGS).

8. References

[1] Chang, C.Y. and Shyu, K.K, "Active noise cancellation with a fuzzy adaptive filtered-X algorithm," *Proc. IEEE, Circuits, Devices and Systems*, Vol. 150, Issue 5, Oct. 2003, pp. 416-422.

[2] Liu Dichen et al., "Study on the measurement of HV insulator's leakage current based on the adaptive noise cancellation theory," *Proceedings of Asia-Pacific Conference on Environmental Electromagnetics*, Nov. 2003, pp. 327-331.

[3] Daesung Kim et al., "Acoustic echo cancellation using blind source separation," *IEEE Workshop on Signal Processing Systems*, Aug. 2003, pp. 241-244.

[4] Widrow, B., et al, "Adaptive noise canceling: principles and applications," *Proc. IEEE*, vol. 63, 1975, pp. 1692-1716.

[5] Van Veen, B.D. and Buckley, K.M, "Beamforming: A versatile approach to spatial filtering," *IEEE ASSP Mag.*, vol. 5, no. 2, Apr. 1988, pp. 4-24.

[6] Brandstein, M. and Ward, D, *Microphone Arrays*, Springer-Verlag, Berlin, Germany, 2001.

[7] Haykin, S., *Adaptive Filter Theory*, 3rd ed., Prentice-Hall, Englewood Cliffs, NJ, 1996.