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 SIGNAL ENHANCEMENT USING SIGNAL-VECTOR ANALOGY

[1] Ms. Madhumati Shridhar Kuntoji, [2] Dr. Veena Devi S.V.

[1] Student, M.Tech, [2] Associate Professor,

[1] Department of Electronics and Communication Engineering, R.V.College of Engineering, Bangalore, Karnataka, India,

Abstract : Sensory neural hearing loss results in loss of frequency selectivity [1]. Less is known about effect of sensory neural hearing loss on level dependence of frequency selectivity which is a prominent feature in normal hearing [1]. Auditory filter shape is nearly linear in hearing impaired whereas it is nonlinear in normal hearing people; hence frequency discrimination capability is poor in hearing impaired listeners, especially in noisy environments. Speech enhancement is utmost important and crucial for hearing impaired. In this paper a simple and computationally efficient speech enhancement algorithm is presented. Most of the speech enhancement algorithms use the magnitude of STFT while keeping the phase as it is [3]. In contrast in this paper the magnitude of STFT of noisy speech is kept as it is while the phase is modified. Modified spectrum of speech is obtained by combining unchanged magnitude spectrum and modified phase spectrum [3]. This modification results into cancellation of low energy components (noise) more, than the high-energy (speech) components during signal reconstruction. This results in reduction of noise, which is essential in hearing aids as a front end operation. Objective speech quality measures, Subjective listening tests on normal hearing and hearing loss subjects and spectrogram analysis of speech samples show that the proposed method gives improved speech quality. In the proposed method, enhancement is achieved by modification of phase angle of DFTs of speech through their imaginary parts. The algorithm is computationally efficient.

Index Terms **– Speech Enhancement, Short Time Fourier Transform, Magnitude Spectrum, Phase Spectrum, Phase Compensation**

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 signal to noise ratio leading to increased 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 STFT of equation (1.1) we get

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

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 literature survey it is found that, some of the speech enhancement methods modify the magnitude spectrum of noisy speech, whereas phase spectrum is kept as it is. The processed magnitude spectrum is combined with unprocessed phase spectrum during signal synthesis. In the proposed work, the magnitude of noisy speech STFT is kept as it is, whereas phase is modified [3]. The unprocessed magnitude is combined with modified phase to get modified spectrum. This results into cancellation of low energy (noise) components more than high energy (speech) components [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.

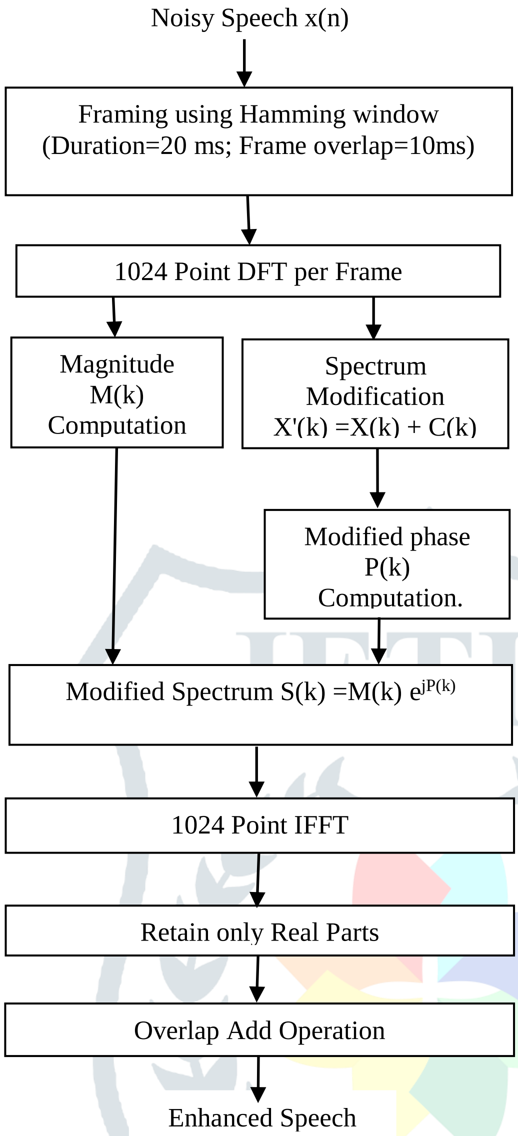

Figure1. Block Diagram of the Proposed Method

2.1 Explanation

The block diagram of the proposed method is shown in Fig.1. The input noisy speech signal x(n) is real hence its DFT obeys conjugate symmetry i.e. $X(k) = X^{*}(N-k)$. IDFT of X(k) results into original noisy speech signal x(n) due to cancellation of imaginary parts of complex conjugate terms. The degree of cancellation or reinforcement of complex conjugates can be controlled by modifying their phase [3]. Here in the proposed work phase modification is done through imaginary parts of DFTs.

2.1.1 Phase Modification through Imaginary Parts of DFTs.

Assuming N to be even, an imaginary constant C(k) given by $C(k) = j\beta \; ; 0 \leq k < N/2$ (2.1)

$C(k) = -j\beta \; ; N/2 \leq k \leq N-1$ (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 noisy speech signal is modified as follows. The noisy speech signal STFT X(k) is modified as

$X'(k) = X(k) + C(k)$ (2.3)

The modified phase of X'(k) is computed and further combined with magnitude of original noisy speech signal to get modified complex spectrum given by

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

The IDFT of above complex spectrum results into enhanced signal. Since conjugate symmetry is disturbed due to phase modification, IDFT of modified spectrum results into complex signal. Only real parts are retained for further processing. The explanation is as follows. Fourier analysis resolves a signal s(t) into a weighted sum of sinusoidal, that is to say a sum of complex conjugates. The magnitude spectrum of DFT of a real signal obeys even symmetry whereas phase spectrum obeys odd symmetry. The complex conjugates sum together to result a real signal due to cancellation of their imaginary parts during signal synthesis. The degree of cancellation or summation of these complex conjugates can be controlled by modifying their phase. Here phase modification is enforced via imaginary parts of complex conjugates. This approach yields reduction in computational complexity. The above process can be visualized using signal – vector analogy. Considering a pair of complex conjugate numbers $C_1=X+jY$ and $C_1^*=X–jY$ both having same magnitude given by

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

and phase angles

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

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

The modified complex numbers are given by

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

$$C_{11}^*=X–jY-j\beta. \tag{2.8}$$

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

The modified complex numbers can be written as $C_{1P}=\sqrt{X^2+Y^2}\ e^{j\tan^{-1}(Y+\beta/X)}$ *and* $C_{1P}^*=\sqrt{X^2+Y^2}\ e^{j\tan^{-1}(-Y-\beta/X)}$ *(2.10)*

the resultant of above two complex conjugate numbers is the function of the factor β and is given by

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

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

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\left(\tan^{-1}\left(\frac{Y}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right)\right)} \tag{2.12}$$

$$\text{Let } \theta=\tan^{-1}\left(\frac{Y}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right) \tag{2.13}$$

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

$$\theta=0$$

$$\cos(0)=1 \tag{2.14}$$

Equation 2.12 becomes

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

Hence the Resultant R is given by

$$R=2\sqrt{X^2+Y^2} \tag{2.16}$$

This is equal to the resultant of same complex conjugate vectors. The implication of above result is spectral components having magnitudes more (Speech components) than magnitude of β the spectral components (speech components) remain unaltered.

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

Equation 2.11 reduces to

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\left(\tan^{-1}\left(\frac{\beta}{X}\right)-\tan^{-1}\left(\frac{Y}{X}\right)\right)} \tag{2.17}$$

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

$$\theta=\tan^{-1}\left(\frac{X(\beta-Y)}{X^2+\beta Y}\right) \tag{2.19}$$

$$R=\sqrt{2(X^2+Y^2)+2(X^2+Y^2)\cos\theta} \tag{2.20}$$

In above equation the value of cos (θ) << 1, hence the Resultant $R \ll 2\sqrt{X^2+Y^2}$ *(2.21)*

Resultant R is far less than the resultant of same vectors having zero phase angles. The implication of above result is spectral components having magnitudes much smaller (noise components) than magnitude of β, the spectral components get suppressed more. This leads to enhancement of signal. Empirically determined values of β as a function of input speech SNR for white Gaussian noise, train noise and babble noise for which enhancement can be achieved are shown in Table. 1.

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

NOISE TYPE			
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 Database

In the experimental evaluation the NOIZEUS speech corpus is used [11]. This includes 30 phonetically balanced sentences from 3 male and 3 female speakers. The quality of data base is equivalent to telephone speech quality. The sampling frequency being used is 8 kHz .The data base is composed of noisy speech samples (train and babble noise) at 0,5,10 and 15 dB SNRs. It also includes pure speech and noise samples. Pure speech samples corrupted by additive white Gaussian noise at 0, 5, 10, and 15dB are generated using MATLAB. For experimentation stimuli corrupted by train, babble, and additive white Gaussian noise at above mentioned four SNR levels are used.

3.2. 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 calculated in both the cases and is plotted in Figure.3 and 4 respectively. Spectrogram analysis is also carried out and is shown in figure.5.

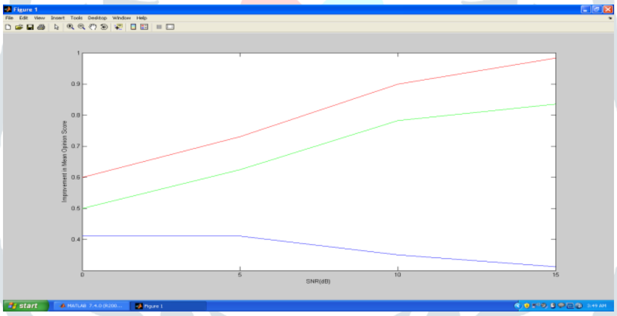

Figure.3. Mean Opinion Score improvement as a function of input speech SNR for white Gaussian noise, train noise and babble noise in case of listening tests on Normal Hearing subjects.

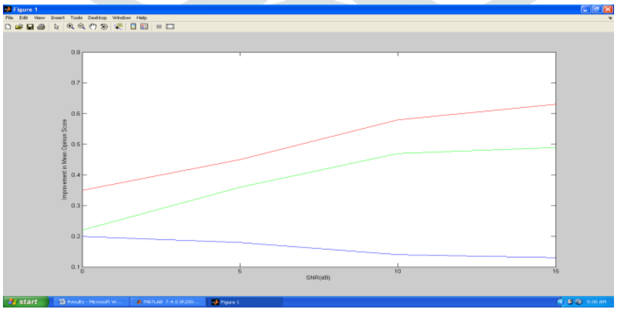

Figure 4. Mean Opinion Score improvement as a function of input speech SNR for white Gaussian noise, train noise and babble noise in case of listening tests on subjects with Moderate Hearing loss.

IV. RESULTS

Improvement of Mean Opinion Scores in case of normal hearing subjects and subjects with moderate hearing loss shown in Fig.3 and 4 indicate that the proposed method performs best in case of additive white Gaussian noise as compared to train and babble noise cases. The results of spectrogram analysis shown in Fig. 5 indicates that the enhanced signal in case of white noise does not exhibit speech distortion, while back ground noise has been attenuated. In case of train and babble noise though the noise is suppressed a small amount of signal distortion is also introduced. This is because using function C(k) as a constant across frequency even though it is function of frequency.

V. CONCLUSIONS

In this paper a computationally efficient speech enhancement algorithm is presented. In the proposed work, the noisy speech signal magnitude spectrum is combined with modified phase spectrum to produce modified complex spectrum. During signal synthesis low energy (noise) components cancel out more as compared to high energy (speech) components, thus resulting in signal enhancement. The proposed method is validated by subjective listening tests (on both normal hearing subjects and subjects with moderate hearing loss), as well as spectrogram analysis. This proposed work can find application in hearing aids as a front end algorithm to suppress the background noise.

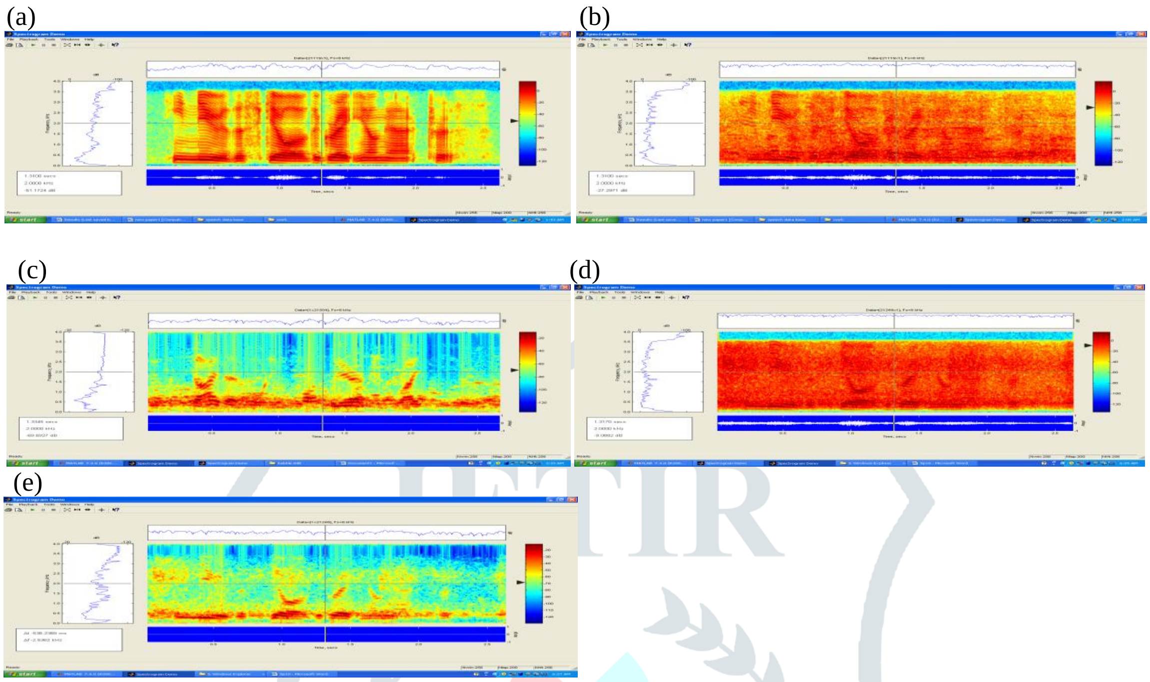

Fig.5. Spectrograms of speech sample sp02.wav "He knew the skill of the great young actress." by a male speaker from NOIZEUS speech data base:(a) clean speech; (b, d) speech corrupted by babble noise, train noise respectively (0 dB SNR); (c, e) corresponding enhanced speech samples.

VI. 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, MitarMilacic, 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.