RESEARCH ARTICLE OPEN ACCESS

Noise Reduction and Echo Cancellation Using Threshold Filters in Hands Free Communication Systems

Ankita Shukla, Dr. Vineeta Saxena Nigam
DDI-PG, Department of Electronics and Communication, UTD-RGPV, Bhopal
Professor, Department of Electronics and Communication, RGPV, Bhopal
Corresponding author: Ankita Shukla

ABSTRACT:-Background noise, far-end acoustic echo, and room reverberation dramatically degrade the performance of many hands-free speech communication systems, in practical environments. For example, for automatic speech recognition system, noises result in the mismatch between the training and testing conditions, further degrading the performance of recognition system in real-world conditions. The threshold filtering can be executed with the help of mask windows of different sizes. For this, first we have to detect the noisy pixel, if the detected pixel is contaminated it can be identified by the homogeneity level of the local region around that pixel. Suppose the signal is corrupted by noise, which means the pixel value, is both 0 and 255 then we must calculate the threshold value and compare it with the neighbouring upper and lower pixels and find out which of them is homogeneous with the threshold value and replace the pixel with that value.
In addition to hands-free speech communication systems, the proposed threshold filter system in this thesis is also useful and preferable to many other applications. For example, for speech recognition systems, it is able to improve the recognition accuracy of the received speech signals in adverse environments.
Keywords:-Echo Cancellation, Noise, Threshold Filter, Signal to Noise Ratio, Error, Commutation Time

Date of Submission: 11-07-2018 Date of acceptance: 25-07-2018

I. INTRODUCTION

In Hands-free communication systems, a microphone often picks up reverberation, background noise, and acoustic echoes together with a speaker's voice which is desired speech signal. Reverberation is due to reflective acoustic environments and leads to degrade the auditory quality of speech signal. This can be remedied by using a DE reverberation technique. DE reverberation consists of recovering a desired speech signal from observed reverberant signals. Several DE reverberation approaches have been proposed [1 -4]. Generally, the research on methods of background noise reduction is being done by the two approaches. One of these is by making use of the single microphone speech enhancement techniques, and the other one is the multi-microphone techniques. Beamforming microphone arrays are very effective since suppress background noise by spatial-temporal filtering without distorting the desired speech signal. Thus the later techniques are preferred to single-microphone techniques. With full-duplex communication, echoes of the loudspeaker signals will join background noise to corrupt the desired speech signal. However, beamforming does not exploit the available loudspeaker signals as reference information for suppressing the acoustic echoes. This is accomplished by acoustic echo cancellation algorithms. In this research work, algorithms will be developed for techniques that allow for removing background noise and acoustic echo from the speech signal before further processing it.

II. ACOUSTIC ECHO CANCELLATION

In order to suppress echo, several conventional acoustic echo cancellation techniques can be applied. These techniques are based on adaptive filtering techniques. Adaptive filters are a powerful signal processing tool which can be used to model the unknown system and track possible system variations. A large set of adaptive filtering techniques has been developed during the last decades, differing in terms of performance (such as convergence speed, tracking, delay, complexity, and stability). In acoustic echo cancellation, the far-end echo path has to be modeled by the adaptive filter. The echo path is acoustic impulse response from the far-end signal emitted by loudspeaker to the microphone(s). Since this acoustic impulse response can be quite long and highly time varying, the adaptive filter will require several hundreds or thousands of filter coefficients and high-performance (fast convergence rate), but low complexity adaptive filtering algorithms are desirable. Moreover, the delay introduced by the algorithm cannot be too large. For acoustic applications, cheap algorithms, such as the least mean squares (LMS) and normalized LMS (NLMS) algorithm are typically used. However, these algorithms exhibit a slow convergence behavior, especially for colored signals such as speech. Therefore, the affine projection algorithm (APA) and its variants have been investigated. These algorithms have a better convergence behavior than NLMS algorithms but at a cost of higher computational complexity. Another class of acoustic echo cancellation based on advanced adaptive algorithms, such as recursive least squares (RLS) algorithm, has been paid more attention in recent years [2, 3]. Traditionally, noise reduction and echo cancellation have been addressed independently, either by first canceling the echo components in all microphone signals and then performing multi-microphone noise reduction, or vice-versa, by first performing multimicrophone noise reduction, followed by a single-channel echo canceller. Both schemes have their own advantages and disadvantages with respect to performance and complexity. Recently, it has been recognized that both problems are better solved using a combined approach, certainly when using multiple microphones. Initial results indicate that a combined approach yields a better performance at a lower complexity.

III. PROPOSED METHODOLOGY

Figure 1 depicts a typical hands-free speech communication environment. The primary goal of the microphone array is to record the desired speech signal uttered by the speaker. The advantage of using microphone array instead of single microphone is that it can be considered as a directional microphone which allow spatial filtering of arriving signals (desired speech signal and different noise signals) and, thus desired speech signal can be enhanced and noise signals can be suppressed. It is clear from Fig. 1.1 that in a hands-free system several types of signal degradation occur. Due to the large distance between the speaker and the microphone array, background noise is also picked up by the microphone array. The background noise typically arises from computer fans, traffic, audio equipment or other speakers present in the room. Not only the direct path signal of the speaker is recorded, but also the signal reflected against the walls, floors and any object present in the recording speech room. These components of the noise are usually termed as reverberation and the signal processing techniques for suppressing these signals are referred to as DE reverberation techniques. A specific type of noise, also depicted in Fig. 1, is far-end echo, emitted by a loudspeaker in the near-end room and coming from the far-end room in a typical hands-free teleconferencing application. Fared signals are also recorded by another set of microphone(s) in the far-end room and are sent to the near-end room, where usually the acoustic coupling between the loudspeaker and the microphone array exists. This will cause the local speaker to hear an echo or a delayed version of his/her own speech through the speaker at the far-end room.

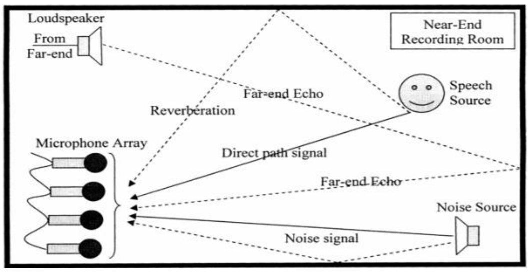

Figure 1: Typical Hands-Free Speech Communication Environment.

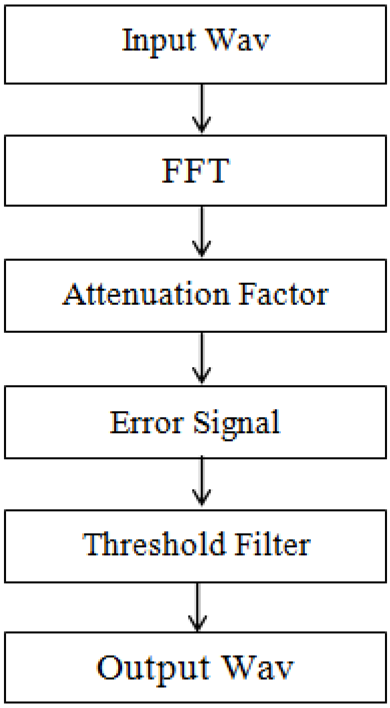

Figure 2: Block Diagram of Proposed Methodology

Threshold filter used for Echo Cancellation

We have developed the simple algorithm in which we perform the noise detection & noise removal process simultaneously. We use the smallest window size which preserves the fine details of signal. The window of size 3x3 chooses for noise detection and noise removal. The window contains total 9 elements which are as follows: Z1, Z2, Z3, Z4, Z5, Z6, Z7, Z8, Z9. First step selects the maximum, minimum and median values of columns and rows. Second step stores these values and selects minimum threshold, maximum threshold and final median value. Third step use threshold values for noise detection and final median value for noise removal.

Step-1:-

First we select all columns of filtering window one by one and then we find three values i.e. Maximum, Minimum and Median in each column. The mathematical expression can be shown as follow: The minimum values of rows and columns are represented as

Min (cln1) = min { Z1, Z4, Z7}
Min (cln2) = min { Z2, Z5, Z8}
Min (cln3) = min { Z3, Z6, Z9}
Min (row1) = min { Z1, Z2, Z3}
Min (row2) = min { Z4, Z5, Z6}
Min (row3) = min { Z7, Z8, Z9}

The maximum value of rows and columns are represented as

Max (cln1) = max { Z1, Z4, Z7}
Max (cln2) = max { Z2, Z5, Z8}
Max (cln3) = max { Z3, Z6, Z9}
Max (row1) = max { Z1, Z2, Z3}
Max (row2) = max { Z4, Z5, Z6}
Max (row3) = max { Z7, Z8, Z9}

The median value of the rows and columns are represented as

Med (cln1) = med { Z1, Z4, Z7}
Med (cln2) = med { Z2, Z5, Z8}
Med (cln3) = med { Z3, Z6, Z9}
Med (row1) = med { Z1, Z2, Z3}
Med (row2) = med { Z4, Z5, Z6}
Med (row3) = med { Z7, Z8, Z9}

Step-2:-

Now we have total nine values (three maximum, three minimum and three median). We will use these values to calculate threshold values (maximum threshold and minimum threshold) and median value. For these calculations, we make three different groups of these nine elements.

Max_group = (Max (cln1), Max (cln2), Max (cln3) Max (row1), Max (row2), Max (row3))
Min_group = (Min (cln1), Min (cln2), Min (cln3) Min (row1), Min (row2), Min (row3))
Med_group = (Med (cln1), Med (cln2), Med (cln3) Med (row1), Med (row2), Med (row3))

First we will calculate max_min by choosing maximum value in min_group and min_max by choosing minimum value in max_group. Then we choose minimum threshold by choosing minimum value in max_group.

max_min = Max (Min (cln1), Min (cln2), Min (cln3) Min (row1), Min (row2), Min (row3))
min_max = Min (Max (cln1), Max (cln2), Max (cln3) Max (row1), Max (row2), Max (row3)}
median_med = Med {Med (cln1), Med (cln2), Med (cln3) Med (row1), Med (row2), Med (row3))

Now these three values (max_min, min_max, median_med) will be furhter sorted and finally we get minimum threshold, maximum threshold and final median value as follows:

Thmax = max {min_max, median_med, max_min}
Thmin = min {min_max, median_med, max_min}
Final_med = med { min_max, median_med, max_min }

These two threshold values will be used for noise detection and final median will be used for noise removal.

Step-3:-

Now we will perform noise detection and noise removal operation using these three values i.e. Thmax, Thmin, and Med_new. We compare the central pixel with threshold values. If the central pixel is in between the Thmin and Thmax, then the pixel will be considered as noise free, then pixel will remain unchanged and window will move or slide to the next pixel. Otherwise pixel will consider as noisy and it will be replaced by median value.

If Thmin ≤Z5≤ Thmax
Then Z5 is unchanged.
Else Z5= Med_new.

Here we are parallelly calculating the threshold values and median value. So there is no need to perform noise detection and noise removal separately.

IV. SIMULATION RESULT

The signal to noise ratio (SNR) and error are used to measure the resulting parameter of wav audio signal.

Signal-to-noise ratio is defined as the ratio of the power of a signal (meaningful information) to the power of background noise (unwanted signal):

$$SNR = \frac{P_{Signal}}{P_{Noise}} \qquad (1)$$

Where P is average power. Both signal and noise power must be measured at the same and equivalent points in a system, and within the same system bandwidth.

The error is defined as,

$$Error = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[y(i,j)-x(i,j)]^2 \tag{2}$$

Where y(i,j) is the output signal and x(i,j) is the input signal.

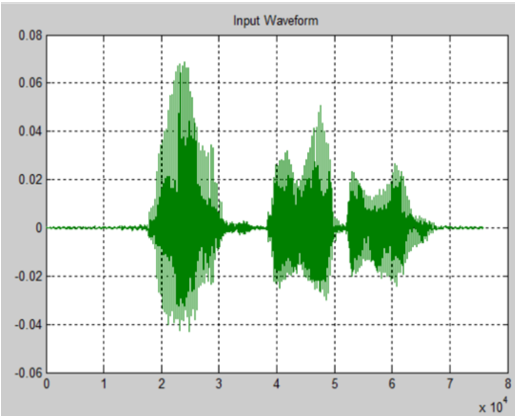

Figure 3: Input Waveform forhi.wav

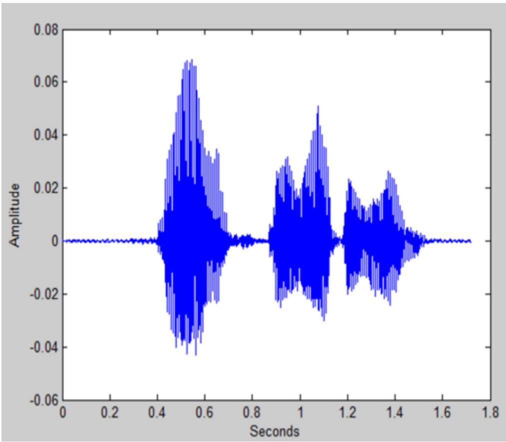

Figure 4: Seconds vs Amplitude for hi.wav

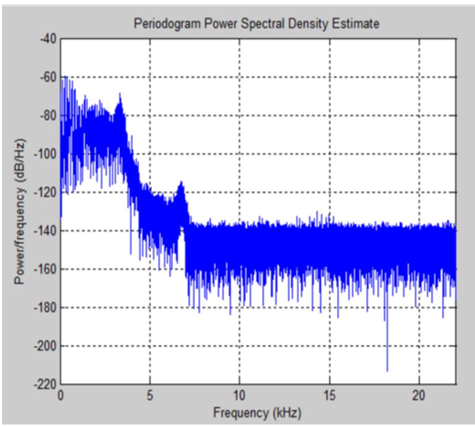

Figure 5: Power Spectrum Density for hi.wav

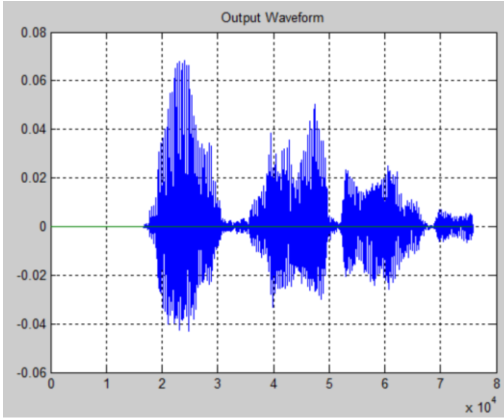

Figure 6: Output waveform with Delay

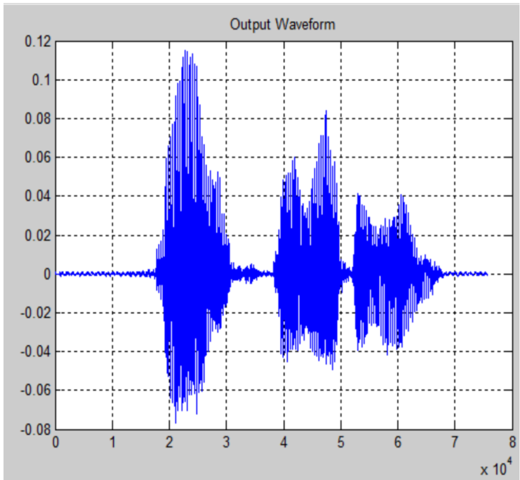

Figure 7: Output Waveform without Delay

Table 1: Simulation Result or Different Wav

Wav	Size	Error	SNR (dB)	CT (sec)
hi.wav	296 KB	0.00005	91.9941	2.452
Yoga.wav	1.31 MB	0.0010	85.3033	3.463
president.wav	519 kB	0.0012	80.5450	2.463
okey.wav	672 kB	0.0047	75.6572	2.891

As shown in table 1 the signal to noise ratio (SNR), error and computation time results are obtained for the proposed threshold filter. From the analysis of the results, it is found that the proposed threshold filter gives a superior performance.

Table 2: Comparison Result for SNR

Wav	Size	Previous Algorithm	Proposed Algorithm
		SNR (dB)	SNR (dB)
hi	296 KB	82.61	91.9941
Yoga	1.31 MB	82.61	85.3033

As shown in table 2 the signal to noise ratio (SNR), result is obtained for the proposed threshold filter and previous filter. From the analysis of the results, it is found that the proposed threshold filter gives a superior performance compared with previous filter.

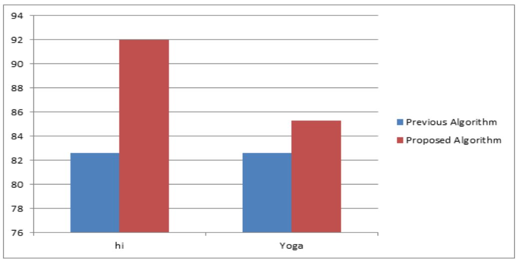

Figure 8: Bar Graph of the Previous and Proposed Filter

Table 2: Comparison Result for Error

	Previous Algorithm	Proposed Algorithm
Wav	Error	Error
hi.wav	0.00009	0.00005
Yoga.wav	0.0070	0.0010
president.wav	0.0034	0.0012
okey.wav	0.0076	0.0047

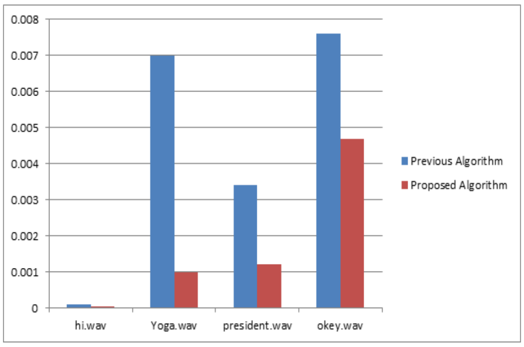

Figure 9: Bar Graph of the Previous and Proposed Filter

V. CONCLUSION

Background noise, far-end acoustic echo, and room reverberation dramatically degrade the performance of many hands-free speech communication systems, in practical environments. For example, for automatic speech recognition system, noises result in the mismatch between the training and testing conditions, further degrading the performance of recognition system in real-world conditions. For hands-free speech communication systems, background noise and far-end acoustic echo signals degrade the quality and intelligibility of received speech signal. Therefore, noise reduction and acoustic echo cancellation has been a fundamental enabling technology and indispensable components for these applications that must recognize or transmit speech in noisy environments. The results of wav clearly indicate that the quality of de-noised wav is better in visual form at that much high noise density. The proposed method improved the quality of de-noised wav especially for random valued impulse noise. SNR & Error has been calculated for the performance analysis and result shows excellent variations in the result.

REFERENCE

[1] K. Murano, S. Unagami, and F. Amano, “Echo cancellation and applications,” IEEE Comm. Mag., vol. 28, no. 1, pp. 49–55, Jan.1990.
[2] SadaokiFurui and M. Mohan Sondhi, “Advances in Speech Signal Processing”, Marcel Dekker, Inc, 1992.
[3] C. Breining, P. Dreiscitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, and J. Tilp, “Acoustic echo control. An application of very-high-order adaptive filters,” IEEE Signal Process. Mag., vol. 16, no. 4, pp. 42–69, Jul. 1999.
[4] Man Mohan Sondhi, David A. Berkley “Silencing Echoes on Telephone network”, Proc. IEEE, vol. 68, no. 8, pp. 948-963, Aug. 1980.
[5] Simon Haykin, “Adaptive Filter Theory”, Prentice Hall, Englewood Cliffs, N.J., 1986.
[6] Bellanger, Maurice G., 2001. “Adaptive Digital Filters”, 2nd edition. Marcel Dekker Inc., New York.
[7] S. L. Gay and J. Benesty, Acoustic Signal Processing for Telecommunication. Norwell, Mass.: Kluwer Academic Publishers, 2000.
[8] Moura, J. M., Flanagan, J. L., and Jayant, N. S, ”The Discipline of Signal Processing [Reflections]”, IEEE Signal Processing Magazine, vol.30, pp. 174–176 (2013).
[9] Breining, C., Dreiscitel, P., Hnsler, E., Mader, A., Nitsch, B., Puder, H., and Tilp, J, ”Acoustic echo control. An application of very-high-order adaptive filters”, IEEE Signal Processing Magazine, vol. 16, pp. 42–69 (1999).
[10] Hallack, F. S., and Petraglia, M. R, ”Performance comparison of adaptive algorithms applied to acoustic echo cancelling”, In: 2003 IEEE International Symposium on Industrial Electronics, vol. 2, pp. 1147–1150 IEEE. (2003).

Ankita Shukla "Noise Reduction and Echo Cancellation Using Threshold Filters in Hands Free Communication Systems ""International Journal of Engineering Research and Applications (IJERA), vol. 8, no.7, 2018, pp.33-37