BACKGROUND NOISE REDUCTION BASED ON WIENER SUPPRESSION USING A PRIORI SIGNAL-TO-NOISE RATIO ESTIMATE IN PYTHON

[1]JAY NEIL GAPUZ, [2]GILBEY'S JHON LADION

Electrical and Electronics Engineering Institute, University of the Philippines - Diliman
Quezon City, Philippines
E-mail: [1]Jayneil.gapuz@live.com, [2]gladion@upd.edu.ph

Abstract - Random, additive noise is a form of degradation a major problem of all analog communication system. In an audio file, listeners can hear it as "hiss" sounds that usually came from different sources. These noises came from inside the device itself and ambient noise coming from the environment which is not correlated to the signal itself. This study focuses on the suppression of background noise in a mono-channel audio file using spectral analysis. Initially, the input audio file was segmented. Then, each segment was analyzed using Fast Fourier Transform and approximation of its SNR was calculated. The signal will then be attenuated by multiplying a suppression value calculated using the Wiener suppression formula in the frequency domain. The original signal was then restored using Inverse Fourier Transform and pass through a moving-average for smoothing. Data were gathered from actual noisy audio files which were the input to the system. The researchers compared the output audio signal to the input audio signal by evaluating the signal to noise ratio of the two.

Index Terms - Background noise, audio segmentation, FFT, window, suppression rule, filter

I. INTRODUCTION

Nowadays, communication system is essential in our every life. The main function of a communication system is to actually transmit signal from one point to another point. However, one major problem the affects the performance of a communication system is noise.

Noise are unwanted signal that affects the clarity of an audio within a communication system. These noises came from different sources which can be seen everywhere. Specifically, background noise is a category of noise that are present in the background of an audio file. People usually perceive it as "hiss" sound. There are actually lots of studies that focuses on the reduction or suppression of background noise. One these studies is the use of short-time spectral attenuation. Background noise suppression is an important step in the recovery of the intelligibility of the audio in the presence of strong background noise. This is also an important part of a microphone.

Short-time spectral attenuation analyzes a signal's characteristic using short-time Fourier transform segment by segment and sets a suppression level based on the signal-to-noise ratio of each segment [1]. A study done by Pascal and Jozue shows the different suppression rule based on the SNR value of a signal. The study focuses on the problem of a single microphone frequency domain noise reduction in noise environment [2]. The study done by Pascal and Jozue tabulates different suppression estimate. The researchers chose the Wiener suppression rule wherein the suppression value is proportional to the SNR estimate of the input signal.

II. PROPOSED SUPPRESSION SYSTEM

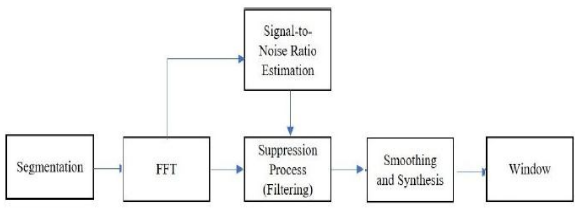

Figure 1. Proposed system Block Diagram

The system is composed of three major processes: the segmentation, suppression or the filtering and the smoothing and synthesis. The input signal was segmented into frames. FFT was used to analyze the spectrum of the signal and to estimate the SNR of the signal. A suppression rule was used and each component of the spectrum was multiplied to an array proportional to the suppression. The process will filter out additive noise since multiplication in the frequency domain is convolution in the time domain which is actually a frequency selective process. The frames were then multiplied to a window to further improve the signal. Lastly, the segments were synthesized. The smoothing process and the windowing method was an optional part of the system to further improve the signal.

A. Segmentation and SNR Estimation

Let x(t) be the noisy audio signal, s(t) be the intelligible signal and n(t) be the background noise added to the signal. We can say that input signal is actually given by Eq. 1.

$$x(t) = s(t) + n(t) \quad (1)$$

In the analysis of the signal, the FFT of the audio input is also equal to the FFT of the signal and the noise component itself. Hence, in order to remove the noise, an attenuation factor must be multiplied to reduce the level of the signal to a significant amount. For efficient analysis, segmentation of the audio is necessary to limit the number of samples to be analyze.

The signal was segmented into frames with a length of 20 millisecond. Each frame has a 10-millisecond overlap. Zero padding was done to make the length of each frame the same. Each segment was analyzed and an estimate value of the segment's signal-noise ratio (SNR) was determined.

$$SNR = \sum E(X,k) / \sum N(X,k) \quad (2)$$

$$thresh = \sigma + \bar{x} \quad (3)$$

Let E(x,k) be the estimated energy of the signal based on the energy of the frequency bins with value above the threshold value, thresh, and N(X,k) be the energy of the noise floor which will be the frequency bins below the threshold value, thresh. Eq. 2 was used to estimate the SNR of the signal. The SNR of the signal is the ratio of the sum of the energy of the peak frequencies and the sum of the energy of the noise floor. Eq. 3 shows the value of the threshold which is the sum of the mean, $\bar{x}$, and the standard deviation, σ, of the frequency bins of the segment. Hence, the energy of the signal was calculated using the Rayleigh's energy theorem shown in Eq. 4.

$$Energy = (1/N) \sum |X(n)|^2 \quad (4)$$

The signal was first converted into a set of Fourier transform for the spectral attenuation using the FFT function.

B. Suppression and Filtering

Let G_0 be the suppression value or attenuation value.

$$G_0 = (SNR - 1) / SNR \quad (5)$$

The suppression equation used by the researcher is shown in Eq. 5. This equation was originally the Wiener estimate of the gain value in the suppression rule.

The idea is to multiply each frequency component to a suppression array generated as function of the SNR. The noise components will be suppressed to a negligible. G_0 was multiplied to the frequency bin being considered as noise component and a factor of 1 was multiplied to the frequency bin being considered as intelligible signal component.

The process of filtering the background noise was accomplished by suppressing the noise spectra through the multiplication process in the Fourier domain. This act is equivalent to the convolution process in the time domain. Hence, the convolution process is the basis of the filtering method and is a frequency selective process.

C. Smoothing and Synthesis

The researcher transformed each frame back to the time-domain by implementing inverse Fourier transform to each frame. All of the frames were then synthesized. The selection in the overlapping region was based on the average of the two points obtained from segmentation process; this creates a smoother output when compared to selecting from the points. The synthesized signal was passed to a band pass filter basing from the type of voice signal used (narrowband or wideband), then performed moving average among adjacent points in an equal length. This further eliminates low and high frequency noise present in the signal.

III. RESULTS AND DISCUSSION

The researcher took samples of audio file from wavsource.com, a collection of actual recording. The selection of the audio file was based on whether the file had a "hiss" sound or background noise. There are five audio files selected from the library. Each file was segmented in 20 ms segment. For practical implement, the researchers Fast Fourier transform. The FFT of the signal was analyzed and the energy of the segment, noise energy (by the use of Rayleigh's Energy Theorem) and an estimate of the SNR were calculated per segment.

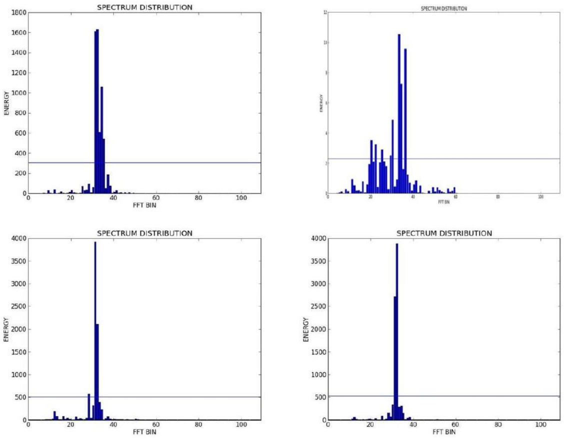

Figure 2. Frequency Spectrum per Segment

Figure 2 shows the frequency spectrum in different segments of one of the audio file. Notice that in Figure 2, the frequency bins above the threshold, represented by the horizontal, were considered part of the signal's total energy and the frequency bins below the threshold were part of the noise component. The signal's energy also varies per segment as well as the noise power. Hence, Table 1 shows the calculated energy, noise energy and estimated SNR value of the segments in one of the audio file.

Table 1. Energy per segment.

Segment	$\sum E(X,k)$	$\sum N(X,k)$	SNR (dB)
1	20.97876	6.161474	5.32095
2	38.26901	19.05076	3.02934
3	41.27603	22.67726	2.60107

4	49.3178	16.7931	4.67872
5	52.71654	10.9991	6.80587
6	49.47636	16.15379	4.86123
7	53.26248	19.02077	4.47193

The suppression value, G_0 was calculated from the value of the estimated SNR. Table 2 shows the suppression value of every segment.

Table 2. Suppression value per segment

Segment	Suppression Value, G_0
1	0.702699
2	0.502188
3	0.450595
4	0.659492
5	0.791353
6	0.673505
7	0.642886

The researcher observed that the value of G_0 is independent from each segment. Hence, applying a suppression in one segment does not affect the other segments. An array of number was generated that will be multiplied to the FFT of the segment to filter out the noise components of the segment. Figure 3 shows a sample spectrum of a segment before and after filtering or suppressing the noise components. The researcher observed that the total energy content of the segment was reduced. With the suppression value introduced, the noise components were not actually reduced to zero. There was a significant loss in the noise energy after the suppression.

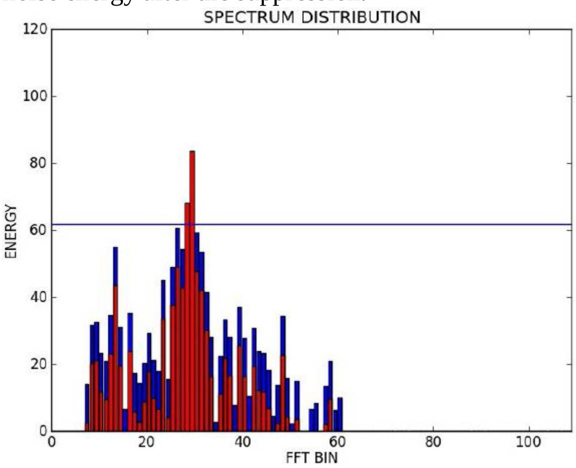

Figure 3. Segment spectrum (dB) before (blue) and after(red) the suppression.

The suppressed segment was converted back to the time-domain using inverse Fourier transform with the same number of point as the segment and passed through a moving-average system for smoothing. The output segments were then synthesized to produce the filtered audio file. Figure 4 shows the speech waveform of the original audio file and the filtered audio file. The original audio file contains additive background noise which can easily be distinguished as amplitude spikes while the filtered audio file, on the other hand, does not have any amplitude spike. The researchers try to compare the two audio files by listening to it.

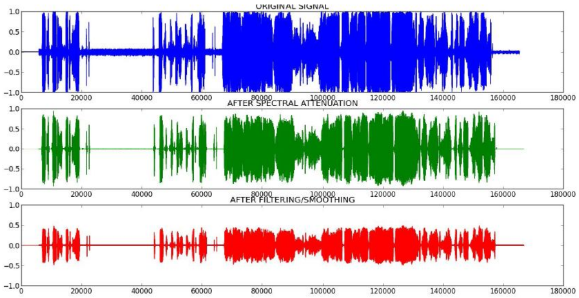

Figure 4. Original audio file waveform (upper), Suppressed audio file waveform (middle) and the Smoothened audio file waveform (lower)

The proposed system produces good quality audio. The audio file was enhanced. The five audio files used by researchers gave a good quality output. However, there were some changes in the quality of the audio. Problems were encountered by the researchers such as generation of silent part and, in some audio files, some part of the speech was also filtered out. Hence, the researcher improved it by adapting the selection of the relevant energy FFT array and the silent discriminator.

CONCLUSION

In this paper, a proposed system for suppressing background noise in the frequency domain was implemented. The researcher presented a method to filter out background noise by applying a suppression rule in the frequency spectra. Good quality restored audios were observed. The results obtained were significant in the field of speech enhancement. Moreover, this method was simple implementation to solve the problem of background noise.

ACKNOWLEDGMENT

The authors would like thank Professor Franz De Leon in sharing his knowledge and expertise in the field Digital Signal Processing.

REFERENCES

[1] M. Kahrs and K. Brandenburg. “Applications of Digital Signal Processing to Audio and Acoustics”. Kluwer Academic Publisher, 2002, pp. 163–169.

[2] J. Ma and P. C. Loizou, “SNR loss: A new objective measure for predicting the intelligibility of noise-suppressed speech”, SciencDirect: Speech Communication, 2010, pp 340–354.

[3] P. Scalart and J. Vieira Filho, “Speech enhancement based on a priori signal-to-noise estimation, IEEE Trans, 1996, pp. 629–632.

[4] E. Verteletskaya and B. Simak, “Noise reduction based in modified subtraction method”, IAENG International Journal of Computed Science, February 2011.

[5] O. Capp`e. “Elimination of the musical noise phenomenon with the Ephraim and Malah Noise Suppressor”. pp. 345–349.

★★★