JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Speech Noise Reduction in Hearing Aids to Improve Perceptual Quality of Speech Using Discrete Fourier Transform

[1]Madhumati Shridhar Kuntoji, [2]Dr. Kiran V ,
[1] Student, M.Tech (Communication Systems), [2]Associate Professor,
[1]Department of Electronics and Communication Engineering
[1]R.V.College of Engineering, Bangalore, Karnataka, India

Abstract : In humans, Sensory Neural Hearing Loss (SNHL) leads to decreased frequency selectivity [1] leading to poor speech perception. This is a most common hearing problem in humans across the world. It is not much known about effect of SNHL on frequency selectivity which is function of speech signal level [1]. Auditory filter frequency response is approximately linear in hearing impaired people but it is nonlinear in normal hearing people. Therefore frequency discrimination ability is very poor in hearing impaired people. This type of hearing problem becomes severe in noisy situations. Speech noise reduction is very much essential for hearing impaired people using hearing aids. Here a low computational intense, noise reduction technique, suitable for hearing aids is presented. Most of the noise reduction algorithms process magnitude of Fourier Transform (FT) of speech signal but phase is kept as it is [3]. In contrast here magnitude of FT of noisy speech signal is kept unprocessed and phase alone is processed. Processed spectrum of speech is obtained by combining unprocessed magnitude spectrum and processed phase spectrum [3]. Phase spectrum processing (modification) is based on the type of noise and signal to noise ratio of speech corrupted by noise. This method of processing results in reduction of low energy components (generally noise) more as compared to high-energy (generally speech) components. Reduction of noise is very much required in hearing aids along with other signal processing operations. Laboratory evaluation of noise corrupted speech and experimentally modified speech samples reveal that present technique is able to reduce noise and improve speech quality. Noise reduction is attained by phase alteration of Fourier transform of input noisy speech signal using imaginary parts of FT.

***Index Terms* - Noise reduction, Hearing problems, Hearing aids, STFT, Magnitude spectrum, Phase spectrum, Symmetry of FT.**

I. INTRODUCTION

Speech intelligibility and clarity, which are very important for people suffering from hearing loss and hearing impairment, can be improved by its enhancement [10, 11, 12, 13]. Hearing aids augmented with speech enhancement algorithms at the front end improve the speech quality leading to improved perceptual speech quality.

Let us consider a noisy speech signal x(n) =s(n) +N (n) (1.1)

Assuming the noise to be additive. Where x(n)=Noisy Speech , s(n)=Clean Speech, N(n)=Noise. STFT is used for the frame wise processing. Taking Fourier transform of equation (1.1) we get

$$X_1(k) = \sum_{m=-\infty}^{\infty} x(m)w(n-m)e^{-i2\pi km/N} \tag{1.2}$$

Here n, k, w(n) are frame duration, discrete frequency index and window function respectively. Hamming window of 20ms duration is used. Taking Fourier transform of equation (1.1) gives X(k) = S(k)+N(k) where X(k), S(k) and N(k) are STFT of noisy speech, clean speech and noise respectively [1]. From earlier works it is observed that, noise reduction techniques modify the modulus of spectrum of noisy speech, whereas phase spectrum is unaltered. The processed magnitude spectrum is combined with unprocessed phase spectrum during signal synthesis. In the present attempt, the absolute value of discrete Fourier transform of noise corrupted speech is not processed but spectrum angle is logically altered [3]. The unaltered spectrum absolute value is clubbed with altered spectrum angle to result processed spectrum. This leads to suppression of small energy (noise) spectral elements higher as compared to more energy (speech) spectral elements [3,14,15]. during signal reconstruction. This leads to enhancement of speech and is comparable with existing algorithms. Speech quality is tested by objective speech quality measures, subjective listening tests and spectrograms. The paper is organized as follows. Section: II gives details of the proposed work.

Section: III deals with relation between α and β. Section: IV deals with experimental details. Section: V presents the experimental details, finally section: VI presents the conclusion and future scope of the work.

II. PROPOSED METHOD

The block diagram of the proposed method is shown in Fig.1.

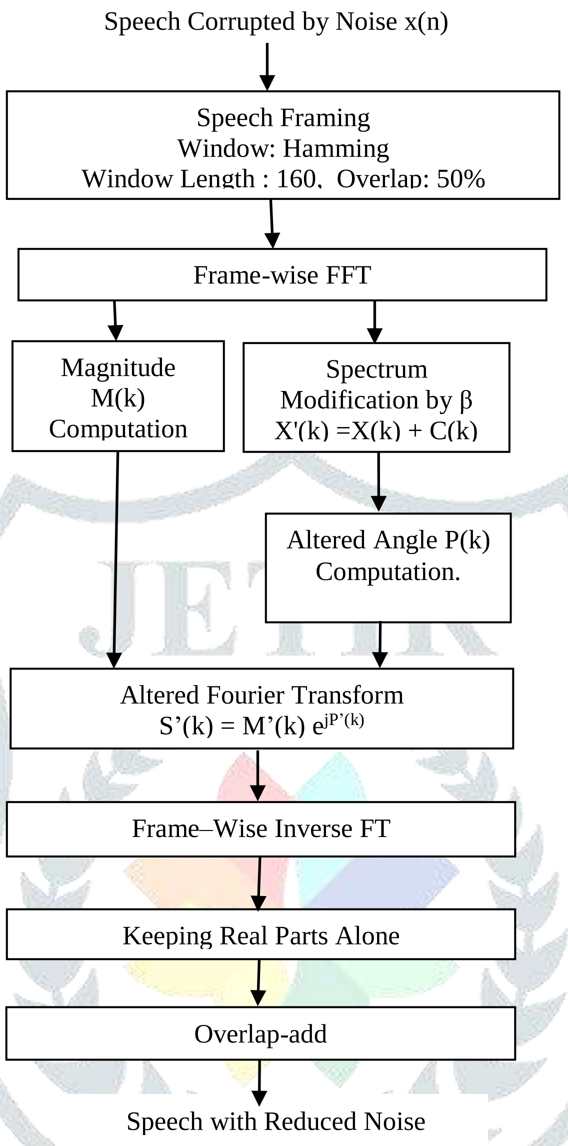


Figure1. Proposed Work Block Diagram

2.1 Details of the Proposed Work

The block diagram of the proposed work is shown in Fig.1. The input noisy speech signal x(n) is real hence it's discrete Fourier transform (DFT) obeys conjugate symmetry property i.e. $X(k) = X^*(N-k)$. Inverse DFT of X(k) results into original noisy speech signal x(n) due to cancellation of imaginary components of complex conjugate terms. The degree of cancellation or addition of complex conjugate terms can be controlled by altering their phase [3]. In the proposed work, phase is altered by modifying imaginary parts of the DFTs.

2.1.1 Phase Modification by Altering Imaginary Parts of DFTs.

Assuming the size of DFT 'N' to be even, an imaginary constant C(k) given by

$$C(k) = j\beta \; ; 0 \le k < N/2 \tag{2.1}$$

$$C(k) = -j\beta \; ; N/2 \le k \le N-1 \tag{2.2}$$

is a complex valued frequency dependent function, anti-symmetric about Fs/2 rad/sample frequency. Here it is assumed to be a constant β, independent of frequency. Using this constant the DFT of noisy speech signal is modified as follows. The noisy speech signal frame-wise DFT, X(k) is modified as

$$X'(k) = X(k) + C(k) \tag{2.3}$$

The altered phase of X'(k) is calculated and further combined with unaltered magnitude of noisy speech signal to get modified spectrum given by

$$Xm(k) = |X(k)| \; e^{j\angle X'(k)} \; ; 0 \le k \le (N-1) \tag{2.4}$$

The IDFT of above spectrum results into approximate enhanced speech signal. Since conjugate symmetry is altered due to altered phase, IDFT of altered spectrum results into complex time domain signal. Only real components are retained for further processing. The details are as follows. The magnitude spectrum of DFT of a real signal obeys even symmetry condition whereas phase spectrum obeys odd symmetry condition. The complex conjugates sum together to result into a real signal due to cancellation of their imaginary components during reconstruction of time domain signal. The degree of cancellation or addition of these complex conjugates can be controlled by altering their phase angles. Here phase alteration is carried out through imaginary parts of the complex conjugates. This technique results in reduction of noise associated with noisy speech signal and also computational complexity involved in the processing is also reduced to a greater extent. The above technique can be explained using signal – vector analogy. Considering a pair of complex conjugate numbers $C_1=X+jY$ and $C_1^*=X-jY$ both having same magnitude component given by

$$M=\sqrt{X^2+Y^2} \tag{2.5}$$

and phase angles given by

$$\varphi_1 = \tan^{-1}(Y/X) \tag{2.6}$$

and $$\varphi_1^* = \tan^{-1}(-Y/X) \tag{2.7}$$

The altered complex numbers are given by

$C_{11}=X+jY+j\beta$ and

$$C_{11}^*=X-jY-j\beta. \tag{2.8}$$

The altered phase angles are given by $\theta_{11}=\tan^{-1}(Y+\beta/X)$ and $\theta_{22}=\tan^{-1}(-Y-\beta/X)$ (2.9)

The resultant of above two complex conjugate numbers depends on the factor β and is given by the equation

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\left(\tan^{-1}\left(\frac{Y+\beta}{X}\right)-\tan^{-1}\left(\frac{Y+\beta}{X}\right)\right)} \tag{2.10}$$

Case I. The Resultant R, When the factor $\beta << \sqrt{X^2+Y^2}$

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\left(\tan^{-1}\left(\frac{Y}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right)\right)} \tag{2.11}$$

Let $$\theta=\tan^{-1}\left(\frac{Y}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right) \tag{2.12}$$

$$\theta=\tan^{-1}\left(\frac{0}{X^2+Y^2}\right)$$

$$\theta=0$$

$$\text{Cos }(0)=1 \tag{2.13}$$

Equation 2.11 becomes

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)} \tag{2.14}$$

Hence the Resultant R is given by

$$R=2\sqrt{X^2+Y^2} \tag{2.15}$$

This is same as the resultant of same complex conjugate vectors. The conclusion of the above result is that spectral components having magnitudes larger (speech components) than the magnitude of the factor β, the spectral components (speech components) remain unmodified.

Case II. The Resultant R, When $\beta >> \sqrt{X^2+Y^2}$

Equation 2.11 reduces to

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\left(\tan^{-1}\left(\frac{\beta}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right)\right)} \tag{2.16}$$

Let $$\theta=\tan^{-1}\left(\frac{\beta}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right) \tag{2.17}$$

$$\theta=\tan^{-1}\left(\frac{X(\beta-Y)}{X^2+\beta Y}\right) \tag{2.18}$$

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\theta} \tag{2.19}$$

In above equation the value of cos (θ) << 1, hence the resultant $R << 2\sqrt{X^2+Y^2}$ (2.20)

Resultant R is far less than the resultant of same vectors having zero phase angles. The conclusion of above result is that spectral components having magnitudes too smaller (noise components) than magnitude of the factor β, the spectral components get suppressed further. This results in enhancement of noisy speech signal. Experimentally determined values of the factor β as a function of input noisy speech SNR for white Gaussian noise, train noise and babble noise for which noisy speech enhancement can be obtained are given in Table. 1.

Table.1. Experimentally determined values of β as a function of input noisy speech SNR for white Gaussian noise, train noise and babble noise.

TYPE OF NOISE			
SNR (dB)	AWGN	TRAIN	BABBLE
0.0	0.0070	0.70	0.490
5.0	0.0090	1.00	0.500
10.0	0.0200	1.80	0.70
15.0	0.0400	2.00	0.90

III. EXPERIMENTAL DETAILS

3.1. Speech Quality Evaluation by Subjective Listening Tests and Spectrogram Analysis

Six normal hearing subjects of age group 20-25 years and six subjects with moderate hearing loss of age group 70-75 years participated in listening tests. The improvement in Mean Opinion Score (MOS) is computed and is plotted in Figure.3 and 4 respectively. Spectrogram analysis is also carried out and is as shown in figure.5.

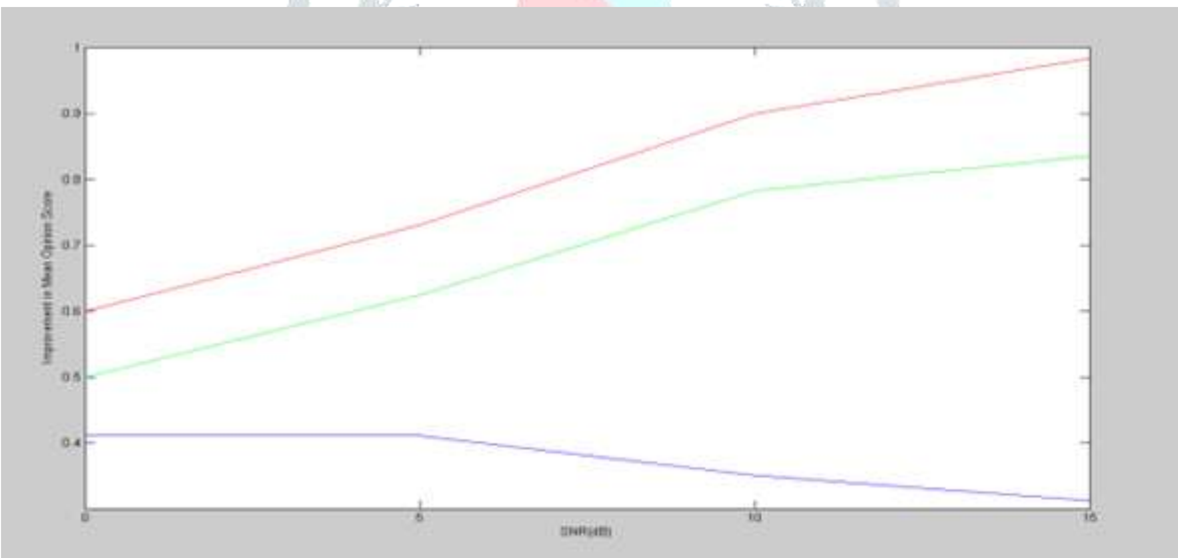


Figure.2. Improvement in Mean Opinion Score as a function of input noisy speech SNR for white Gaussian noise, train noise and babble noise in case of listening tests on Normal Hearing subjects.

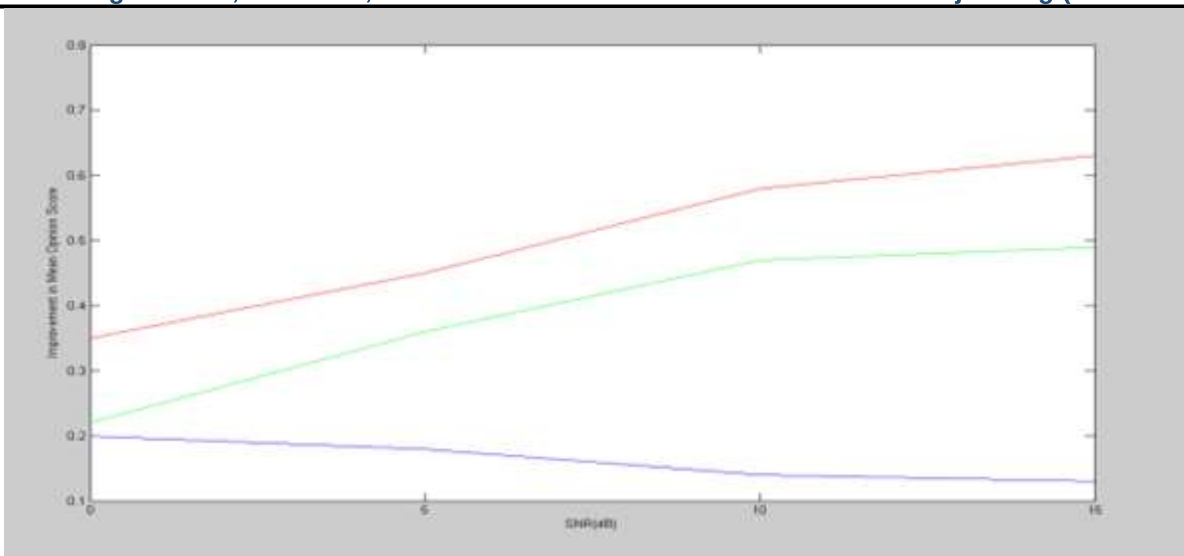


Figure. 3. Improvement in Mean Opinion Score as a function of input noisy speech SNR for white Gaussian noise, train noise and babble noise in case of listening tests on subjects with moderate hearing loss.

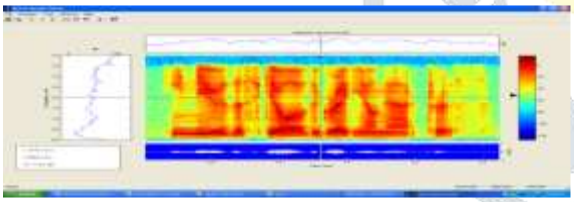
IV. RESULTS

Improvement of Mean Opinion Scores in case of normal hearing subjects and subjects with moderate hearing loss shown in Fig.2 and Fig.3 reveals that the proposed technique of speech enhancement performs satisfactorily in the presence of AWGN compared to the presence of train and babble noise. The spectrograms shown in Fig. 4, indicates that the enhanced speech signal in case of AWGN does not show speech distortion, while back ground noise has been reduced. In case of train and babble noise though the noise is suppressed, a small quantity of signal distortion is also introduced. This is due to using the function C(k) as a constant across frequency, which practically varies with respect to frequency.

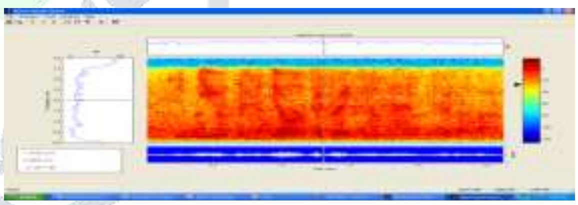
V. CONCLUSIONS

In this paper a computationally efficient noisy speech enhancement algorithm is presented. In the proposed work, the noisy speech signal magnitude spectrum is combined with altered phase spectrum to produce altered spectrum. During signal reconstruction, small energy (noise) components gets canceled out more as compared to more energy (speech) components, thus resulting in speech signal enhancement. The proposed method is tested by subjective listening tests (on both normal hearing and subjects with moderate hearing loss) and spectrogram analysis. This proposed work will find application in hearing aids as a front end technique to suppress the background noise associated with noisy speech signals.

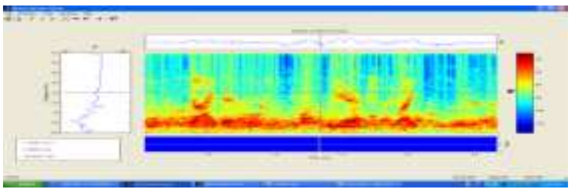
(a)



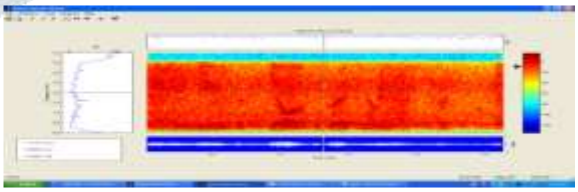
(b)



(c)



(d)



(e)

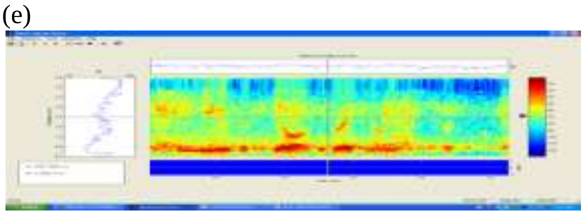


Fig.4. Spectrograms of (a) clean speech; (b, d) speech corrupted by babble noise, train noise respectively (0 dB SNR); (c, e) corresponding enhanced speech samples.

REFERENCES

[1] Richard J. Baker, Stuart Rosen "Auditory filters nonlinearity in mild/moderate hearing impairment" Journal of Acoustical Society of America, 111(3). March 2002, pp 1330-1339.

[2] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction,"IEEE Trans .Acoustics, speech, signal processing, vol. ASSP-27, NO.2,PP.113-120, Apr.1979.

[3] Y.Ephraim and D. Malah, "Speech enhancement using a minimum mean –square error log-spectral amplitude estimator, "IEEE Trans. Acoustics, Speech Signal Processing, vol.ASSP-33, no.2, pp.443-445, Apr.1985.

[4] Kamil Wojcicki, Mitar Milacic, Anthony Stark, James Lyons, and Kuldip Paliwal, "Exploiting conjugate symmetry of the short – time Fourier spectrum for speech enhancement, "IEEE Signal processing letters, Vol.15.2008. pp.461-464

[5]. K.Paliwal and A.Basu, "A Speech enhancement method based on kalman filtering", in proc. IEEE Int.Conf. Acoustics, Speech, and Signal Processing (ICASSP'87), Apr.1987, vol.12, pp.297 – 300.

[6].Y. Ephraim and H. V. Trees, "A Signal Subspace approach for speech enhancement" IEEE Trans. Speech Audio Processing., vol.3, no.4. pp. 251-266, July.1995.

[7]. J.Allen and L.Rabiner, "A Unified approach to short- time Fourier analysis and synthesis", Proc. IEEE, vol. 65, no.11, pp. 1558 -1564, Nov.1977.

[8].R.Crochiere, "A weighted overlap-add method of STFT analysis/ synthesis", IEEE Trans. Acoustics. Speech, Signal Processing., Vol ASSP-28, no.2, pp, 99- 102, Apr.1980

[9].D. Griffin and J. Lim," Signal estimation from modified Short-time Fourier Transform," IEEE Trans.Acoustics, speech, Signal processing. Vol. ASSP-32, no.2, pp.236 -243, Apr. 1984.

[10].M. Portnoff, "Short-time Fourier analysis of sampled speech, "IEEE Trans. Acoustics., Speech, Signal process. Vol. ASSP-29, no.3, pp.364-373, Jun.1981.

[11].Richard J Baker, Stuart Rosen, "Auditory filter non linearity in mild/moderate hearing impairment., Journal of Acoustical Society of America, 111(3), March 2002, pp.1330-1339.

[12].Pick G.F., Evans E.F. "Dissociation between frequency resolution and hearing threshold" In hearing- Physiological basis and Psychophysics, Edited by R. Klinke and R.Hartmann.pp.393-398.

[13].Rosen S. Baker R.J, Darling A.M. "Auditory filter non linearity at 2kHz in normal listeners" Journal of Acoustical Society of America, 103, pp.2539-2550

[14].Tyler R.S, Hall J.W, Glassberg B.R, Moore B.C.J, Patterson R.D. "Auditory filter asymmetry in the hearing impaired". Journal of Acoustical Society of America, 76, pp.1363-1368.

[15].B.C.J Moore. "Speech processing for the Hearing impaired: Successes, Failures and Implications for Speech mechanisms" Speech communication, Elsevier 41(2003) pp.81-91.

[16].Jesko L.Verhey, Anne Kristin Anweiler, Volker Hohmann, "Spectral loudness summation as a function of duration for hearing impaired listeners, International Journal of Audiology.2006; 45; pp.287-294.