 ISSN : 2230-7109 (Online) | ISSN : 2230-9543 (Print)

Noise Estimation for Speech Enhancement

[1]Shubham Nayak, [2]Ashvini Paliwal, [3]Aditya Mandhane, [4]Jagdish Kene
[1,2,3,4]Dept. of ECE, Shri Ramdeobaba College of Engg. & Management, Nagpur, Maharashtra, India

Abstract
This technical paper begins by introducing the concept of Speech enhancement, including a brief process of removal of noise for enhancing the speech. It discusses the main theoretical concepts of speech enhancement including the types of noise, algorithm used for enhancing speech, role of enhancement and procedures followed to get cleaned speech. This project is provided with various voice samples and output is provided at very back of the paper.
The main purpose of speech enhancement technique is to eliminate the background noise from speech signal to improve the quality of the speech signal. Speech signal is often corrupted by additive background noise like train noise, market noise, buzzing sound by internal components of mobile, wind etc. In such noisy environment listening at the end user is very difficult. This paper presents speech enhancement method using a priori noise estimation algorithm. Based on Peak Signal to Noise Ratio and Mean Square Error criterion the system evaluates the amount of surrounding noise present in speech. The proposed algorithm helps to improve the quality of speech signal at the listener end.

Keywords
Speech Enhancement, Apriori Noise Estimation, SNR Estimation, Posterior SNR Estimation.

I. Introduction

Speech enhancement algorithms have developed considerably, and significant progress has been made in detecting noise from degraded speech signal [1-2]. Estimating noise from voice signal is very difficult when the signal to noise ratio is very low or the random noise is non-stationary. This results into speech distortions and unnatural sounding in enhanced speech signal. Thus the system has to be enhanced so as to get fine response of output speech signal even when the signal to noise ratio is very low.
In this paper, we focus on single channel subtractive-type speech enhancement. This type of enhancing technique is quite tedious as noise and speech signal are collectively present in the same channel.
The single channel subtractive-type speech enhancement has two major drawbacks: (1) the introduction of a musical residual noise with an unnatural structure in the enhanced speech. (2) a voice activity detector cannot correctly segment noisy speech signal at very low SNR.
The solution proposed in this paper uses improving power spectral density estimation based on minimum statistics to track the varying of SNR; and exploits the masking properties of the human auditory system to overcome the limitations of one channel subtractive-type enhancement systems in additive background noise at very low SNR [3-4]. This allows one to find the best tradeoff between the amount of noise reduction, the speech distortion and the level of residual noise [3].

II. The Enhancement Algorithm

Consider a band limited noisy speech signal y(i). which is the sum of a clean speech signal s(i) and an additive stationary background noise signal d(i). The noisy speech can be expressed by:

y(i) = s(i) + d(i)

where i denotes the sampling time index. We further assume that s(i) and d(i) are statistically independent and zero mean [3].

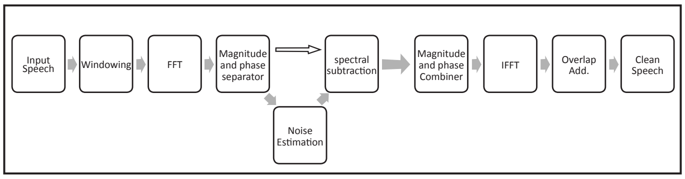

Fig. 1:

The processing is done on a frame-by-frame basis in the frequency domain.

Below represents the block diagram for speech enhancement:

In the frequency domain, this may be denoted as:
$Y(j\omega) = S(j\omega) + D(j\omega)$ =>
$S(j\omega) = Y(j\omega) - D(j\omega)$
where $Y(j\omega)$, $S(j\omega)$, $D(j\omega)$ are Fourier transforms of y(t), s(t), d(t), respectively.

The statistic parameters of the noise are not known, thus the noise and the speech signal are replaced by their estimates:
Since the speech is assumed to be uncorrelated with the background noise, the short-term power spectrum of $y[i]$ has no cross-terms. Hence,

$|Y(\omega)|^2 = |S(\omega)|2 + |D(\omega)|^2$

The speech can be estimated by subtracting a noise estimate from the received signal.

$$|\widehat{S}(\omega)|^2 = |Y|(\omega)|^2 - |\widehat{D}(\omega)|^2$$

The estimation of the noise spectrum $|\hat{D}(\omega)|^2$ is obtained by averaging recent speech pauses frames:

$$|\widehat{D}(\omega)|^2 = \frac{1}{M}\sum_{j=0}^{M-1} |Y_{SP_j}(\omega)|^2$$

where M is the number of consecutive frames of Speech Pauses (SP)
Further improvement is done to increase efficiency of noise estimation by using apriori noise estimation technique which is given by:

$$SNR_{(posterior)} = \frac{|D(w)|}{|D^{\wedge}(w)|} - 1$$

For apriori estimation,

$$SNR_{(estimate)} = (1-a)\{\frac{(1-a)}{SNR(posterior)}\}$$

Where a is incremental factor. a=0.05
Final estimation is given by:

$$|\hat{D}(\omega)| = \text{Max}\ \{(1-l)(\frac{1}{SNR(apriori)+1})^{b1}\}^{b2}$$

Where b1 and b2 are gain factor.
l is frequency band. *l* = 3
b1, b2 are 0.5 and 1 respectively.

Clean speech is obtained by subtracting final estimated noise from input speech.

III. Results & Discussion

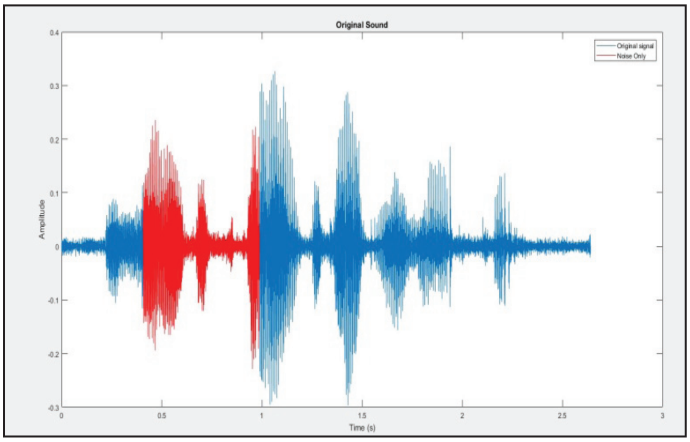

Fig. 2:

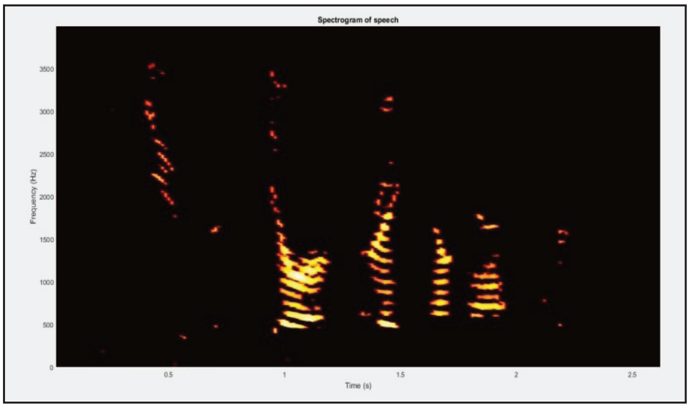

Fig. 3:

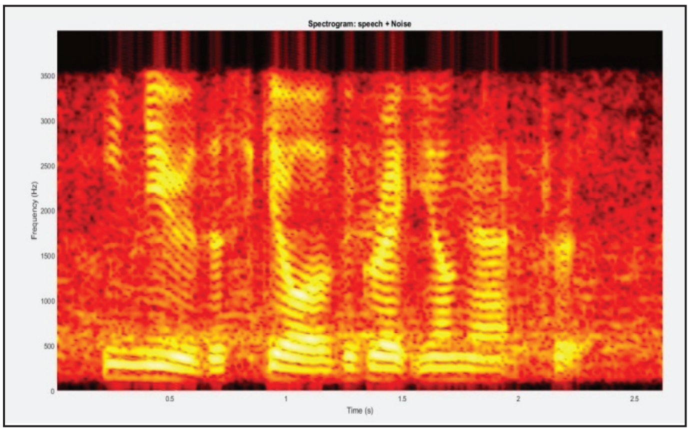

Fig. 4:

IV. Conclusion

Apriori noise estimation is the efficient algorithm to remove the non stationary noise from background in speech signal but it introduces annoying residual noise. In this paper, a simple implementation of spectral estimation property of non-stationary noise and masking of auditory system is proposed.

References

[1] S. F. Boll,"Suppression ofacoustic noise in speech using spectral subtraction", IEEE Tms. Acourt. Speech, Signal Processing, Vol. ASSP-27, pp. 113-120, 1979.
[2] M. Bemuti, R. Schrvanz, J. Makhoul,"Enhancement of speech corrupted by acoustic noise," In Pmc.IEEEICASSP, Washington, DC. Apr. 1979. pp. 208-2 I I
[3] Nathalie Virag,"Single channel speech enhancement baed on masking properties of the human auditory system", IEEE Trans. Speech and Audio Processing, Vol. 7, pp. 126-137, March. 1999.
[4] R.Martin,"Noise power spectral density estimation based on optimal smoothing and minimum statistics", IEEE Trans. Speech and Audio Processing, Vol. 9, pp. 504-512, 200l.

Dr. Jagdish Deorao Kene completed his bachelor degree in Electronics Engineering in 2001, from Manoharbhai Patel Institute of Engineering and Technology, Nagpur University, Nagpur and Master degree in Electronics Engineering in 2005, from Yashavantrao Chowan College of Engineering, Nagpur University, Nagpur, M. S. India. He completed his Ph.D. degree in Electronics Engineering, in the year 2015 from VNIT, Nagpur. He is currently associated with Ramdeobaba College of Engineering & management, Nagpur, as Assistant Professor in Electronics Communication Engineering Department having total experience of 15 years. His research work is related to Performance evaluation and optimization solution of physical layer by implementing various error correction coding techniques in mobile Wi-Max environment. He has published 9 Journal Papers, 5 papers in International Conferences in his research area. He also published 5 papers in International Conference and more than 12 have been published in National Conferences in his academic carrier. He is member of Professional societies like ISTE. He believes that Trust and Honesty is the secrets of success.

Mr. Shubham Kirti Nayak is pursuing Electronics and Communication Engineering From Ramdeobaba College of Engineering and Management, Nagpur. He is currently a student of BE final year. He is currently working on developing noise estimation algorithm for speech enhancement for mobile communication.

 ISSN : 2230-7109 (Online) | ISSN : 2230-9543 (Print)

Mr. Aditya Mandhane is pursuing electronics and communication engineering from Ramdeobaba College of Engineering and Management, Nagpur. He is currently a student of BE final year. He is currently working on developing noise estimation algorithm for speech enhancement for mobile communication.

Mrs. Ashvini Paliwal is pursuing Electronics and Communication Engineering from Ramdeobaba College of Engineering and Management, Nagpur. She is currently a student of BE final year. She is currently working on developing noise estimation algorithm for speech enhancement for mobile communication.