Effect of Speech enhancement using spectral subtraction on various noisy environment

Mr. Avinash C. Sathe

M.Tech. Student Dept. Of Electronics and Telecommunication Engineering,
TPCT's college of engineering,
Osmanabad, Maharashtra, India

---***---

Abstract - *Analysis Modification Synthesis (AMS) plays a key role in many audio signal processing applications, separating the audio stream into time intervals with speech activity and time intervals without speech. Many features have been introduced into the literature that reflect the existence of language. Therefore, this article presents a structured overview of several established speech enhancement features targeting different characteristics of speech. Categorize features in terms of their exploitable properties. B. Evaluate performance in a background noise environment, different input SNR categories, and some dedicated functions. Our analysis shows how to select promising VAD features and find reasonable tradeoffs between performance and complexity. To estimate clean speech using the Fast Fourier Transform (FFT), we emphasize the noise spectrum estimated during speech, subtract it from the noisy speech spectrum, and consider the average amplitude of the clean spectrum. and tried to develop a new method to minimize the spectrum of loud sounds. The noise reduction algorithm uses MATLAB software to semi-duplicate the noisy speech data (overlap-add processing) and use FFT to calculate the corresponding amplitude spectrum to remove noise from the noisy speech. and performed by reversing the audio in time. Reconstructed with the Inverse Fast Fourier Transform (IFFT).*

Key Words : *Analysis Modification Synthesis (AMS), Inverse Fast Fourier Transform (IFFT), Fast Fourier Transform (FFT) etc.*

1.INTRODUCTION

An important goal of speech communication systems is to transmit speech signals in a way that is correctly understood by the recipient. Examples can be found in the field of telephony and public address systems. Unfortunately, speech intelligibility can be affected by background noise. For example, poor speech intelligibility can be disruptive during a telephone conversation, but potentially dangerous in the context of a toll collection system voice alarm. as shown. At 1, background noise from either side of the communication channel can impair speech intelligibility for near-end listeners. In other words, noise can come from both the far end and the near end. To eliminate the negative effects of far-end noise, we usually apply single-channel noise reduction algorithms (see [1] for an overview). However, speech can also be preprocessed before playback to make it more audible in the presence of near-end background noise, which is the focus of this work. There are many ways to break free from ringing signals. Telephones are increasingly used in noisy environments such as cars, airports, streets, trains and stations. Now let's pull back the noise using the spectral minimization method. The purpose of this work is to build a real-time system that can reduce the background caused by audio signals. This process is called speech enhancement. In many language systems, the background lowers the language standard. It is usually difficult to remove noise without distorting the audio. Therefore, the performance of speech enhancement devices is limited between speech distortion and noise reduction. from the loud sound spectrum. In this method, the following he assumes three situations. Channels are on the market in which the sound adds to the audio signal and the sound is uncorrelated. During this work, we used spectral subtraction techniques to try to reduce noise spectrum estimation errors and improve corrupted speech. We proposed a method that was tested on a real language data framework in the MATLAB environment. Real speech signals from the speech data database were used for various experiments. We then propose a technique to reduce the difference between the estimated noise spectrum and the noise spectrum. In general, we find that systems with only one medium perform under the most difficult conditions in response to varying audio data and unwanted noise, where advanced noise intelligence is not available. Usually, when the speech is polite, men consider the noise to be steady. They usually tolerate non-stationary noise during periods of speech activity, but in practice the power of speech signals decreases dramatically when the noise is non-stationary.

2. Methods of noise reduction

Before you start formatting your work, first write and save the content as a separate text file. Please complete all content and organization edits before formatting. See Sections A-D below for additional proofreading, spelling, and grammar information.

Keep text and graphic files separate until you have finished formatting and styling the text. Do not use hard tabs and limit the use of hard line breaks to the single line break at the end of a paragraph. Do not add pagination anywhere on the paper. Do not number text headers. The template will number them for you.

A. Conventional spectral subtraction

Many additional modifications of various conclusions were invented to improve the language. The single we use is considered the conclusion of phantom subtraction. This guy works within the Epidemic Bureau and trusts Phantom of Aid to be mentioned when Sound Phantom and Noise Phantom are added. The action is shown in the image below and in two main parts It has been constructed.

B. Noise estimation

A minimum statistical noise estimate is used to estimate the background noise.

where * means the convolution process and is the bias value

Background spectrum estimation. Subtract noise spectrum from noisy speech

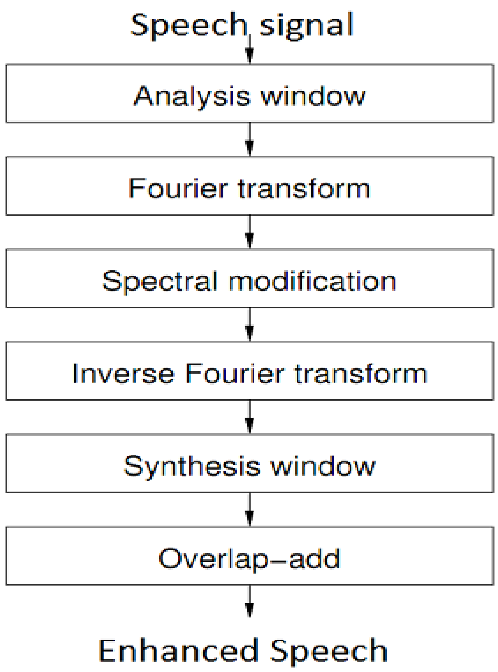

Fig.1 Conventional Spectral Subtraction

Figure 1 shows the conventional spectral subtraction using the overlap add method.

The AMS method [6], shown in Figure 1, is an efficient method of signal amplification. AMS uses the following procedure: We first frame the input audio signal using a suitable window function, then perform his STFT of the windowed frames using some frame shifts. The third is the inverse Fourier transform, and the fourth is signal acquisition by the overlap-and-addition (OLA) method [7].

Consider an additive noise scenario like Equation (1). One equation in a row.

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

where x(n), s(n), and N(n) are windowed samples of noisy speech, clean speech, and spurious noise, respectively. n Discrete time index. Since speech signals are inherently non-stationary, speech in AMS frames is processed over short frame durations using short-term Fourier transforms. From the STFT definition, the spectrum of noise degrades speech.

Not all spectral magnitude values appear to be positive after subtraction. There are several ways to remove unwanted components. An inverse Fourier transform using the phase components directly from the Fourier transform unit and an overlap-add are then performed to reconstruct the speech estimate in the time domain. The basic concept is to minimize the noise of the input noise signal.

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

$$x(j) = (1/N)\sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)} \quad (2),(3)$$

Unfortunately, we don't know the exact level of the noise signal, so we reduce the amplitude and leave the level of X unchanged. After subtraction, the spectral magnitude is definitely not positive. There are several ways to remove the negative component present in the spectrum. To convert the frequency-domain signal to the time-domain signal, the phase of the noisy speech signal is combined with the processed amplitude spectrum and an inverse discrete Fourier transform (IDFT) is initiated. time domain.

	Input SNR (dB)			
Noisy s	0dB	5dB	10dB	15dB
Ses01M_script01	LLR=0.624	LLR=0.543	LLR=0.497	LLR=0.4542
	SNRseg= -1.028139	SNRseg=1.450995	SNRseg=2.853354	SNRseg=3.937311
	WSS=92.23 PESQ=1.98	WSS=81.82PESQ=2.18	WSS=77.46 PESQ=2.29	WSS=74.119 PESQ=2.40

Table 1. shows composite objective measure (COM) at different input SNR (dB) using di erent noise estimation techniques for NOIZEUS utterances

Figure. 2. shows the generalised Short Time Fourier Transform incorporating spectral subtraction

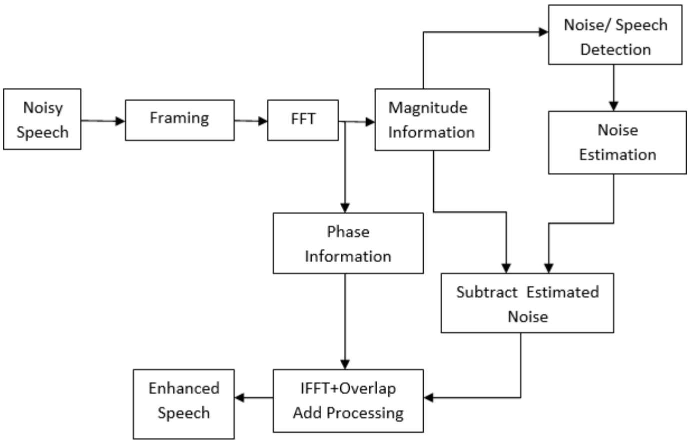

Fig. 2. Conventional Spectral Subtraction

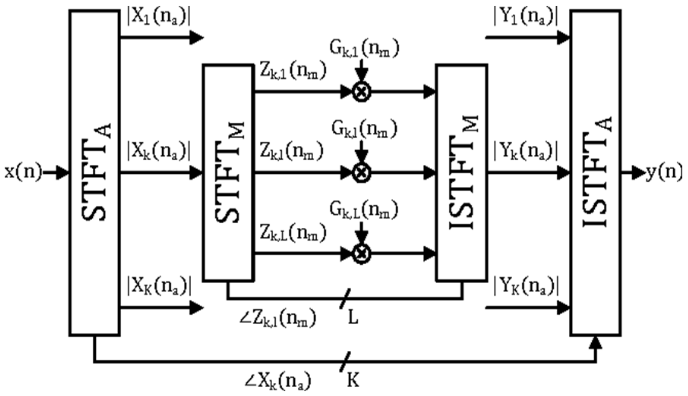

Fig. 3. Short Time Fourier Transform of speech Enhancement incorporating Overlap-Add Processing

3. Experimental Evaluation

There are many public or commercially available speech and noise databases for evaluating speech enhancement algorithms. In this work, we used the language corpus database NOIZEUS [7] for our experiments. The basic premise of databases like NOIZEUS is to provide researchers with more realistic noise recordings at various input SNRs. The speech corpus consists of 30 phonetically balanced her IEEE sentences from 6 speakers (3 male, 3 female). The experiments used corpus noise stimuli constructed by real-time noise environments at different input SNRs, such as airport, chatter, car, restaurant, station, and train background noise. Thirty sets of mean values for the NOIZEUS corpus were calculated for each treatment type, noise type, and input SNR. C. Complex objective measures Speech quality is also assessed by the Composite Objective Scale (COM). Several objective measures of composite quality are derived from multiple regression analysis. We compared the spectral behavior of representative features from time-frequency spectrogram analysis. Capturing an enhanced audio signal compared to a noisy spectrum at various input SNRs such as 0dB, 5dB, 10dB, 15dB without exploiting the temporal context of the audio information was observed.

4. Spectrogram analysis

Spectrograms, in particular, have different colors. Here red and yellow indicate two different intensity phases. The entire loud audio input signal is full of extreme intensity colors because yellow is loud, red is high in the center intensity, and black is a weak crowd with little intensity. Especially during quiet times of the resulting audio signal, the noise result can be more.

	Input SNR (dB)			
Noisy s	0dB	5dB	10dB	15dB
Ses01M_script01	Csig=2.816	Csig=3.113	Csig=3.270	Csig=3.4063
	Cbak=1.87	Cbak=2.19	Cbak=2.37	Cbak=2.510
	Covl=2.224	Covl=2.499	Covl=2.647	Covl=2.7754
	loss=0.8467	loss=0.8073	loss=0.7759	loss=0.75763

Table 2 shows PESQ scores for proposed speech enhancement technique at different input SNR (dB) using di erent noise estimation techniques for NOIZEUS utterances

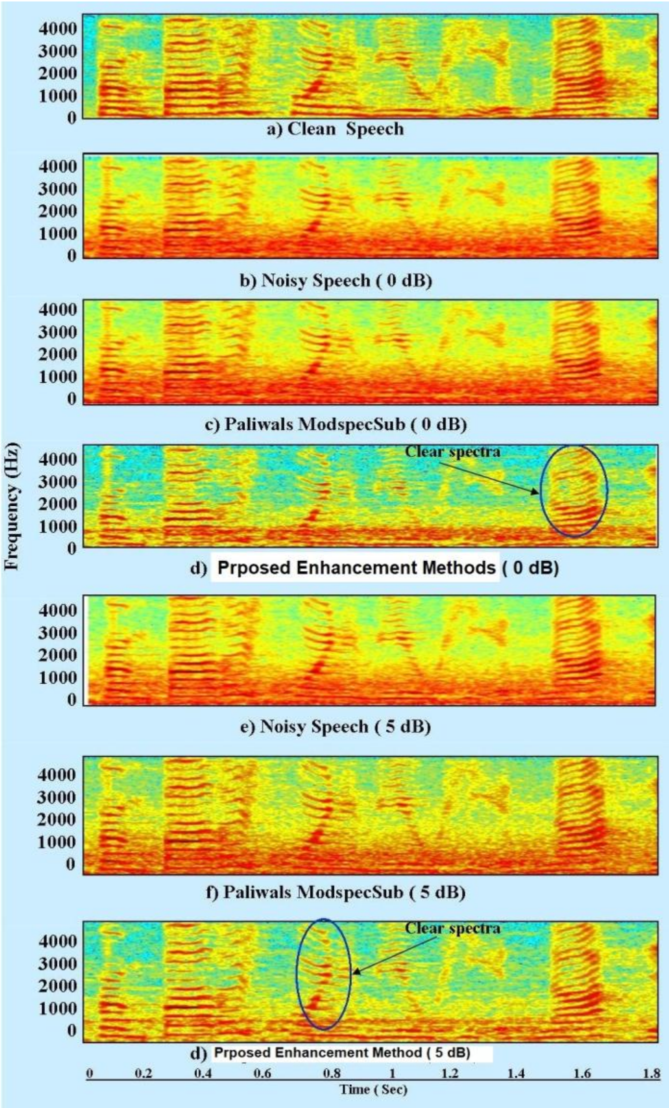

Figure 4. Spectrogram of pure audio signal

We compared the spectral behavior of representative features from time-frequency spectrogram analysis. Capturing an enhanced audio signal compared to a noisy spectrum at various input SNRs such as 0dB, 5dB, 10dB, 15dB without exploiting the temporal context of the audio information was observed. To illustrate the proposed performance, consider the spectrogram plot in Figure 5. Spectrograms are shown for clean speech (5.5a) and noisy speech (5.5b, 5.5e) with input SNRs of 0 dB and 5 dB, respectively. Figure 5.5b shows that the low frequencies, where most of the audio energy resides, are corrected more than the high frequencies. For speech stimuli constructed with the proposed method (5.5d), a significant reduction in noise is observed at certain low frequencies compared to the Mod-SpecSub method (5.5c).

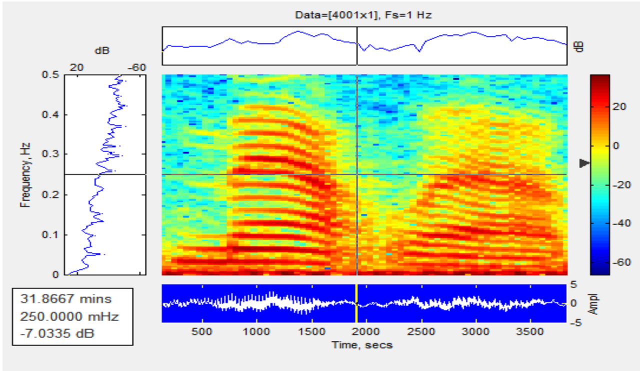

Fig. 5. Spectrogram analysis of speech signal at 0db input SNR noise: a) clean b) noisy signal 0dB c) Paliwals Modspecsub 0dB d) proposed enhancement method 0dB e) noisy signal 5dB f) Paliwals Modspecsub 5dB g) proposed enhancement method 5dB

V. CONCLUSIONS

In this article, we summarized and analyzed several spectral subtraction functions using different noisy environments. We have outlined established approaches for the purpose of classifying features according to the language extension properties used. Our analysis showed that performance was taken into consideration. We identified objective evaluations such as the Perceived Speech Quality Score (PESQ), Csig, Cbak, overall signal quality Covl and loss of signal context as important aspects for improving speech enhancement performance. .

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 Volume: 04 Issue: 03 | Mar -2017 www.irjet.net p-ISSN: 2395-0072

© 2017, IRJET | Impact Factor value: 5.181 | ISO 9001:2008 Certified Journal | Page 2488

[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] P Loizou. Noizeus: A noisy speech corpus for evaluation of speech enhancement algorithms.Speech Commun, 49:588{601, 2017..

[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.