Prof. Vijaya Sugandhi / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 3, Issue 1, January-February 2013, pp.074-076

Spectral Analysis in Speech Processing Techniques

Prof. Vijaya Sugandhi
Deptt. of Electrical & Elex. Engg. Shri Vaishnav[SM] Institute of Tech. & science,Indore(M.P.)

ABSTRACT

The corruption of speech due to presence of additive background noise causes severe difficulties in various communication environments. This paper addresses the problem of reduction of additive background noise in speech. The proposed approach is a frequency dependent speech enhancement method based on the proven spectral subtraction method. Most implementations and variations of the basic spectral subtraction technique advocate subtraction of the noise spectrum estimate over the entire speech spectrum. However, real world noise is mostly colored and does not affect the speech signal uniformly over the entire spectrum. This method provides a greater degree of flexibility and control on the noise subtraction levels that reduces artifacts in the enhanced speech, resulting in improved speech quality.

Key words: spectrum, noise subtraction.

I INTRODUCTION

Speech processing has been a growing and dynamic field for more than two decades and there is every indication that this growth will continue and even accelerate. During this growth there has been a close relationship between the development of new algorithms and theoretical results, new filtering techniques are also of consideration to the success of speech processing. One of the common adaptive filtering techniques that are applied to speech is the Wiener filter.

A spectrogram is a time-varying spectral representation (forming an image) that shows how the spectral density of a signal varies with time. Also known as spectral waterfalls, sonograms, voiceprints, or voicegrams, spectrograms are used to identify phonetic sounds, to analyses the cries of animals; they were also used in many other fields including music, sonar/radar, speech processing, seismology, etc.The instrument that generates a spectrogram is called a spectrograph.

The frequency spectrum of a time-domain signal is a representation of that signal in the frequency domain. The frequency spectrum can be generated via a Fourier transform of the signal, and the resulting values are usually presented as amplitude and phase, both plotted versus frequency. Any signal that can be represented as amplitude that varies with time has a corresponding frequency spectrum. This includes familiar concepts such as visible light (color), musical notes, radio/TV channels, and even the regular rotation of the earth. When these physical phenomena are represented in the form of a frequency spectrum, certain physical descriptions of their internal processes become much simpler. Often, the frequency spectrum clearly shows harmonics, visible as distinct spikes or lines that provide insight into the mechanisms that generate the entire signal.

II BACKGROUND

In the past decades, research in the field of speech enhancement has focused on the suppression of additive background noise [3] [4] [5]. From the point of view of signal processing, additive noise is easier to deal with than convolutive noise or nonlinear disturbances. The ultimate goal of speech enhancement is to eliminate the additive noise present in speech signal and restore the speech signal to its original form. Several methods have been developed as a result of these research efforts. Most of these methods have been developed with some or the other auditory, perceptual or statistical constraints placed on the speech and noise signals. However, in real world situations, it is very difficult to reliably predict the characteristics of the interfering noise signal or the exact characteristics of the speech waveform. Hence, in effect, the speech enhancement methods are sub-optimal and can only reduce the amount of noise in the signal to some extent. Due to the sub-optimal nature of these methods, some of the speech signal can be distorted during the process. Hence, there is a trade-off between distortions in the processed speech and the amount of noise suppressed.

III. PROPOSED METHODOLOGY

The following simplified model for voiced speech production, where the speech signal s(n) is formed as the convolution [1]

$$s(n) = e(n) * \theta(n), \qquad (1.1)$$

where e(n) is the excitation source and θ (n) the vocal tract response. For speech recognition, extracting the vocal tract response and discarding the excitation information from the resulting signal is useful, as the information relevant for distinguishing the spoken words is mainly in the vocal tract response, while the excitation source primarily contains the irrelevant pitch information.

If we assume that y(n) , the discrete noisy input signal, is composed of the clean speech signal s(n) and the uncorrelated additive noise signal d(n) , then we can represent it as:

$y(n)=s(n)+d(n)$ (1.2)

Processing is done on a frames-by-frame basis. Analysis of overlapping frames of the noisy signal is implemented by using the Discrete Fourier Transform (DFT) preceded by a Hamming window. The power spectrum of the noisy signal can be written as:

$|Y(k)|^2 \approx |S(k)|^2+|D(k)|^2$ (1.3)

Since the noise spectrum D(k) cannot be directly obtained, a time-average of the power spectrum Dˆ(k) is calculated during a period of silence.

Assuming that noise is uncorrelated with the speech signal, an estimate of the modified speech spectrum can be given as:

$|\hat{S}(k)|^2= |Y(k)|^2 - |\hat{D}(k)|^2$ (1.5)

From Eq. (1.5) it can be seen that the subtraction process involves the subtraction of an average estimate of the noise from the instantaneous speech spectrum. Due to the error in computing the noise spectrum, we may have some negative values in the modified spectrum. These values are set to zero. This process is called half-wave rectification.

Distortions due to half / full wave rectification

The modified speech spectrum obtained from Eq. 1.5 may contain some negative values due to the errors in the estimated noise spectrum. These values are rectified using half wave rectification (set to zero) or full-wave rectification (set to its absolute value). This can lead to further distortions in the resulting time signal.

Modifications to spectral subtraction

Several variants of the spectral subtraction method originally proposed by Boll [2] have been developed to address the problems of the basic technique, especially the presence of musical noise. Still other methods based on this method have been developed that perform noise suppression in the autocorrelation, cepstral , logarithmic and sub-space domains. A variety of pre and post processing methods have also proved to help reduce the presence of musical noise while minimizing speech distortion.

IV. RESULTS

This result computes an Audio Spectra in MAT LAB. The fft function computes the FFT of a specified signal. In general, either the magnitude or phase values of the FFT coefficients are find, which are in Matlab can be determined using the abs and angle functions. A variety of windows can be applied to a signal before the computation of the FFT using the functions hann, hamming, blackman. Time-domain windows can help minimize spectral artifacts related to signal truncation. The specgram function computes a time-frequency plot of a signal where color represents spectral magnitude amplitude.

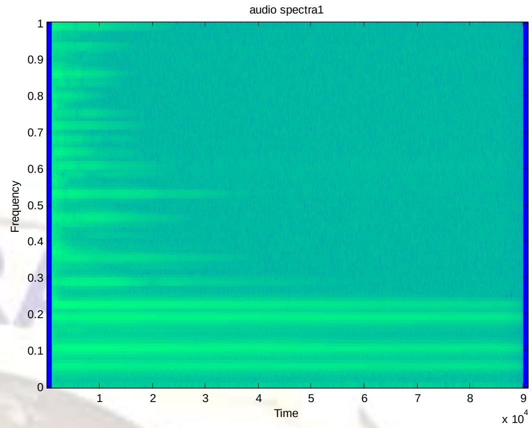

Fig 1 Audio Spectra

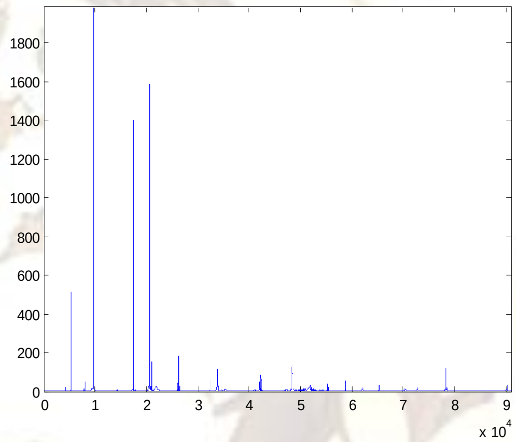

Fig 2 Spectral Analysis

V. CONCLUSION

The work in this paper addressed the problem of enhancing speech in noisy conditions. A multi-band spectral subtraction method, based on the direct estimation of the short-term spectral amplitude of speech and the non-uniform effect of noise on speech, was proposed. The results establish the superiority of the proposed method over the conventional spectral subtraction method with respect to speech quality of the enhanced signal and reduced residual noise. The major contributions of this paper are development of a multi-band speech enhancement strategy based on the spectral subtraction method. Speech processed by the new algorithm shows reduced levels of residual noise and good speech quality.

VI Acknowledgment

This author thanks to Principal and Administrative Officer of SVITS Indore,Prof.Hamant Chouhan, Prof. Dev Kumar Rai thanks to author’s family for their motivation support and encouragement.

Reference

[1] Deller, J. R., Hansen, J. H., and Proaxis, J. G. Discrete-Time Processing of Speech Signals. IEEE Press, 2000.
[2] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Process., vol.27, pp. 113-120, Apr. 1979.
[3] J. Deller Jr., J. Hansen and J. Proakis, "Discrete-Time Processing of Speech Signals", NY: IEEE Press, 2000.
[4] H. Levitt, "Noise reduction in hearing aids: An overview", Journal of Rehabilitation Research and Development, vol. 38, No. 1, January/February 2001.
[5] J. Lim and A. Oppenheim, "All-pole modeling of degraded speech," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 26, No. 3, pp. 197-210, June 1978.

Ms.Vijaya sugandhi received B.E.and M.E.(digital techniques &instrumentation) Degree with in Electrical engineering in 2006 and 2011 from Rajiv Gandhi Technical University, Bhopal, Madhypradesh, India