iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 2017 **Issue:** onference **Month of publication:** September 15, 2017

DOI:

Development of Algorithm to Improve Speech Perception for Hearing Impaired

Mr.Mayur N Ingale[1], Mr.Swapnil S Kharat[2], Mr. Udayraj Chaudhari[3]
[1, 2, 3] Dept.of Electronics & Telecommunication Engg.,SSJCET,Asangaon, India[M.S.]

Abstract: Voice conversation is one of the most important tools for communication whereas some people don't have profit from this opportunity because of their hearing loss. Approximately 10% of the population suffers from some hearing loss, however only a small percentage of these categories actually use a hearing aid device. The problem associated with auditory system is Sensorineural hearing impairment deals with widening of auditory filters in human ear. It leads to increased unwanted masking effect, loudness recruitment and degraded speech perception. Binaural dichotic presentation can reduce the effect of spectral masking and can help in improving speech perception by person suffering from stated loss. The various speech splitting techniques based on filter like critical band filters, comb filters having complementary magnitude response, constant bandwidth filters and 1/3 octave bandwidth filters have been proposed by many researchers in the last several years. However, there are not enough analysis and comparison on the previous researches. Many systems have critical problems understanding different voices and do not have robust vocabularies. This paper is based on study of these earlier speech processing algorithms for restoring a sensor neural hearing loss. And to give optimistic solution to minimize bilateral inner ear losses significantly using new algorithms such as wavelet transform & source subtraction.

Keywords: Sensorineural hearing loss, Spectral splitting, Wavelet Transform, Spectral Subtraction.

I. INTRODUCTION

The peripheral auditory system of the human ear behaes like a bank of bandpass filters called auditory filters or critical bands. One of the hearing impairment in humans is sensorineural hearing loss due to defects in the cochlea and auditory nerve. The characteristics of this loss are elevated hearing threshold, abnormal growth in loudness perception with increase in intensity, reduced frequency and temporal resolution and increased spectral and temporal masking. Reduction in spectral contrasts, results in broadening of auditory filters. The peaks and valleys of the speech spectrum are broadened affecting the perception of speech because of masking. Increased temporal masking results in the increase of forward and backward masking of weak acoustic segments by strong ones, which also affects speech intelligibility. Purely conductive losses are those resulting from dysfunction of the ear canal or middle-ear structures, so that less acoustic energy reaches the auditory receptors in the cochlea.[1] But sensorineural loss is difficult to cure and it becomes progressively worse with time. This leads to a decrease in frequency resolving capacity of the auditory system of the ears. The sensorineural impairments are characterized by high frequency hearing loss, increase in the threshold of hearing, compression in dynamic range, severity of temporal masking, and loss of spectral resolution due to spread of masking[5]. From different investigations it was noticed that splitting the speech into different bands and presenting the alternate bands to each ear as shown in figure was advantageous for people with moderate bilateral sensorineural hearing impairment, with residual hearing in both ears. Also in this project, a novel technique for speech enhancement is described. The method uses a combination of the Wavelet transform with "Wiener filtering" in the wavelet domain & spectral subtraction. Results indicate that the proposed method provides better speech enhancement for hearing impaired. human ear is a very complex organ. The function of the ear is to transduce acoustic energy into electrical energy which may be perceived by the brain as sound. The basic structure is divided into three parts. Outer ear, middle ear and inner ear. The outer ear consists of the pinna and the external auditory canal (EAC), it receives the signal and directs it towards the middle ear. The middle ear is an air-filled pouch that contains the eardrum and ossicles. The normal eardrum is sealed at its edges and effectively separates the outer and middle ears. It encodes the signal at different frequencies, makes different nerve cells resonate and transmit short pulses to the brain. Figure 1 shows a pictorial view of the human ear.

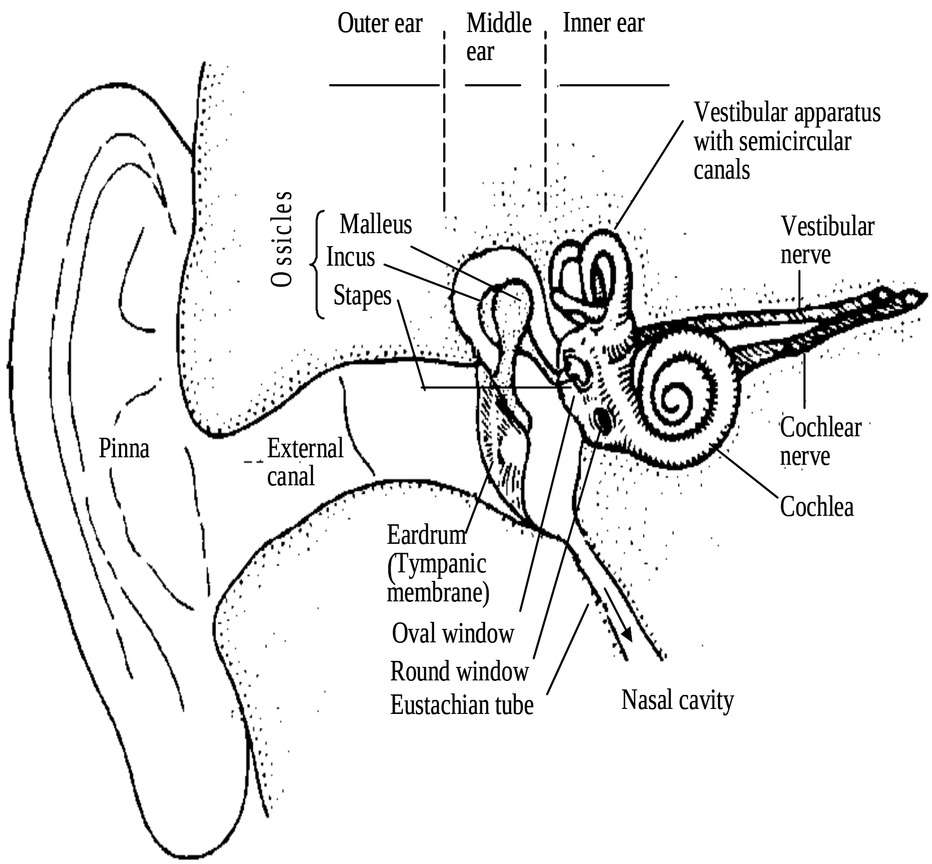


Fig.1. Structure of human ear.

The frequency range of hearing for a normal human is 20 Hz to 20 kHz. Human hearing is most sensitive in the range of 1 kHz to 4 kHz. There are two important parameters which effects hearing. One is loudness that is the intensity of sound and the other one is pitch, which is the frequency of the fundamental component in the sound. Human ears are designed to perceive sound on a logarithmic frequency and amplitude scale. The doubling of frequency will be perceived the same no matter what the frequency will be, for e.g doubling from 200Hz to 400Hz will be perceived the same as from 2kHz to 4kHz. Same is true for amplitude, if we double the amplitude with the same ratio no matter what the amplitude is, it will be perceived as same. We normally express sound level in dB. Normal speech is around 60dB. If the ear is exposed to 85dB the ear is at risk and the sound pressure level is harmful. Above 140dB the hearing will most certainly get damaged.

A. Hearing losses

The problems in any one part of the auditory system can result in different types of hearing loss(HL). These problems can occur in one or both ears in which case we may refer to them as unilateral or bilateral, respectively. The losses are conductive, sensorineural, retro cochlear.

B. Conductive Hearing Loss(CHL)

Conductive hearing losses are created when it is difficult for sounds to reach an otherwise normally functioning cochlea. This means that a problem exists somewhere between the outer and middle ears. Excessive earwax in the ear canal, fluid and/or infection in the middle ear (otitis media), damaged eardrums, and broken ossicular chains are examples of problems that can cause can conductive losses. Most (not all) hearing losses of a conductive nature can generally be repaired or medically treated. Losses of up to 60 dB can result from uncorrected conductive hearing losses. CHL can be treated surgically as it associated with outer ear.

C. Sensor neural Hearing Loss(SHL)

In sensorineural hearing loss the problem is either associated with the inner hair cell, auditory nerve, or both. But the term SHL implies a disorder in the inner ear. SHL are the most common of the different types of hearing loss (HL). The hair cells are susceptible to damage from ototoxicity, from aging, and from long-term exposure to loud sounds. Age-related HL (presbycusis) may have genetic implications. Many syndromes that interrupt, arrest, attack, or breakdown the structure of the inner ear can cause an array of problems. The inner and outer hair cells play different roles in hearing activity. The inner hair cells communicate directly with the auditory nerve, whereas the outer hair cells amplify soft sounds via changes in fluid pressure in the inner ear. It is entirely possible for the inner and outer hair cells to be affected independently of one another. For example, most people with HL have outer hair cell damage and little to no inner hair cell damage. People with only outer hair cell damage make very good hearing aid candidates, because the hearing aid replaces the amplification role of the outer hair cells. However, the ability to understand speech degrades with increasing damage to the inner hair cells. As there is currently no method for repairing damaged cochlear hairs, treatment for SHL usually involves amplifying the incoming sound.

D. Retro cochlear Hearing Loss(SHL)

Retrocochlear hearing losses refer to hearing disorders associated with the auditory nerve to the auditory centers of the brain. Progressive neural disorders, such as multiple sclerosis (demyelination), can obscure sound information being sent by the cochlea. Tumors developing on the auditory nerve can wipe out large frequency regions within the range of speech sounds, making it difficult understand speech. In the event that hearing loss is due to problems with the auditory nerve, an individual may elect to receive an auditory brainstem implant (ABI).

E. Fixed Hearing Loss(MHL)

As name indicates one can have one or more than above stated loss. MHL occurs when there are interruptions in both the Conductive & sensorineural pathways.

II. LITERATURES FROM VARIOUS RESEARCHERS

To improve speech perception (SP) by person suffering from bilateral hearing impairment, several studies have investigated the dichotic presentation, by spectrally splitting the speech signal using different techniques.

Laura A. Drake (1993) developed wavelet based multiband dynamic compression technique to compensate loudness recruitment problem. In that work they calculated gain separately for each wavelet coefficient based on its level of intensity. Also results are compared with multiband amplitude compression and found same results in both the schemes and concluded that wavelet processing was more efficient and pleasant. The scheme is using synthesized vowels in which the first formant was presented to left ear and the second formant to the right ear. They used comb filters which having constant bandwidth of 700Hz and stop band attenuation 40dB. The peoples in age group 39-69 were subjects for testing bilateral and moderate SHL. The gain of the filters was adjusted depending on the hearing loss of individual subject. The scheme was found good for SP even in noisey environment [1].

Further work is carried by splitting the speech into two complementary spectra by using bank of critical band filters to reduce spectral masking in cochlea. They splitted speech signal in nine bands using 128-coefficient linear phase FIR comb filters for binaural dichotic presentation using band pass filters. The critical bandwidths selected are the auditory filter bandwidths reported by Zwicker[2]. The processing was done by digitally filtering the speech signal, digitized with 12-bit resolution at 10 k Sds. Testing is done with vowel-consonant-vowel & consonant vowel syllables for twelve English consonants The listening tests carried on five subjects with SHL and found increase in recognition scores and reception of speech. The proposed design was able to improve speech quality, response time, recognition scores and transmission of consonantal features as observed in the listening tests [3].

The inter-aural switching method with switching frequency of 50Hz with step & trapezoidal fading(duty cycle 70%) is used to switch speech bands for dichotic presentation. For temporal splitting, they used a symmetrical inter-aural switching with a frequency of 50 Hz to switch the odd and even bands alternately between the two ears. They have reported that temporal splitting did not contribute to improvement in scores. The scheme is used to reduce the problem such as temporal masking & to increase speech intelligibility. The five normal subjects associated with HL are participated in tests their ages ranged between 20-35 years. All the subjects had pure tone hearing thresholds less than 20 dB HL in the frequency range of 125-6000 Hz. The test stimuli were digitally recorded in an acoustically isolated room, using a microphone, amplifier, and lowpass filter with cutoff frequency of 4.8 kHz, at a sampling rate of 10 k Sa/s with 16-bit resolution. Test results showed that there was maximum improvement in information transmission & reduced temporal masking. They also reported that slower transitions resulted in better reception features [4].

There was still performance gap exists in speech recognition between those who have received cochlear implant and people with a normal hearing capability. The new method investigates the application of an improved signal processing method called bionic wavelet transform (BWT). This method is based upon the auditory model and allows for signal processing. Consonant recognition results in 15 normal hearing subjects show that the BWT produces significantly better performance than the earlier methods. Based on neural network simulations, a better recognition rate in processing the speech signals can be achieved by BWT in comparison with BPF. BWT has better tradeoff between time and frequency resolutions, which would result in the reduction of the number of channels required for CIs and the reduction of the stimulation duration required for a word [5].

Then to improve results the comb filters based on auditory critical bands to minimize spectral distortion are implemented & evaluated. The design used 256-coefficient linear phase FIR comb filter using frequency sampling technique to obtain low pass band ripples such as 1 dB & high stop band attenuation of 30 dB. The comb filter is designed such way that to reduce masking and improve SP. The listening tests involving closed set identification twelve vowel-consonant-vowel syllable were conducted and compared with comb filters with sharp transitions. The new comb filter gives better speech recognition scores, better transmission and reduced spectral masking [6]. P.N.Kulkarni in (2008) proposed a scheme in which critical bands corresponding to auditory filters based on psychophysical tuning curves were used. Filter banks corresponding to eighteen critical bands over 5kHz frequency range were used. 512-coefficient linear phase FIR Comb filters were designed having adjustable magnitude response at transition crossovers for minimizing any change in perception intensity. Also reduction in pass band and increase in stop band was considered. The SP gets degraded due to increased spectral and temporal masking in persons with SHL.

Work was carried out on filters with constant, critical and 1/3 octave bandwidth. Binaural dichotic scheme was performed using each type of filters for comparative study. Results indicated that constant bandwidth filters help in speech intelligibility. SP was significantly improved for critical band based filter and 1/3 octave filters [7]. Some efforts are being made in implementing filters on FPGA for hearing impaired. The decimation filter used for hearing aid applications were implemented on FPGA. The implementation was carried out using the canonical signed digit representation. Each digital filter structure was simulated using MatLab, and its complete architecture was captured using DSP blockset and Simulink. The filter was implemented on Xilinx FPGA using Virtex-II technology. The resulting architecture was hardware efficient and consumed less power compared to conventional decimation filters [8].

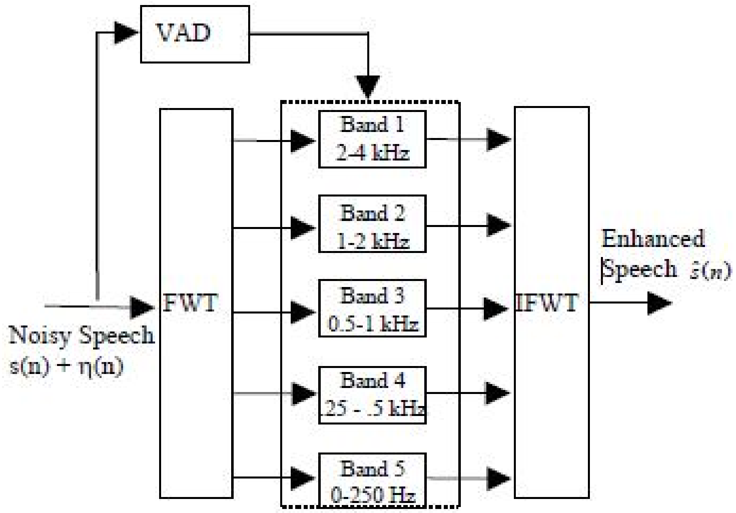
III. PROPOSED METHODOLOGY

In the scheme of the spectral splitting, sensory cells corresponding to alternate filter bands on the basilar membrane are always stimulated. In the scheme of spectral splitting for binaural dichotic presentation, two complementary comb filters with pass bands corresponding to critical band auditory filters were used. In their scheme, 18 bands were used to cover the frequency range of 5 kHz. Comb filters were 128-coefficient linear phase FIR filters, designed for sharp transitions between bands. Comb filters with responses corresponding 18 critical bands were designed as 512-coefficient linear phase FIR filters, using frequency sampling technique. In the investigation of first scheme of splitting, these comb filters are used. Other method uses a combination of the Wavelet transform with "Wiener filtering" in the wavelet domain & spectral subtraction for better speech enhancement for hearing impaired.

IV. FWT-BASED SPEECH ENHANCEMENT SYSTEM

In this Section, the speech enhancement method proposed in the project will be described in more detail. The system architecture is shown in Fig. 10.1 The first stage is the processing of the noisy speech signal using a Fast Wavelet Transform (FWT). A four level FWT decomposition is performed resulting in five subbands covering the frequency range from 0 - 4000 Hz. The output of fast wavelet structure is set of wavelet coefficient. A noisy speech input signal results in noisy wavelet coefficients at the output of the FWT decomposition. Fig.2 shows FWT-based Speech Enhancement system.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September2017- Available at www.ijraset.com



V. WAVELET DENOISING USING WIENER FILTERING

were compressed. This method is based on "Weiner filtering" in the wavelet domain, as opposed to the well-known Weiner filtering in the frequency domain[13] . Wiener gain, ki, is calculated using the following equation:

ki=Si2/Si2+Ni2

where Si2 is the speech energy and Ni2 is the noise energy in band i.

Noise segments were detected using a voice activity detector, and the noise power was calculated as follows:

Ni2=1/L ∑j=1L(1/M∑l=1m [ni(l,j)]2)

where M is the number of wavelet coefficients in a frequency band, L is the number of noisy frames over which the noise power is averaged and ni(l,j) are the wavelet coefficients associated with sub-band i for the noise input during frame j. The signal plus noise power was calculated during speech regions as the sum of the squares of the wavelet coefficients. The noise power is then subtracted from the speech plus noise power to obtain the speech power. With these power values, the Wiener gain can be calculated for Each frequency. If the estimate of the noise power is greater than the estimate of the signal plus noise power, then ki for that band may be set to zero or a small value. If dij is the jth noisy wavelet coefficient in band i, then the denoised wavelet coefficient is given by dij(denoised) = dij . ki

These denoised wavelet coefficients are used to reconstruct the speech signal. When the speech power is much greater than the noise power, which would normally be the case during a voiced speech frame, then ki »1. If the speech and noise powers are comparable in value, then ki » 0.5; this would normally be the case during onset and offset of voicing, or in unvoiced regions.[13]

VI. SPECTRAL SUBTRACTION ALGORITHM

Speech signals from the uncontrolled environment may contain degradation components along with required speech components. The aim of speech enhancement is to improve the quality and intelligibility of degraded speech signal. Main objective of speech enhancement is to improve the perceptual aspects of speech such as overall quality, intelligibility and degree of listener fatigue. Improving quality and intelligibility of speech signals reduces listener's fatigue; improve the performance of hearing aids, cockpit communication, videoconferencing, speech coders and many other speech systems. Quality can be measured in terms of signal distortion but intelligibility and pleasantness are difficult to measure by any mathematical algorithm. Perceptual quality and intelligibility are two measures of speech signals and which are not co-related. The proposed method also consist of speech signal enhancement using basic spectral subtraction among the all available methods because spectral subtraction algorithm is the historically one of the first algorithm, proposed for background noise reduction[15]. The greatest asset of Spectral Subtraction Algorithm lies in its simplicity.

The standard spectral subtraction method is described in the following equations. A short-term noise spectral magnitude is subtracted from a degraded speech signal by

$$Y_i(\omega) = H_i(\omega) X_i(\omega)$$

Where

$$Hi(\omega)= Di(\omega)/Xi(\omega)$$
$$Di(\omega)= | Xi(\omega)| - |Bi(\omega)|$$

The The Xi(ω) is a short-term spectral estimation of speech (frame), Bi(ω) a short-term estimation of noise (frame i), Xi(ω) a smoothed-out estimate of the corrupted magnitude at time i, Bi(ω)| a smoothed-out estimate of the noise magnitude at time i, and Yi(ω) is a clean speech estimate.

The magnitude | Xi(ω)| and |Bi(ω)|can be computed from following equations

$$|Bi(\omega)| = \lambda b \, |Bi\text{-}1(\omega)| + (1\text{-} \lambda b)| \, Bi(\omega)|$$
$$|Xi(\omega)| = \lambda x \, |Xi\text{-}1(\omega)| + (1\text{-} \lambda x)| \, Xi(\omega)|$$

Where the values of the memory factors λb, λx are found in intervals $0.1 < \lambda x < 0.5$ and $0.5 < \lambda b < 0.9$.

For better quality the noise suppression technique should have low algorithmic delay. Its computational complexity should be low to permit its implementation on a low-power processor. Spectral subtraction is a single-input speech enhancement technique developed for use in audio codecs and speech perception[15]. It involves estimating the noise spectrum, subtracting it from the noisy speech spectrum, and resynthesizing the enhanced speech signal.

A used technique using cascaded-median based continuous updating of the noise spectrum, without using voice activity detection, is presented for speech enhancement by spectral subtraction.

Presented spectral subtraction technique is for real-time speech enhancement in the aids used by hearing impaired listeners. For reducing computational complexity and memory requirement, it uses a cascaded-median based estimation of the noise spectrum without voice activity detection. The technique is implemented and with sampling frequency of 12 kHz, processing using window length of 30 ms with 50% overlap, and noise estimation by 3-frame 4-satge cascaded median. Spectral subtraction for enhancement of speech corrupted by additive noise involves estimating the magnitude spectrum of the noise and using it for estimating the magnitude spectrum of the speech signal. The enhanced magnitude spectrum along with the phase spectrum of the noisy speech is used to resynthesize the enhanced speech.

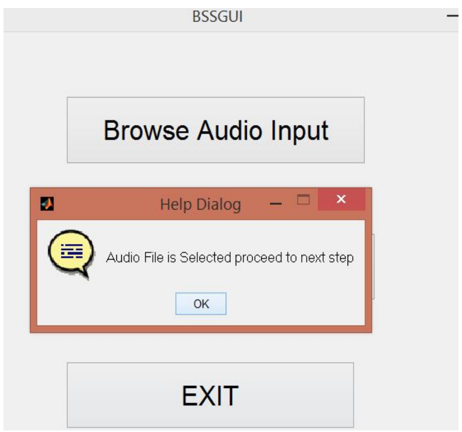
Using Discrete wavelet transform algorithm noise signals present in the information are efficiently minimized but for better results DWT is combined with spectral substraction algorithm to make better speech perception for hearing impaired listeners.

VII. RESULTS

In proposed speech enhancement algorithm two different types of signals has been tested as follows. In the first case, clean speech was passed through the enhancement algorithm followed by chirp signal.

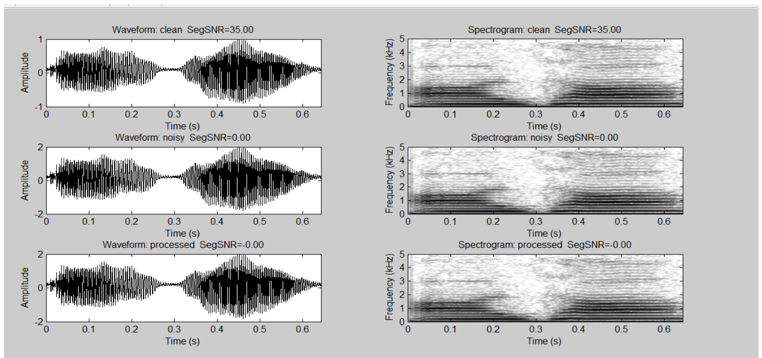
A. For Clean signal

After running the main implemented file you will get the GUI shown below



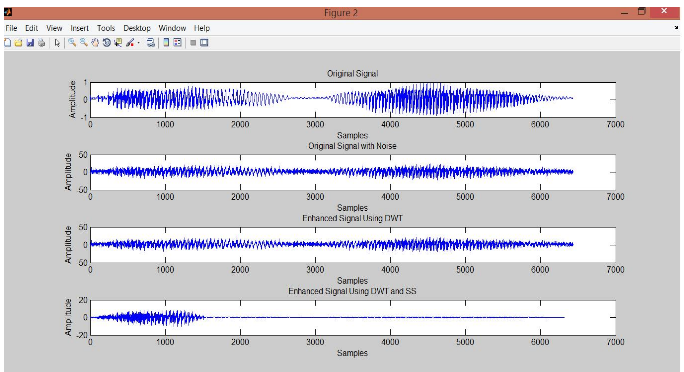
B. GUI for main

Now after pressing Browse Audio Input button of GUI, select Clean.wav file from database and hit Enter. Clean,Noisy & processed signals frequency spectrum & their respective spectrograms are displayed as shown in figure by pressing Process DWT with SS button.



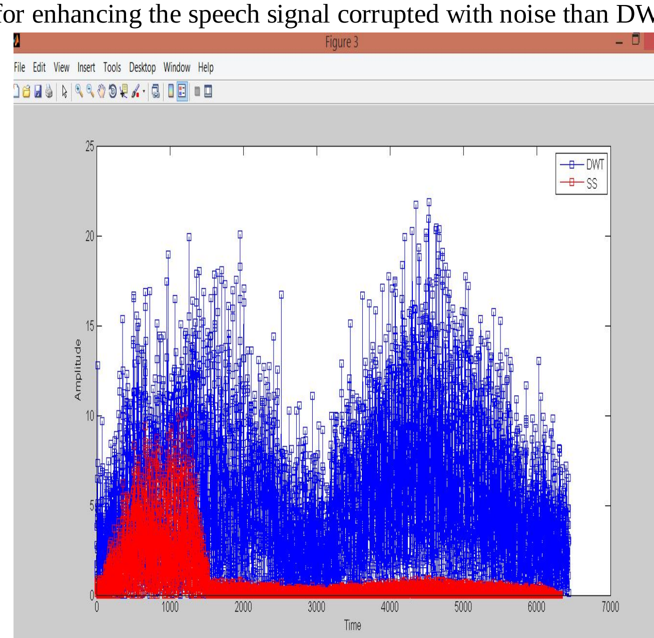
Frequency Spectrum & Spectrogram of clean signal

Figure Below represents the spectrum for Original signal, Original Signal with Noise,Enhanced signal using DWT and Enhanced signal using DWT and SS.



Frequency Spectrum for Original signal, Original Signal with Noise, Enhanced signal using DWT and Enhanced signal using DWT and SS

Above figure shows in combination of DWT and SS you will get better enhanced signal than using DWT alone.
The comparison of DWT and DWT+SS methods for Enhanced signals using histogram. Here also you will find that the combination of DWT and SS works efficiently for enhancing the speech signal corrupted with noise than DWT alone.



Histogram representation of DWT & SS

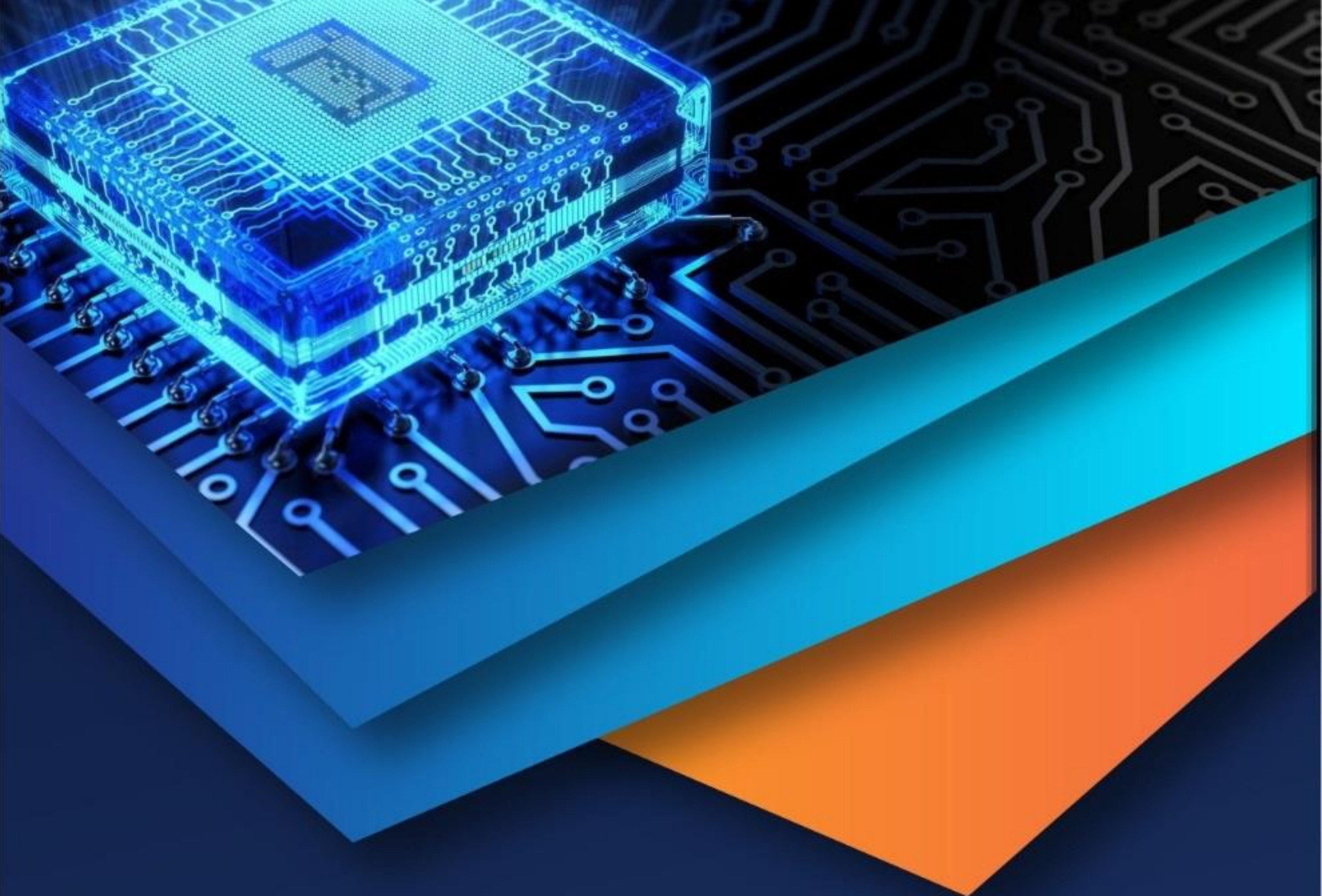
International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor:6.887
Volume 5 Issue IX, September2017- Available at www.ijraset.com

The obtained results for above mentioned methods are listed in tabulated form in Table-V below

Sr. No	Signal Type	METHODS					
		DWT			DWT+SS		
		PSNR	Entropy	Mean	PSNR	Entropy	Mean
1	Clean	68.66	0.39	8.40	105.54	1.81	0.37
2	Chirp	65.38	0.94	13.71	91.61	2.74	0.92
3	Noisy	68.66	0.67	8.42	105.29	1.8	0.3
4	Processed	65.41	0.6	13.66	91.61	2.98	0.92
5	Late-Noise	79.68	3.18	2.49	115.5	5.67	0.11

REFERENCES

[1] CHABA. (1991). "Speech-perception aids for hearing-impaired people: Current status and needed research," J. Acoust. Soc. Am. vol. 90, 637–683.
[2] Drake, L. A., Rutledge, J C and Cohen, J. (1993). "Wavelet analysis in recruitment of loudness compensation" IEEE Trans. On Signal Processing, vol. 41, 3306─3309
[3] E. Zwicker, "Subdivision of audible frequency range into critical bands (Frequenzgruppen)", J. Acoust. Soc. Am., vol. 33, p. 248,1961.
[4] Chaudhari, D. S., and Pandey, P. C. (1998). "Dichotic presentation of speech signal with critical band filtering for improving speech perception," Proc. IEEE Int. Conf. Acoust., Speech and Signal Processing (ICASSP'98), Seattle, Washington, AE 3.1.
[5] D. S. Jangamashetti and P. C. Pandey(2000), Proc. 4th World multi conference on systemics, cybernetics and informaics (SCI'2000), Orlando, Fl. USA 2000, pp.346-353
[6] Jun Yao, & Yuan-Ting Zhang*(2002), "The Application of Bionic Wavelet Transform to Speech Signal Processing in Cochlear Implants Using Neural Network Simulations", IEEE Transactions on biomedical engg.,vol. 49, no. 11,November 2002.
[7] Alice N.Cheeran P. C. Pandey (2002), "Design of comb filters based on auditory filter bandwidths for binaural dichotic presentation with sensorineural hearing impairment", 0-7803-7503-3/02/ IEEE DSP,971-974.
[8] P. N. Kulkarni & Pandey(2008)," Optimizing the Comb Filters for Spectral Splitting of Speech to Reduce the Effect of Spectral Masking", IEEEInternational Conference on Signal processing, Communications and Networking Madras Institute of Technology, Anna University Chennai , Jan 4-6, 2008. pp69-73.
[9] Jaya chandwani & chaudhari (2013)," Improving Speech Perception to Sensorineural Hearing Impairment Using Filters", International Journal of Advanced Research in Computer Science and Software Engineerisng, Volume 3, Issue 3, March 2013.
[10] Jedrzej Kocinski. Speech Intelligibility Improvement Using Convolutive Blind Source Separa- tion Assisted by Denoising Algorithms. Speech Communication, Elsevier, 2007, 50 (1), pp.29.
[11] Chaudhari, D. S., and Pandey, P. C. (1998). Splitting of Speech Signal by Critical Band Filtering for Bilateral Sensorineural Hearing Impairment, xviiith annual conference of indian association of biomedical scientists (iabms)New Delhi, India, Oct. 22-24, 1997
[12] Philipos C. Loizoua and Arunvijay Mani, "Dichotic speech recognition in noise usingreduced spectral cues"(April 2003).
[13] E. Ambikairajah, G. Tattersall and A. Davis. "Wavelet Transform-Based Speech Enhancement"
[14] Babak Nasersharif & A Akbari, "Application of Wavelet Transform and Wavelet Thresholding in robust Sub-Band Speech Recognition",Eusipco(2004).
[15] Santosh K. Waddi, Prem C. Pandey, and Nitya Tiwari "Speech Enhancement Using Spectral Subtraction and Cascaded-Median Based Noise Estimation for Hearing Impaired Listener"(2013).



10.22214/IJRASET

45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY