Speech Enhancement Based on Spectral Subtraction Involving Magnitude and Phase Components

Miss Bhagat Nikita [1], Miss Chavan Prajakta [2], Miss Dhaigude Priyanka [3], Miss Ingole Nisha [4], Mr Ranaware Amarsinh [5]

[1] *Final Year BE Student at College of Engineering Phaltan, Maharashtra, India*
[2] *Final Year BE Student at College of Engineering Phaltan, Maharashtra, India*
[3] *Final Year BE Student at College of Engineering Phaltan, Maharashtra, India*
[4] *Assistant Professor, E&TC Department PES's College of Engineering, Phaltan, Maharashtra, India*
[5] *Assistant Professor, E&TC Department PES's College of Engineering, Phaltan, Maharashtra, India*

---***---

Abstract - *In this paper, we developed a new method to improve the performance of noisy speech signal by using speech enhancement technique based on single-band spectral subtraction. Spectral subtraction is used to remove noise from noisy speech signals in the frequency domain. This method consists of the spectrum of the noisy speech using the Fast Fourier Transform (FFT) and subtracting the calculated average magnitude of the noise spectrum from the noisy speech spectrum. We applied spectral subtraction to the noisy speech signal. The denoise algorithm was implemented using Matlab software by storing the noisy speech data into half overlapped (overlap add processing) and calculate the corresponding magnitude spectrum using the FFT, removing the noise from the noisy speech, and reconstructing the speech back into the time do-main using the inverse Fast Fourier Transform (IFFT).*

***Key Words*: Speech Enhancement, Spectral subtraction, FFT, Noise Estimation, overlap add processing, IFFT.**

1.Introdction

In speech enhancement, the number of noise are added into original clean speech signal. There are number of methods to remove noise signal. Telephones are increasingly being used in noisy environments such as cars, airports, streets, trains, station. So, we try to remove the noise using spectral subtraction method. The aim of this project is to implement a real-time system that will reduce the background noise in a speech signal. This process is called speech enhancement. In many speech communication systems, background noise causes the quality of speech to degrade. In most of speech processing applications such as mobile communications, speech recognition and hearing aids removing the back- ground noise in a noisy environment is inevitable. So, speech enhancement as a necessity for related applications has been widely studied in recent years. The main objective of speech enhancement is to reduce the noise from noisy speech, such as the speech quality or intelligibility. It is usually difficult to reduce noise without distorting speech and thus, the performance of speech enhancement systems is limited between speech distortion and noise reduction.

There are several techniques such as Wiener filtering, wavelet-based, adaptive filtering and spectral subtraction is still a useful method. In spectral subtraction method, we have to estimate the noise spectrum and subtract it from noisy speech spectrum. In this method, three following conditions are assumed: (1) the noise is additive (2) speech signal and noise are uncorrelated (3) one channel is available. In this paper, we try to reduce the estimation error of noise spectrum for enhancement of corrupted speech with spectral subtraction technique. We Proposed method has been tested on real speech data by computer simulation in MATLAB environment. Real speech signals from Speech Data database were used for experiments. Then we propose a method to reduce the difference between estimated noise spectrum and noise spectrum.

1.1 Single Channel Speech Enhancement

In general, single channel systems constitute by depending on different statistics of speech and unwanted noise that, work in most difficult situations where no prior knowledge of noise is available. Usually the methods assume that the noise is stationary when speech is active. They normally allow non-stationary noise between speech activity periods but in reality, when the noise is non-stationary, the performance of speech signal is dramatically decreases.

1.2 Spectral Subtraction Basic

Many different algorithms have been developed for speech enhancement: the one that we use is known as spectral subtraction algorithm. This technique operates in the frequency domain and makes the assumption that the spectrum of the input signal can be expressed as the sum of the speech spectrum and the noise spectrum. The procedure is illustrated in the diagram below and contains two important parts:

1. Estimating the spectrum of the background noise

2. Subtracting the noise spectrum from the noisy speech

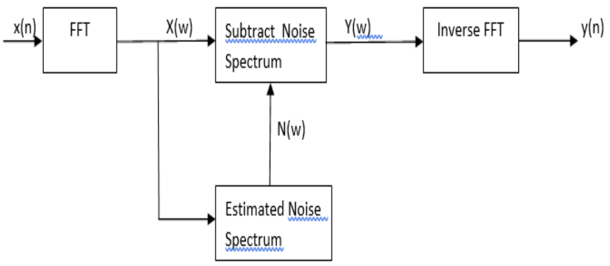


Fig 1: Spectral Subtraction Method

Suppose x(n) is a noise corrupted input speech signal, x(w) is spectrum of noisy signal, N(w) is spectrum of estimated noise and Y(w) is the spectrum of processed speech. y(n) is the original clean speech signal. so, the processed spectrum of clean speech signal can be represented as:

$$Y(w) = X(w) - N(w)$$

1.3 Block Diagram of Speech Enhancement Using Spectral Subtraction

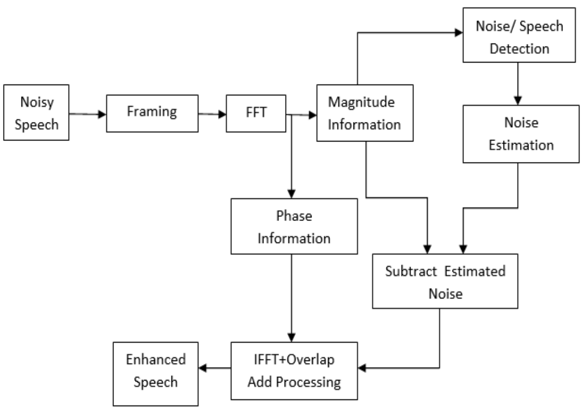


Fig 2: Block Diagram of speech Enhancement

Overlap-Add Processing

The Overlap add method is used to break continues signals into smaller segments for easier processing. The overlap Add method is based on the: (1) decompose the signal into simple components, (2) process each of components, (3) recombine the processed components into the final signal. To perform frequency-domain processing in spectral subtraction, it is necessary to split the continuous time- domain signal up into overlapping chunks called frames. After processing, the frames are then reassembled to create a continuous output signal. Each frame is 50 % overlapped. FFT is applied to each frame.

FFT's are generally used to determine the frequency components of a signal buried in a noisy time domain signal. The functions Y = FFT(x) and y = IFFT (X) implement the transform and inverse transformation given for vectors of length N by:

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)}$$

$$x(j) = (1/N)\sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)}$$

Where,

$$\omega_N = e^{(-2\pi i)/N}$$

The proposed noise reduction method was compared to spectral subtraction based on noise level reduction and improvement introduced in speech signal. Evaluate the quality of noisy speech enhanced by spectral subtraction and proposed noise reduction method.

Estimating the noise spectrum:

The noise spectrum cannot be calculated previously, but can be estimated during period when no speech is present in the noisy speech signal. When someone is speaking, he/she definitely have to pause for breath from time to time. We can assume these gaps in the speech to estimate the background noise.

We calculate the average magnitude that has been calculated in any frame over the past few seconds or so. We calculate this magnitude under the assumption that 'no one ever talks continue for more than ten seconds without a break', this spectral minimum corresponds to the minimum noise amplitude.

Subtracting the Noise Spectrum:

After subtraction, the all values of spectral magnitude are not possible to be positive. There are some possibilities to remove the negative components. An inverse Fourier transform, using phase components directly from Fourier transform unit, and overlap add is then done to rebuild the speech estimate in the time-domain.

The basic idea is to subtract the noise from the input noisy signal:

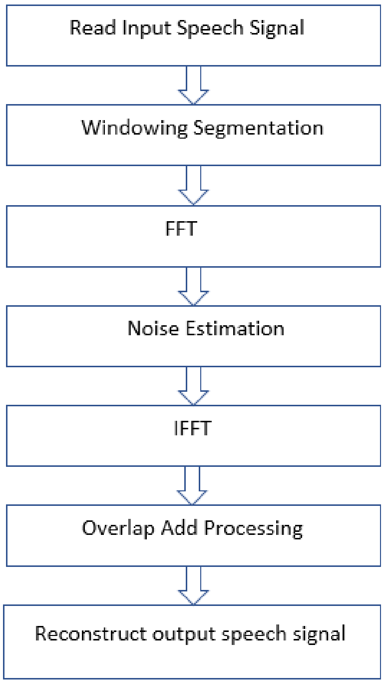
$$Y(w) = X(w) - N(w)$$

Unfortunately, we do not know the exact accurate phase of the noise signal so we subtract the magnitudes and leave the phase of X alone. After subtraction, the spectral magnitude is not guaranteed to be positive. There are some possibilities to remove the negative components present in spectrum.

In order to transform the frequency domain signal to time domain signal, the phase of the noisy speech signal is combined with processed magnitude spectrum, and then Inverse Discrete Fourier Transform (IDFT) is applied and overlap add is then done to rebuild the speech estimate in the time-domain.

2. Propose workflow and Simulation Results

2.1 Flowchart



The above flowchart shows the proposed methodology, it consists of collecting the data of noise speech and then it passes through the window to removes the artifacts present in the Noisy signal and then by using FFT algorithm extracting the phase and magnitude of the noisy speech signal, in this methodology we are concentrating on magnitude part of the noise signal to make it distinct pure speech signal. By using spectral subtraction method noise is estimated and subtracted with desired magnitude value. Then applying IFFT algorithm and overlap add processing to get the pure speech in time domain.

2.2 Simulation Results

The noises used for experiments are following: AWGN noise generated by computer, train noise, and car noise. Each frame is 50 % overlapped. FFT is applied to each frame. In addition, subjective speech quality and intelligibility evaluation tests have been held to evaluate the quality of noisy speech enhanced by conventional spectral subtraction and proposed noise reduction method. The difference noise reduction methods based on spectral subtraction and speech distortion introduced by processed waveforms and spectrogram. Three types of noise were used to generate noisy speech signals with the different noise level (SNR= [0dB 5dB 10dB]).

1.Result of inbuilt audio file in MATLAB

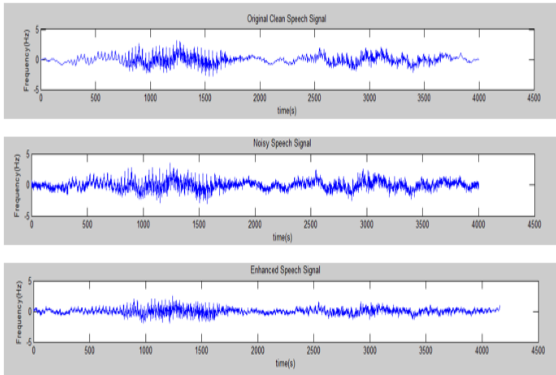
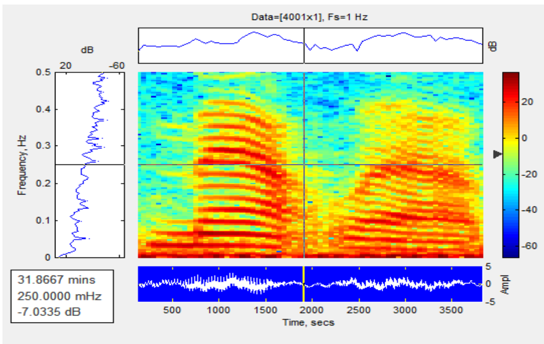
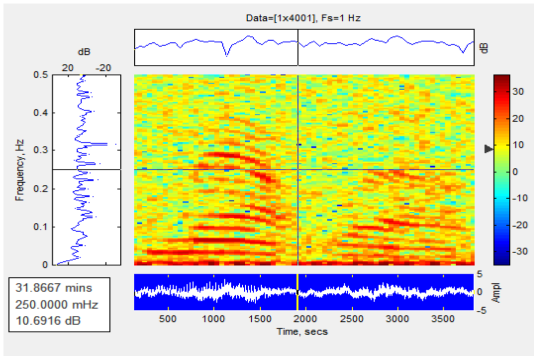


Fig 3: speech enhancement results:
(a) Clean speech
(b) Noisy speech
(c) Enhanced speech by SS method

2. Spectrogram results:



Fig(4) Spectrogram of clean speech signal



Fig(5) Spectrogram of noisy speech signal

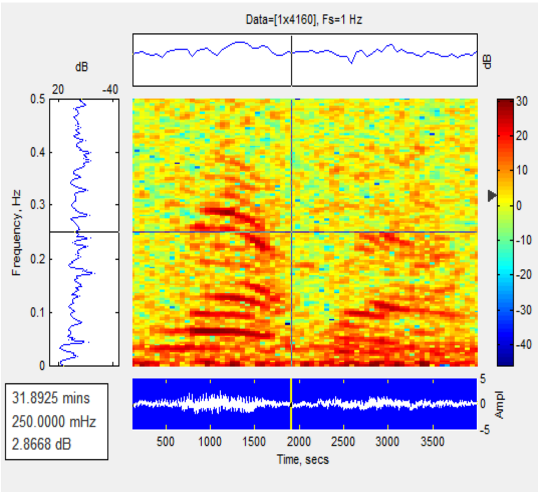


Fig (6) Spectrogram of Enhanced speech signal

The above all spectrogram shows different color. where the Red and yellow color shows the two different intensity levels. The entire all input noisy speech signal is filled with color of varying intensity because of the presence of noise. followed by yellow being high, red being very intermediate intensity, and black being the least with almost zero intensity. From above all the results During the silence period of the speech signal, the effect of surround of noise will be more. On comparing Figure 5 with Figure 6, the effect of noise is reduced.

2.. Result of database of car noise at 10dB

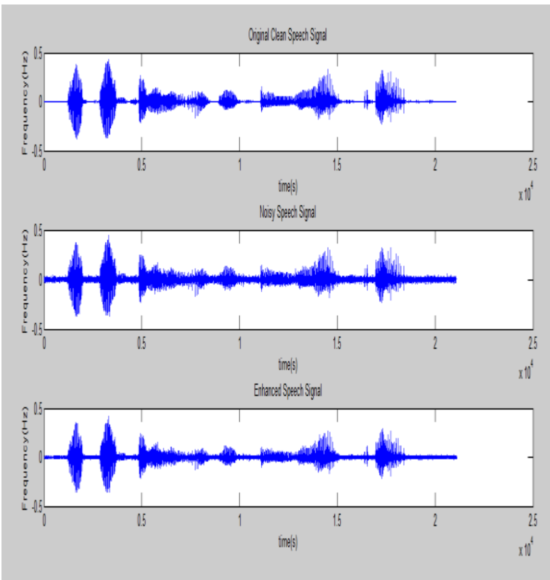
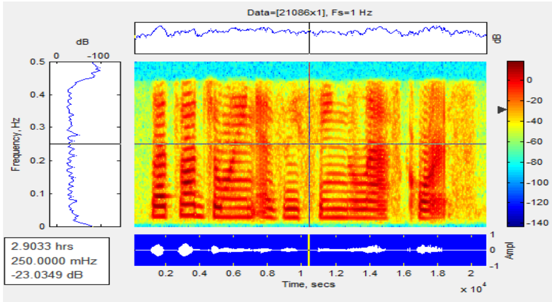


Fig 7: speech enhancement results:
(a) Clean speech
(b) Noisy speech
(c) Enhanced speech by SS method

Spectrogram results:



Fig(8) Spectrogram of clean speech signal

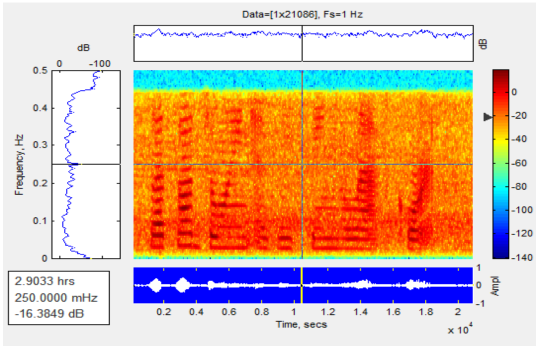


Fig (9) Spectrogram of noisy speech signal

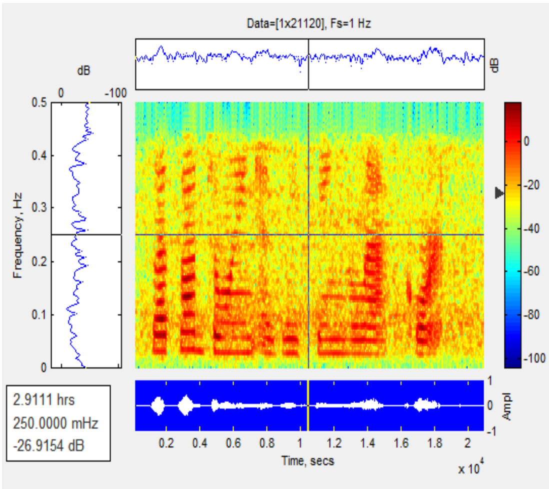


Fig (10) Spectrogram of Enhanced speech signal

On comparing Figure 9 with Figure 10, the effect of noise is reduced. Another interesting observation of the proposed algorithm is witnessed by comparing the spectrogram results obtained in Figure 9 and where there is more number of yellow lines in the female speech signal during the higher frequency range. This verifies that the female audio will have higher frequency content when compared to the male audio.

3. Results of train noise and Comparison of noise level in dB for Female and male speech signal

female speech signal	I/P Noise Level (dB)	O/P Noise Level (dB)	Reduced level of noise
Train Noise (at 0 dB)	3.6448	-4.852	8.4968
Train Noise (at 5 dB)	-15.1722	-17.0717	1.8995
Train Noise (at 10 dB)	-25.5382	-41.2185	15.6803

Male speech signal	I/P Noise Level (dB)	O/P Noise Level (dB)	Reduced level of noise
Train Noise (at 0 dB)	-3.9138	-6.2895	2.3757
Train Noise (at 5 dB)	-22.4948	-19.3090	3.1858
Train Noise (at 10 dB)	-26.1571	-31.5079	5.3508

Table -1: Reduced Noise level

From Table 1, it can be observed that the input and output noise level are different for male and female noise. This proves that there is heavy influence of noise in the speech. The improvement in noise reduction level. This spectral subtraction method makes the speech signal audible but along with the noise. The noise has not been reduced completely and it can be verified when comparing the input noise level and output noise level.

Noise level of the input speech signal, output speech signal and the improvement in reduction in noise level has been compared using a tabular form. Table 1 shows the comparison of noise level in dB for male and female speech signal.

3. CONCLUSIONS

In this paper speech enhancing method based on improved spectral subtraction algorithm is introduced. It can be seen from the experimental results that proposed method effectively reduces background noise in comparison with commonly used spectral subtraction type algorithm. This method makes the speech signal audible but along with the little bit noise. The noise has not been reduced completely and it can be observed when comparing the input and output noise level. Simulation results show that the algorithm works efficiently in reducing the background noise, which when mixed in original clean speech signal. Currently this algorithm used for stationary sound. This could further be modified and extended to make it work for non-stationary noise. From this method to reduce the background noise up to 70% and we can implement this technique into many embedded systems related to speech processing or communication purpose. For a good comparison, we have shown the results and the spectrograms of clean, noisy and processed speech. Clean speech sentence that is processed here, has been chosen from database.

ACKNOWLEDGEMENT

We have taken efforts to complete this speech enhancement software based project. Primarily, we submit our gratitude and sincere thanks to our guide Prof.Ms N.N. Ingole, for their constant motivation and support during the course of the work in the last six month. We truly appreciate and value their esteemed guidance and encouragement from the beginning to the end of this work. We also thankful to Prof. S. Badigare for supporting us in conception of project work as well as providing guidelines time to time.

REFERENCES

[1] S. Kamath, and P. Loizou, A multi-band spectral subtraction method for enhancing speech corrupted by colored noise, Proceedings of ICASSP-2002, Orlando, FL, May 2002. R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[2] S. F. Boll, Suppression of acoustic noise in speech using spectral subtraction, IEEE Trans. on Acoust. Speech & Signal Processing, Vol. ASSP-27 , April 1979, 113- 120.

[3] M. Berouti, R. Schwartz, and J. Makhoul, Enhancement of speech corrupted by acoustic noise, Proc. IEEE ICASSP , Washington DC, April 1979, 208- 211.

[4] H. Sameti, H. Sheikhzadeh, Li Deng, R. L. Brennan, HMM-Based Strategies for Enhancement of Speech Signals Embedded in Nonstationary Noise, IEEE Transactions on Speech and Audio Processing, Vol. 6, No. 5, September 1998.

[5] Yasser Ghanbari, Mohammad Reza Karami-Mollaei, Behnam Ameritard "Improved Multi-Band Spectral Subtraction Method For Speech Enhancement" IEEE International Conference On Signal And Image Processing, august 2004.

[6] Pinki, Sahil Gupta "Speech Enhancement using Spectral Subtraction-type Algorithms" IEEE International Journal of Engineering, Oct 2015.

[7] S. Kamath, A multi-band spectral subtraction method for speech enhancement, MSc Thesis in Electrical Eng., University of Texas at Dallas, December 2001.

[8] Ganga Prasad, Surendar "A Review of Different Approaches of Spectral Subtraction Algorithms for Speech Enhancement" Current Research in Engineering, Science and Technology (CREST) Journals, Vol 01 | Issue 02 | April 2013 | 57-64.

[9] Lalchhandami and Rajat Gupta "Different Approaches of Spectral Subtraction Method for Speech Enhancement" International Journal of Mathematical Sciences, Technology and Humanities 95 (2013) 1056 – 1062 ISSN 2249 5460 .

[10] Ekaterina Verteletskaya, Boris Simak "Enhanced spectral subtraction method for noise reduction with minimal speech distortion" IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing .

[11] D. Deepa, A. Shanmugam "Spectral Subtraction Method of Speech Enhancement using Adaptive Estimation of Noise with PDE method as a preprocessing technique" ICTACT Journal Of Communication Technology, March 2010, Issue: 01.

[12] M. Berouti, R. Schwartz, & J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," Proc. ICASSP, pp. 208-211, 1979.