Analysis of Speech Enhancement Incorporating Speech Recognition

Jyoti Londhe[1]

Dept. Of Electronics and Telecommunication Engineering, TPCT's collage of engineering , Osmanabad, Maharashtra, India.

---***---

Abstract — In this paper review work, we have tried to develop a novel way to improve the structure of a speech test during a speech, and to measure clear speech will be removed from the confusing speech limit. Automatic speech recognition consists of a sound speech spectrum using Fast Fourier Transform as well as a thoughtful reduction in the volume of the audio sound from the audio sound. Using the MATLAB software the audio reduction algorithm is developed to store audio speech data in half-time and calculate the spectrum of various sizes using Fourier Transform, reduce noise in audio speech, and recreate retrospective speech during high performance. Fast Fourier (FFT).

Keywords – Speech and voice Enhancement, Spectral noise Reduction, FFT, Sound Measurement, IFFT, Spectrograms.

Introduction

A clean and clear speech signal is linked to the amount of sound in speech development. There are many ways to make a speech signal without a clamor signal. In many places Phones and cell phones get noisy air Domains like cars, airports, roads, trains, stations. Therefore, we attempt to eliminate the clamor signal by using a spectral reduction method. The main purpose of this paper is for real-time device to reduce or decrease background sound with a visual speech signal, this is called speech development. Variety of languages of speech are present, in that background noise lowering speech. Applications like mobile communication can be learned a lot in recent years, speech improvement is required. The purpose of this speech development has to reverse the noise which input from a noisy speech, like as the speech phase or accessibility. It is often hard to remove the background sound without interrupting the speech, therefore, the exit of the speech enhancement system is not allowed between speech contradiction and noise reduction.

There may be other techniques same as to Wiener filtering, wavelet-based, dynamic filtering and optical output remain a useful method. In order to reduce the spectral, we must measure the clamor spectrum and reduce it from the clamorous acoustic spectrum. Completely this approach, there are the following three scenarios to consider: sound adds speech signal and sound is not related to a single channel in the market. During this paper, we attempted to reduce the audio spectrum in order to improve distorted speech by using spectral output. We have described the method tested in the actual speech data frame in the MATLAB area. The signals we receive from Real speech signals are a website used for various tests. Then we suggest how to remove noise between the average noise level and the spectrum noise.

In general, only one medium system is set up based on a variety of speech data and unwanted screaming that, it works in difficult situations where no previous clamor intelligence is available. Genres often assume that sound is stable whenever the speech is alert. They usually allow for disturbed sound during speech operations but in actual manner, when the sound is not moving, the performance of the speech signals is greatly increased.

I. WAYS TO REDUCE NOISE

A. How to Remove the Spectral

Many additional variations of the various conclusions are designed to improve speech. The one we are using is thought to be the end of the phantom release.This type works within the scope of the spread and creates the hope that the help cycle is said to be due to the phantom additon sound and clamor phantom. The action is specified within the image below and contains two main components.

B. Convolutional Denoising Autoencoder (CDAE)

Convolutional Denoising Autoencoder (CDAE) Promotes the same function of default encoders by adding convolutional encoding and decoding layers. CDAE has a 2D layout in the dialog and adjusts the input to the 2D alignment structure using the embedded space and the desktop as shown in the picture. Thus, in the study we proposed, So from this method we can propose that the CDAE is best suited for speech development which is shown in Eq.

(3), on anyother map included,

$$hi = f(Wi * \alpha + ai) \quad (3)$$

where, the process of conversion and the value of bias is Measuring the background of in the spectrum.

Reduce to spectrum noise from the sound speech spectrum

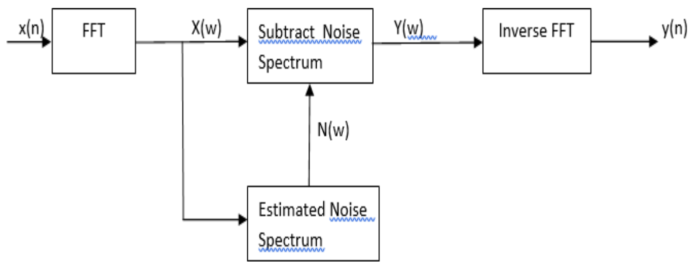


Fig 1: Spectral Subtraction Method

Assuming that x (n) can be sound and the modified speech signal is considered to be x (w) is a sound signal spectrum will be, N (w) the spectrum of rated intervals and the range of speech processed by Y (w) and the pure speech signal the first is y (n). therefore, the processed spectrum as to the unpolluted speech signal will be provided as follows:

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

Next Drawing Speech Block Impression By Spectral Removal.

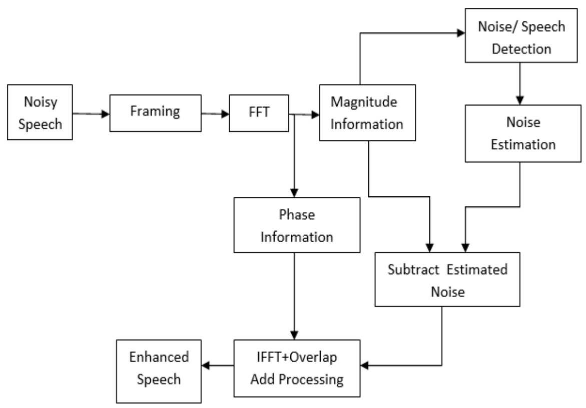


Figure 2: Speech development block diagram

Here, we can create a continuous output signal. Every frame is 50% overwhelmed. FFT (Finite Fourier transform) is used in every frame.

FFTs are often useful in determining a signal hidden in a noisy time zone in parts of frequency. The functions Y = FFT (x) and y = IFFT (X) suggest the alternating rotation given the vectors length N by

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

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

Here,

The proposed method of noise reduction is provided for visual removal that relied on the reduction of the background sound level and the development of the speech signal. Evaluate the volume of sound enhanced by reduced visual acuity and the type of sound reduction recommended.

II. Sound spectrum measurement:

The spectrum sound cannot be calculated in advance, but it will be about the time for some instance speech is not present within the sound speech. When the people are speaking, one should definitely pause to take a deep breath. We can include these gaps within speech signal to balance the background.

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

We can calculate the average size which is calculated for any frame in the last few seconds.

III. Removing Noise Spectrum:

After removing, all data of spectral signal which is appear unlikely in positive values. Some chances are available for removing unwanted parts. Fourier conversions, using section segments directly from the Fourier conversion unit, and additional additions are made to reconstruct the speech rate within the time zone.

The basic idea will be to reduce the noise from the given signal i.e-input signal:

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

Suggest workflow and simulation results

1. FLOWCHAR

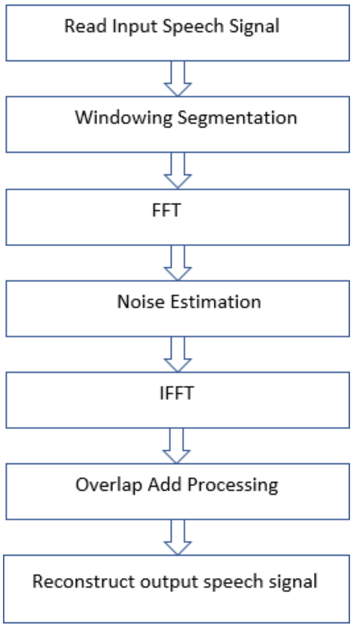


Fig. 3 Flow Chart

The Overlap add-on method is used to interrupt and keep the signals into smaller segments to make the process easier. The scattering method depends on: (1) signal decode in basic parts, (2) the procedure of every part, (3) reassembling the cleared parts in an additional speech. In order to execute domain of frequency process to reduce visibility, it's important to divide the constant wisdom signal speech into compact fragments which is called by frames. After completing procedure, the frames are then reassembled.

The flowchart above shows the proposed method, which involves collecting audio speech information and passing through a window to extract existing art objects within the Clamorous signal and use the FFT algorithm that detects the phase and magnitude of the audio signal, during which the Technology focuses on the size of the clamor to produce a different refined speech signal. By using the type of visual reduction the noise is measured and reduced by the value of the required magnitude. Then use the IFFT algorithm and the overlap adds the process to wish for a refined expression in the time zone.

I. Simulation Results

The following tests are used using screaming: Computer, train sound, and car audio using this AWGN sound produced. Every frame is 50% spaced FFT is applied. In particular, the quality of speech that is consistent with the test of comprehension test is held to judge the quality of sound speech that is improved by standard visual removal and the proposed noise removal method. Audio-type audio based on visual deletion and speech modification created by activated waveforms and spectrogram. Different sound level with sound speech signal is generated using 3 types of shouting.

The result of the audio file built into MATLAB

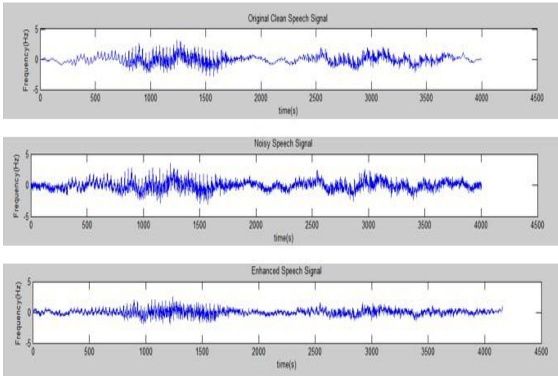


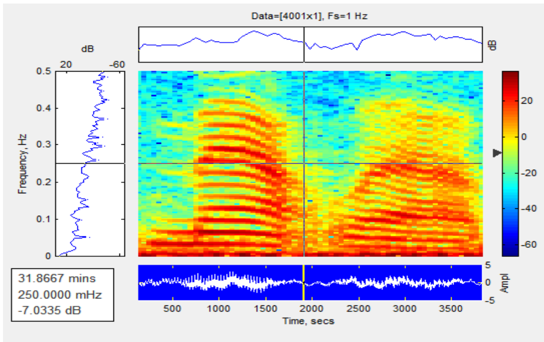
Fig 3: Results of speech enhancement:

Noise free speech

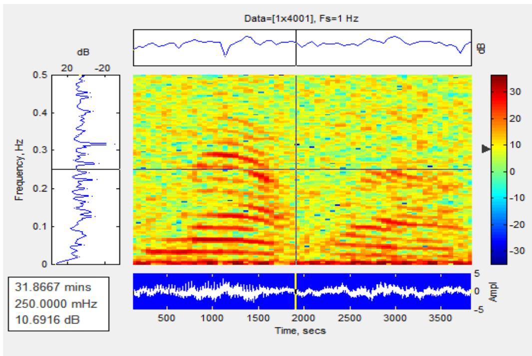
Noisyspeech

Enhancedspeechusing SSmethod

3.1 Results of Spectrogram:



Fig(4) Clean speech signal spectogram



Fig(5) Spectrogram of noisy speech signal

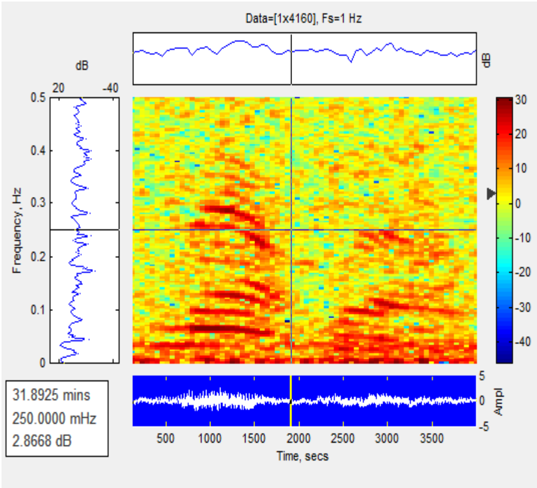


Fig (6) Enhanced speech signal

The above fig.6 shows the different colors of the spectrogram, while the red and yellow colors show two dissimilar energy level. All of the sound given input signals are full color of a very strong force due to the screaming pattern associated with being large yellow, red to medium power, and black to a small number with almost zero power. Starting mainly from the effects at the exact time of the speech signal, the effect of the shouting cycle. By combining Fih.5 with Fig.6, the effect as to clamor is reduced.

II. RESULTS OF DATABASE OF CAR NOISE AT 10DB

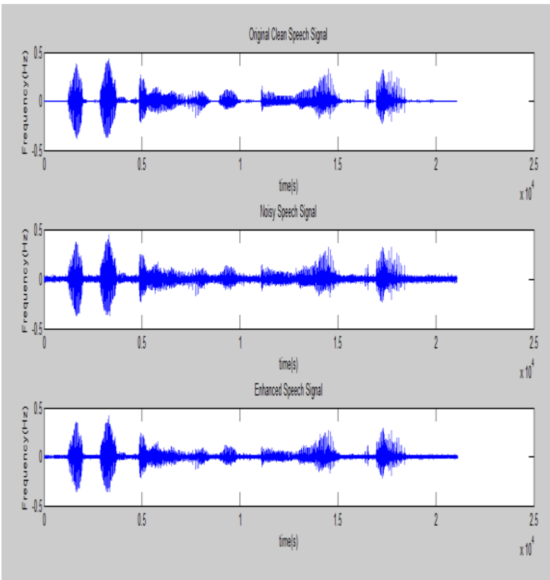


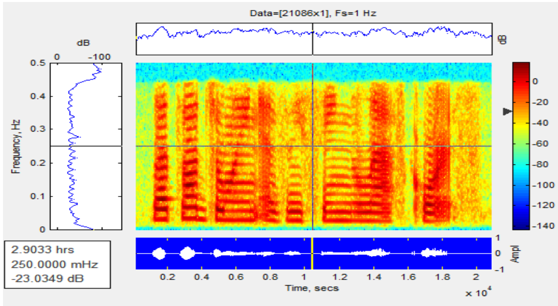
Fig 7: Results of speech enhancement:

Clean speech

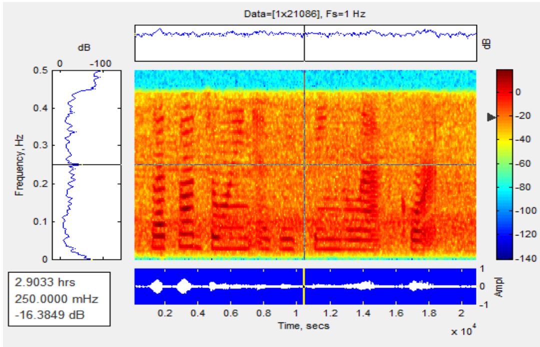
Noisyspeech

Enhancedspeech bySSmethod

III. Results of Spectrogram :



Fig(8) cleanspeechsignal



Fig(9) noisy speech signal

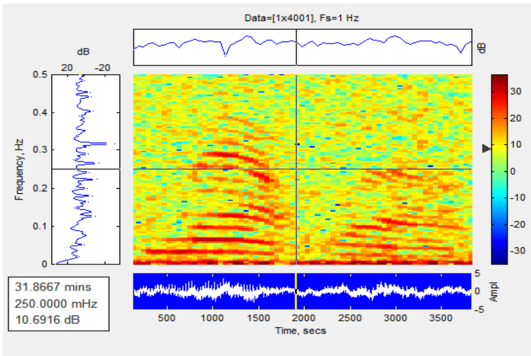


Fig (10) Enhanced speech signal

When matching Figure 9 with Figure 10, the noise effect decreased. Another interesting test of the proposed algorithm is seen in contrast to the spectrogram results found in Figure 9 and here an additional number of yellow lines left out the female speech signal suffering from high frequency. This proves that a woman's voice can have a very high frequency when separated from a male tone.

IV. RESULTS THAT WE GOT FROM TRAIN NOISE AND COMPARED WITH BACKGROUND NOISE IN DB USED FOR FEMALE AND MALE SPEECH SIGNALS

Signals of female speech	background Level (dB)	Noise Level in (dB)	Reduced level of noise
Train Noise (at 0dB)	3.6448	-4.852	8.4968
Train Noise (at 5dB)	-15.1722	-17.0717	1.8995
Train Noise (at 10 dB)	-25.5382	-41.2185	15.6803

Table -1: Reduced Noise level for male

Male speech signal	Noise Level in (dB)	Noise Level in (dB)	Noise Reduced level
Train Noise (at 0dB)	-3.9138	-6.2895	2.3757
Train Noise (at 5dB)	-22.4948	-19.3090	3.1858
Train Noise (at 10 dB)	-26.1571	-31.5079	5.3508

Table -2: Decreased Noise level for male

From Reduced noise level for male, It will be found the amplitude stages of male and female sound getting input and output different. This indicates that it has a significant effect on sound speech. It promotes the clamor reduction phase. This type of spectral reduction creates a signal of speech but also of shouting. Noise reduction was not properly guaranteed and can be guaranteed if separated by input and output.

The Clamor Stage of the input and output speech signal and amplitude reduction updates are different from using the table. Table 1 shows the background stage division for the male and female speech signal site.

V. CONCLUSIONS

In this presentation paper it is helpful to develop a better spectral reduction algorithm. It has recently been observed in the fictional results that the suggested type significantly reduces low noise compared to the algorithm of many ways to reduce visibility. This type of speech causes the signal to be visible but is accompanied by a negative screaming. The sound is not very low and it is very low compared to input and output. These forces will also be adjusted and expanded to accommodate static noise. From this type of domain design we have reached 70% and that we can build this system for embedded organizations that are subject to speech processing or communication purpose. For better comparisons, we showed as results and therefore spectrograms as to clear, sound and prepared speech. Delete the site-prepared speech sentence.

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 Improvement of Speech Signals Embedded in Non-stationary Noise, IEEE Transactions on Speech and Audio Processing, Vol. 6, No. 5, September 1998. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056

[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.

[13] Prince Priya Malla, Amrit Mukherjee, Sidheswar Routary, G Palai "Design and analysis of direction of arrival using hybrid expectation-maximization and MUSIC for wireless communication" IJLEO International Journal for Light and Electron Optics, Vol. 170,October 2018.

[14] Pavan Paikrao, Amit Mukherjee, Deepak kumar Jain "Smart emotion recognition framework: A secured IOVT perspective", IEEE Consumer Electronics Society, March 2021.

[15] Amrit Mukhrerjee, Pratik Goswani, Lixia Yang "DAI based wireless sensor network for multimedia applications" Multimedia Tools & Applications, May 2021, Vol.80 Issue 11, p16619-16633.15p.

[16] Sweta Alpana, Sagarika Choudhary, Amrit Mukhrjee, Amlan Datta "Analysis of different error correcting methods for high data rates cognitive radio applications based on energy detection technique" , Journal of Advanced Research in Dynamic and Control Systems, Volume:9 | Issue:4.