Audibility Extender & Smart Compression Technique in Digital Hearing Aid

[1] Sheetal Nana Patil
Assistant Professor
Department of Information Technology,
R C Patel Institute of Technology, Shirpur (MS) India

[2] Prashant G Patil
Research Scholar
Department of Electronics Engineering
Manoharbhai Patel Institute of Technology, Gondia (MS)

Abstract: Speech intelligibility improvement in hearing aid users is a challenging and complex issue for researchers. Several techniques had been developed, modified & implemented by many researchers during last few decades. Many researchers suggest an innovative and useful solution to overcome these difficulties, but benefits of technologies has numerous limitations in few performance areas. Children, from the age group of 5-10 years with hearing aid face several challenges in their lives, struggling to learn and communicate in mother tongues, sign language and attitudes face from parents and society. Children adopting hearing aid face both negative and positive attitudes. The nature and harshness of the disability affect development directly by imposing limitations on the child's functioning. Emotional and social responses in the child and child's social environment are indirectly affected by the disability. To overcome above challenges, there is a need to develop a novel technique for improving performance of hearing aid and speech intelligibility of user.

***Index Terms* – Hearing Aid, Speech Intelligibility, Compression, High Frequency, Hearing Loss**

I. INTRODUCTION

Hearing loss is associated with age and exists in infants, teenagers, young adults and elders. Hearing is the primary sense of action by which we learn speech and mother tongue first. The ability to hear clearly after 6 to 8 months of birth is important to learn speech, language skills and auditory processing skills. Normal person without any hearing loss can perceive a wide range of sounds in terms of pitch and loudness level in any background environment.

High-frequency hearing loss (Partial Deafness) occurs due to damaged sensory hearing cells in cochlea die as shown in Fig 1.1. These hair cells are responsible for interpreting the sounds from ears and convert them into electrical impulses & brain understand as noticeable sound. Cochlea is located in the inner part of ear which perceives high-frequency sounds and the hair cells that perceive low-frequency sounds are located near the top. Because of this, hearing loss typically affects the higher frequencies before it affects the lower frequencies. Speech intelligibility is a degree of understandable speech for different background conditions. Intelligibility is affected by the loudness level, excellence of the speech signal, amount of background noise, and echo. Speech intelligibility is applicable to various fields, including phonetics, human factors, acoustical engineering, audiometer and assistive devices for hearing disabled.

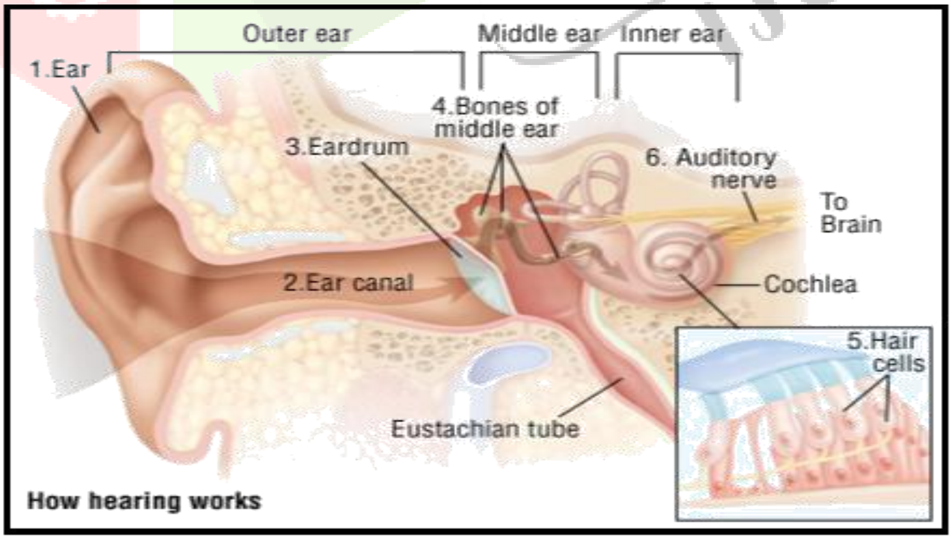

Figure 1- Ear structure with Outer, Middle and Inner Ear functioning

Over the past years many techniques were proposed and implemented by several researchers to improve speech intelligence ability, but unable to improve in all aspects like noise control, intensity control, gain adjustment , speech background, gender wise performance variation and speech frequencies impact. Hence, this research work is planned to develop the novel algorithm for improving speech intelligence ability of hearing aid. Today, 83% of all hearing aids available in commercial market are based on digital signal processing. Digital Signal Processing (DSP) is the ever growing and most powerful technology, which has already shaped the science and engineering at the end of twenty-first century. Revolutionary changes have already been made in a broad range of fields of communications, medical imaging, radar and sonar and high fidelity music reproduction, to name just a few. Many researchers have suggested that they have reached the limit of what Digital Signal Processing can do; Digital Signal processors are as powerful as they can be and no more digital features left to design. Redesign and modification of processor will affect performance of hearing aid. The motivational factors discussed in the previous section encouraged to formulate the aim of this research work. The aim of this research work is to design, develop and implement a novel algorithm to improve speech intelligibility of Marathi Speech in Hearing Aid users.

- Speech processing in background noise
- Automatic gain control (AGC)
- Acoustic feedback cancellation
- Real time processing delay
- Parameter estimation & setting for algorithm
- Selection of comfort hearing zone.
- Audiologist Correct analysis

In Multiband compression approach different stationary and active characteristics are applied for various frequency bands. This is applicable to HA users who needs different compression at different frequencies. Multiband frequency compression method normally use syllabic compression, more number of channels (N=2-16). Multiband frequency compression cuts height of spectral peaks, boost the level of spectral valleys and flatten the spectral shapes. Since spectral peaks and valleys help in identifying speech sounds, spectral flattening makes it harder for the hearing-aid users to identify the place of diction of consonants. Challenges faced during improvement in speech intelligibility are classified into two aspects; scientific and developmental aspects. Scientific aspects are primary, which are considered before design, development and implementation of any method or technology.

II. LITREATURE REVIEW

Ying-Yee Kong and Ala Mullangi develops frequency lowering system for enhancing place of articulations in which they proposed 8 octave multiband analysis filter structure with centre frequency from 1 KHz to 5 KHz, adjacent filter outputs are combined to form 4 analysis bands. Output levels of these bands were measured with averaging times of 20 ms, and were then used to determine the output levels of low-frequency narrow-band noise signals.

The noise synthesis filters whose centre frequencies ranged from 397 to 794 Hz. In addition, the levels of each of the four low-frequency noise bands were attenuated to minimize any masking effect on the original speech signal. The four low-frequency narrow-band noise signals were then summed and added to the original speech signal. The results were achieved a high level of accuracy for fricative classification. Overall accuracy was similar for fricative classification for both the design 79% correct and test 82% correct sets. Joshua M. Alexander implemented both NLFC and Frequency transposition (FT) algorithm for individual recognition variability measurement. If FT is active, the algorithm continually searches for the most intense spectral peak in source band. The source band is one-third-octave band frequencies from 630 to 6,000 Hz. Number of 10 source and target bands are used in this methodology. Results show that consonants recognition increases for FT method as compared with NLFC method. Ching et al; have explained FT based Speech recognition of hearing-impaired listeners. Where the speech scores measured using the Speech Intelligibility Index (SII). Scaling the index by a multiplicative proficiency factor was found to be inappropriate, and alternative modifications were explored. The data were best fitted using a method that combined the standard level distortion factor, which accounted for the decrease in speech intelligibility at high presentation levels based on measurements of normal-hearing people with individual frequency-dependent proficiency. Matthias Milczynski et al; has developed the Frequency transposition signal processing approach named F0 modulation. F0 provides an enhanced temporal pitch cue by amplitude modulating the multichannel electrical stimulation pattern at the F0 of the incoming speech signal. Word and sentence recognition tests were carried out in quiet and in noise.Danielle Glista et al; proposed multichannel nonlinear frequency compression (NLFC) algorithm in which compressing is applicable above cut off frequency, constant gain of 10-30 dB is added for 2-4 KHz frequency zone. Fig 2.2 shows nonlinear frequency compression method with operational parameters. Speech thresholds were calculated together with octave and inter-octave frequencies between 250 and 6000 Hz. They conducted four objective tests for English language such as aided speech sound detection, consonant recognition, plural recognition, and vowel recognition. People with a precipitous high frequency hearing loss often miss the high frequency information even when they wear hearing aids. Sometimes it is because the high frequency gain available on the hearing aid is not sufficient to reach audibility before feedback occurs; sometimes the severity of the hearing loss in the high frequency region is so great that it is UN aid able or “dead” from the complete depletion of inner hair cells. In the former case, audibility may be achievable at the expense of a smaller vent diameter on the hearing aid. This could compromise wearer comfort because of an increase in the occlusion effect. In the latter case, acoustic stimulation of the unaidable region may decrease further the already depressed speech understanding. The loss of audibility of high frequency sounds often compromises speech understanding and the appreciation of music and nature’s sounds (such as bird songs).

The early attempts on frequency lowering were designed to achieve easy implementation of existing technologies rather than to achieve the desired signal processing results. Many were not even practical enough to be implemented into hearing aids. While lowering the frequencies, these methods also altered other aspects of speech known to be important for perception. Some of these approaches created unnatural sounding speech, distorted gross temporal and rhythmic patterns, and extended durations (slow playback) of the speech signals. Others created reversed spectrum (amplitude modulation based techniques) which is difficult to even recognize as speech by inexperienced listeners. In vocoder-based systems, both analysis and synthesis were often carried out using only a limited number of frequency bands, which resulted in unnatural speech sounds.

III. PROPOSED MODIFIED COMPRESSION METHOD

A key objective when fitting hearing aids to children is to maximize the audibility of high frequency speech cues which are critical in the understanding of spoken language. Recent advances in digital signal processing have enabled the development of hearing aids which offer linear frequency transposition as a new way of accessing these important speech sounds. Transposition of acoustic information from higher to lower frequencies may help people with severe or profound high frequency hearing loss, especially when a 'dead region' is present. The Audibility Extender is different from other frequency lowering schemes in several aspects

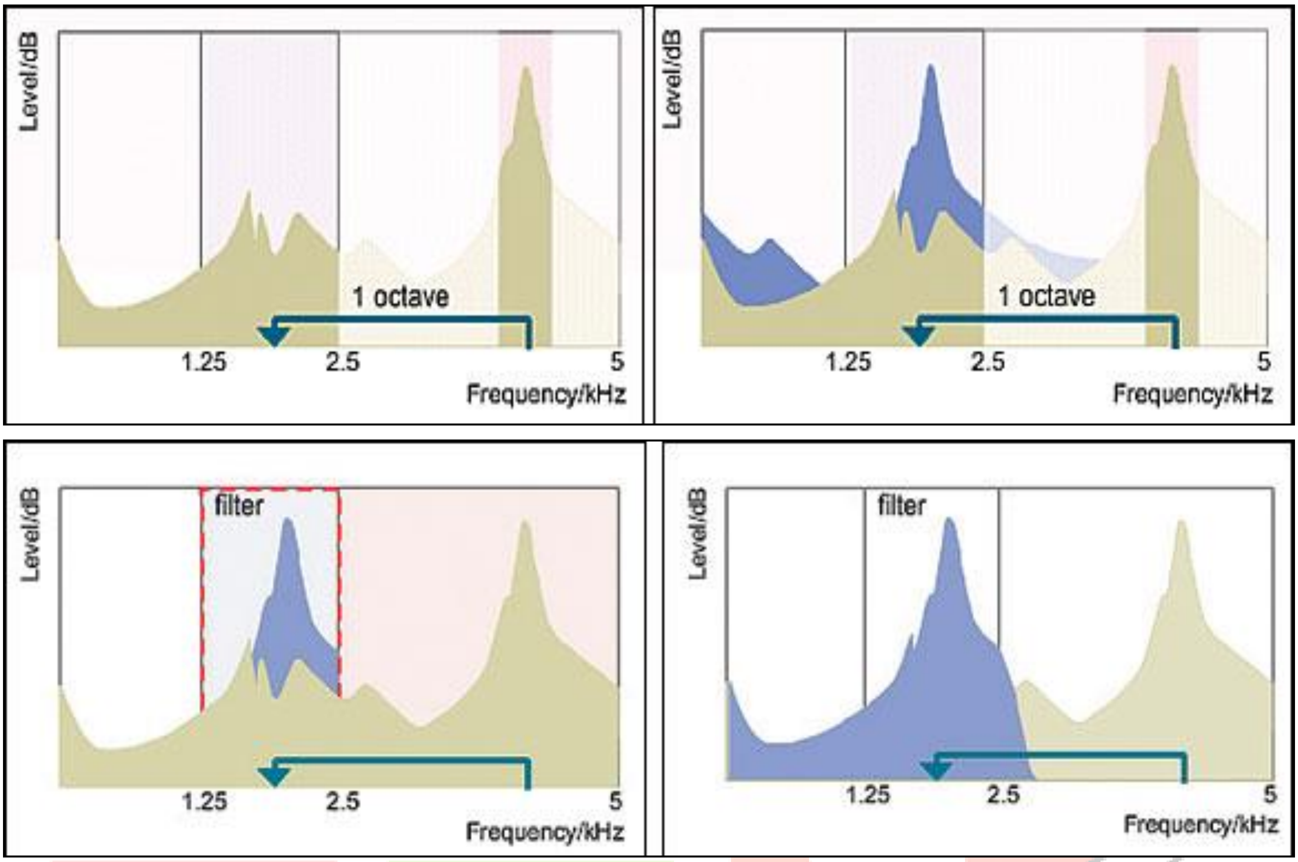

Figure 2- Target & Source Band Octaves for Speech Signal Processing in Proposed Algorithm.

It transposes only the high frequency sounds (above the start frequency) regardless of their voicing characteristics (e.g., voiced or voiceless). Thus, it is equally effective on periodic and periodic sounds. Systems that are active only for voiceless signals may miss high frequency periodic signals including music and bird songs.

- It is active during all segments of speech and not at specific linguistic segments, (eg, voiced versus voiceless).
- Typically only one octave (although two octaves may be allowed) of high frequency sounds above the start frequency is transposed to a lower octave.
- Frequencies higher and lower than the transposed region are filtered. This limits the amount of masking and avoids the need for compression.
- For simple stimuli, it preserves the transition cues and the harmonic relationship between the transposed signal and the original signal.
- This preserves as much of the original signal as possible. The transposed signal is mixed with the original signal to give a richer, more "natural" sound perception.
- Systems that do not overlap the transposed sounds would risk "exaggerating" any unnaturalness of the transposed sounds.
- By transposing frequencies linearly, the temporal structure of the signal is preserved. Thus, it can be easily recognized as the original source signal but at a lower frequency.

A key objective when fitting hearing aids to children is to maximize the audibility of high frequency speech cues which are critical in the understanding of spoken language. Recent advances in digital signal processing have enabled the development of hearing aids which offer linear frequency transposition as a new way of accessing these important speech sounds. Transposition of acoustic information from higher to lower frequencies may help people with severe or profound high frequency hearing loss, especially when a 'dead region' is present.

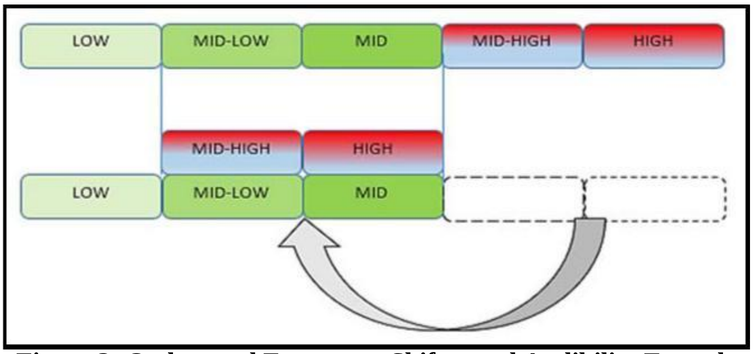

Figure 3- Orthogonal Frequency Shifter and Audibility Extender

Parameters Selection in Extender-

Target Band: The start frequency is the frequency where the hearing loss is unaidable. Thus, instead of amplifying that frequency, the AE transposes it without amplification. The target band in FT should be carefully determined. If the target band is too low, transposition might disturb the existing low-frequency components and negatively affect hearing for users with residual hearing at low frequencies. If the target band is too high, transposition is not meaningful since it will not preserve high-frequency information at sufficiently low frequencies that hearing-impaired listeners might be able to access.

Source Band: A fixed source band in FT may not adequately represent different speech signals with varying frequency components. Thus, we propose an adaptive source band on a frame-by-frame basis. It was shown that the first spectral moment M1, which can represent the centred of the spectrum, well characterizes the consonants.

Gain: Amount of amplification needed to the input signal to reach audible level of hearing aid user. Input decibel and gain should be controlled in a way to limit output dB level up to sweet hearing threshold level. Maximum loudness will affect uncomfortness in HA users.

Frequency Response Curve: Frequency response is a behavioural nature of hearing aid output relates to input at fixed intensity. Input Level & processing gain is kept constant for correct estimation of HA frequency response.

Peak Clipping: Peak clipping is a technique which controls or limits the maximum output of a hearing aid. Output dB level in HA is linearly increases with Input dB level. Once the output dB level reaches a Maximum level, the hearing aid will unable to produce a louder speech. Maximum output is the highest possible signal that a hearing aid is capable of delivering, regardless of the input level or the gain of the hearing aid.

Compression Threshold (CT) and Threshold Knee (TK): The Sound Pressure Level above which the hearing aid begins compression is referred as the compression threshold. High Threshold knee (TK) is used to control the output of HA for the loudness discomfort levels, maximize listening comfort, while Low TK used to improve audibility of soft speech and loudness perception.

Compression Ratio (CR): Compression ratio is defined as the change in input level required to produce 1dB change in output level. Compression ratios should be greater than 1:1 or less than 1:1.

Second approach is the combined spectral subtraction and wavelet packet based thresholding (SSWPT) method to suppress the noise without introducing any perceptible distortion in the signal for the speech that has been corrupted extremely by high levels of noise like cock pit noise in air ground communication scenario. In this approach dual band spectral subtraction (DSSS) method with adaptive noise estimator is used as the pre-processor for reducing the noise initially.

The harmonic product spectrum (HPS) is the frequency spectrum of a harmonic signal. When reducing sampling rate, the spectrum with harmonics align with fundamental frequency. F0 is fundamental pitch, which is affected by age and gender. It is an important parameter for classifying male and female speech. Voiced speech of male varies from 80 to 190 Hz, adult female 165 to 255 Hz and for child more than 190 Hz. In our classification-based alphabets separation the fundamental pitch (F0) by gender wise, background wise is used during training to neural network. Pitch Range (PR) based feature set for investigation and background classification is used for detection of non speech sound detection.

Suppose x (n,) s (n) and v(n) are the observed noisy speech, the clean speech and noise signals respectively. We define a filter and such that our estimate of the clean speech signal is

$$s(n)=a(n).x(n) \qquad (1)$$

The optimal filter vector a = [a0 . . . aN−1]

Wiener filtering, spectral subtraction, subspace methods and Kalman filtering are popularly used approaches for noise reduction. In the following subsections, we discuss these methods, and study their differences and similarities. The Wiener filter obtains a least squares estimate of under stationary assumptions of speech and noise. The construction of the Wiener filter requires an estimate of the power spectrum of the clean speech and the noise In the Wiener filter approach, the optimal estimator is designed to minimize the mean squared error. Speech enhancement aims to improve speech quality by using various algorithms. It may sound simple, but what is mint by the word quality. It can be at least clarity and intelligibility, pleasantness, or compatibility with some other method in speech processing.

- Listen to it in the same environment and with the same hardware as the intended application.
- Listen to the whole range of speakers (female, male, children, and different ethnic groups) and languages as the desired application.
- Listen with the whole range of background noises in the intended application.

IV. EXPERIMENTAL RESULTS & OBSERVATIONS

VADs are accurate only at low noise levels. Low-energy parts of a signal are most likely non-speech (energy of noise is smaller than energy of noise + speech).Stationary parts of a signal are most likely non-speech (speech is non-stationary). In this section the review different types of noise removal techniques is described. The speech signal can be degraded because of the noise such as be periodic noise, wide band noise, and interfering speech. Stationary filters, adaptive filters, or transform domain filters are used for removing the periodic noise.

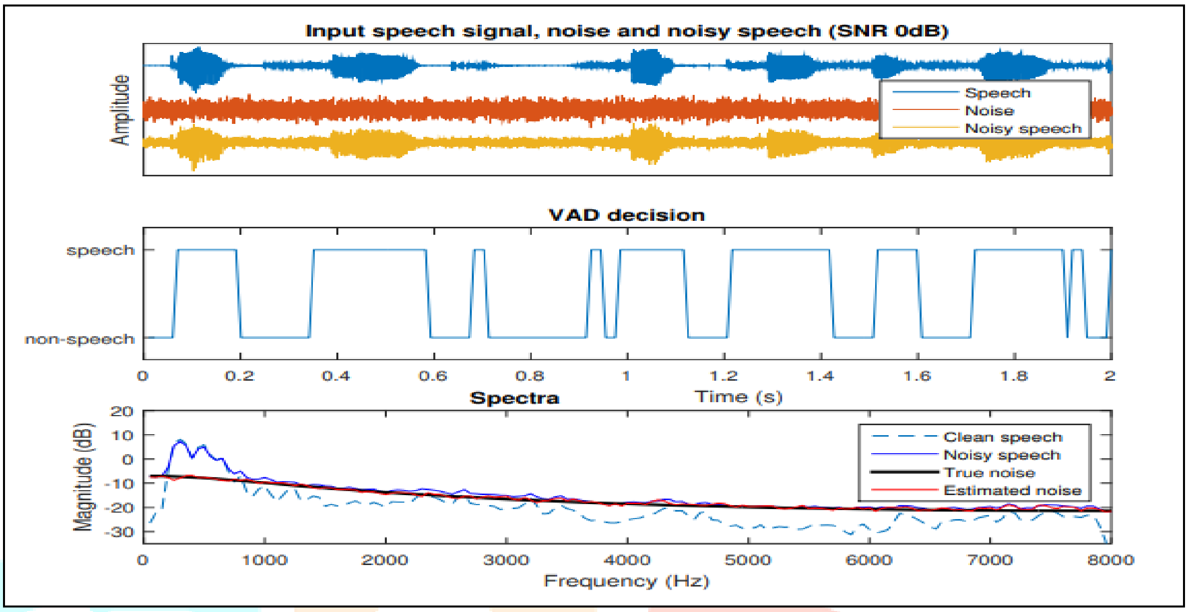

Figure 4- Noisy Speech Enhancement, VAD Decision & Spectra.

A working VAD (voice activity detection) in hand, giving values of zero and one as indicators of the voice activity in each frame, enables us to update the estimate of the background noise spectrum during the frames that have zero VAD.

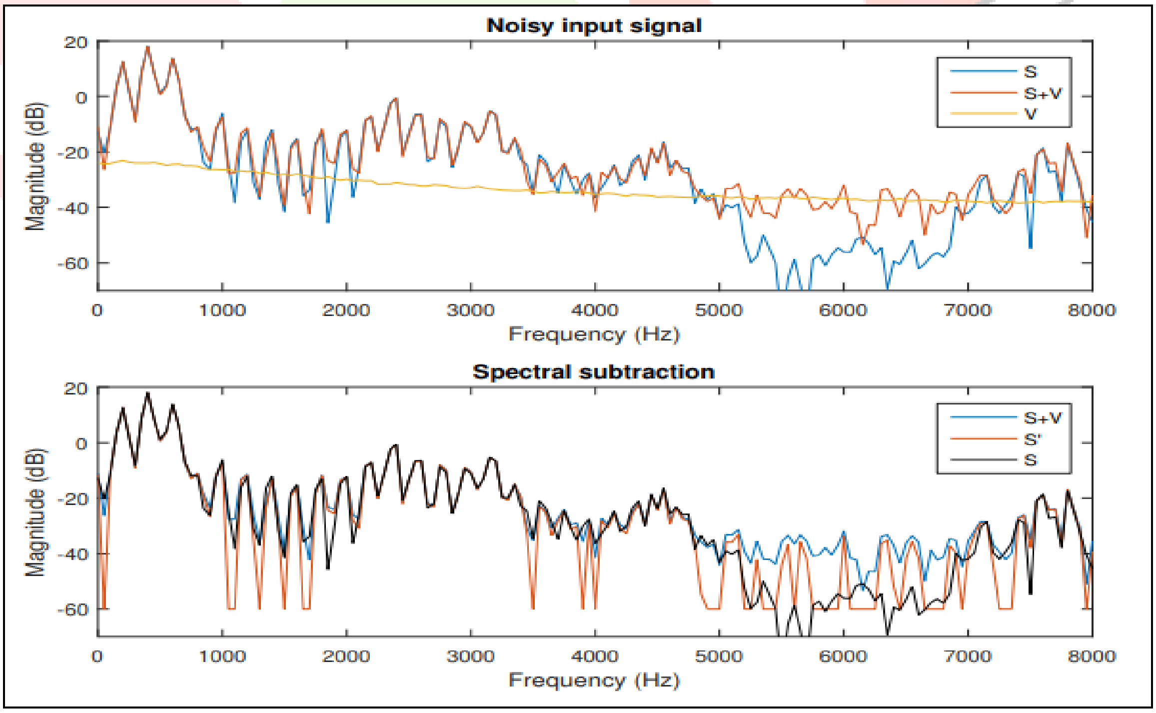

Figure 5- Noisy Input Speech & Spectral Subtraction Method

Speech enhancement aims at improving the quality of noisy speech. This is normally accomplished by reducing the noise (in such a way that the residual noise is not annoying to the listener), while minimizing the speech distortion introduced during the enhancement process. In this technique, the modulation domain has been investigated as an alternative to the acoustic domain for speech enhancement.

More specifically, it determines how competitive the modulation domain is for spectral subtraction as compared to the acoustic domain. For this purpose, the traditional analysis, modification, synthesis and framework to include modulation domain processing has been extended. Then it compensates the noisy modulation spectrum for additive noise distortion by applying the spectral subtraction algorithm in the modulation domain. Using an objective speech quality measure as well as formal subjective listening tests, it has been showed that the proposed method results in improved speech quality. Furthermore, the proposed method achieves better noise suppression than the MMSE method.

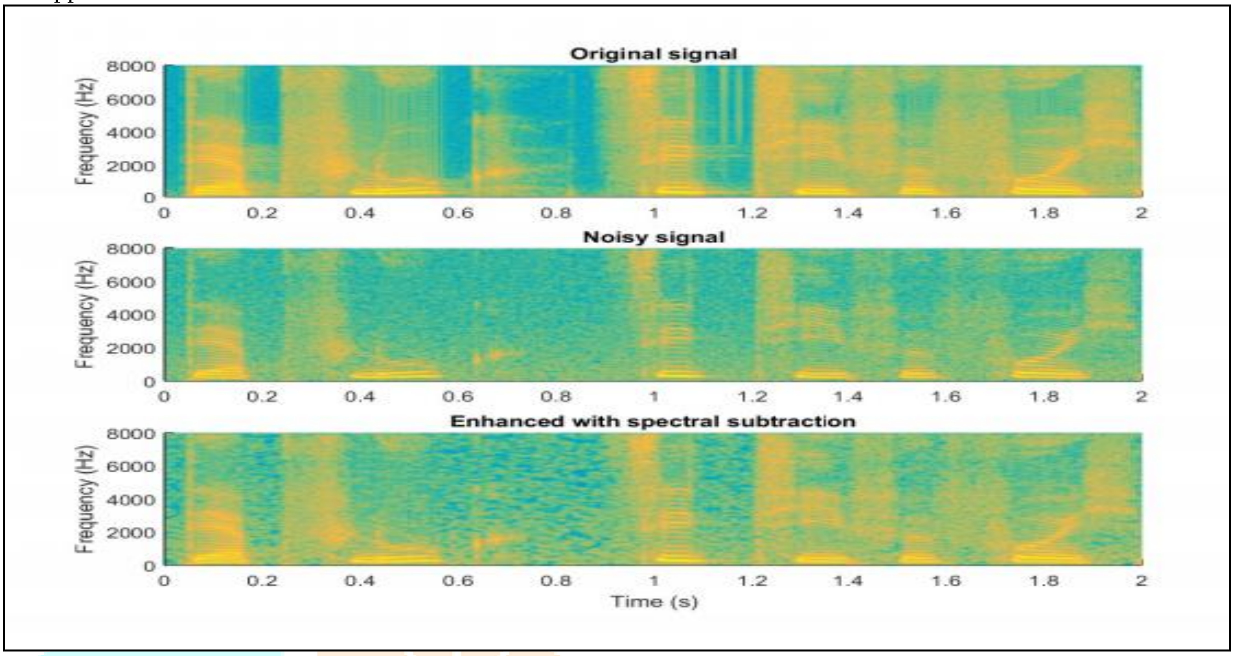

Figure 5- Spectrum of Original, Noisy & Enhanced Signal.

When the background noise is suppressed, it is crucial not to harm or garble the speech signal or at least not very badly. Another thing to remember is that quiet natural background noise sounds more comfortable than more quiet unnatural twisted noise. If the speech signal is not intended to be listened by humans, but driven for instance to a speech recognizer, then the comfortless is not the issue. It is crucial then to keep the background noise.

The bilinear approach reformulates the spectral envelope representation from e.g. line spectral frequencies to a two-factor parameterization corresponding to speaker identity and phonetic information. The spectral vector ysc, uttered by speaker s and corresponding to the phonetic content class c, is represented as a product of a speaker-dependent matrix As and a phonetic content vector b, using the asymmetric bilinear model.

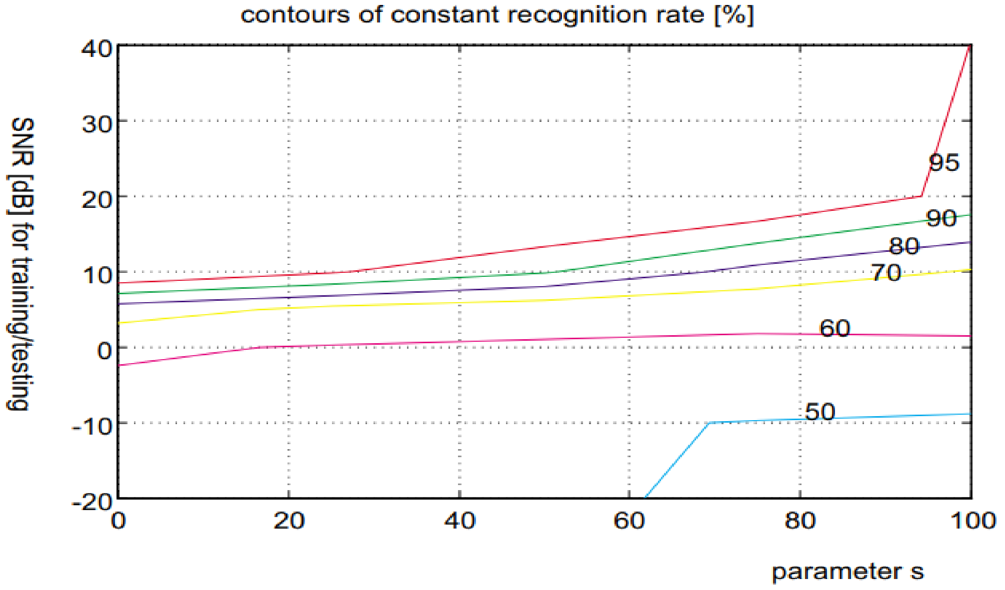

Figure 6- Consonant Recognition Rate Vs SNR of Input Speech.

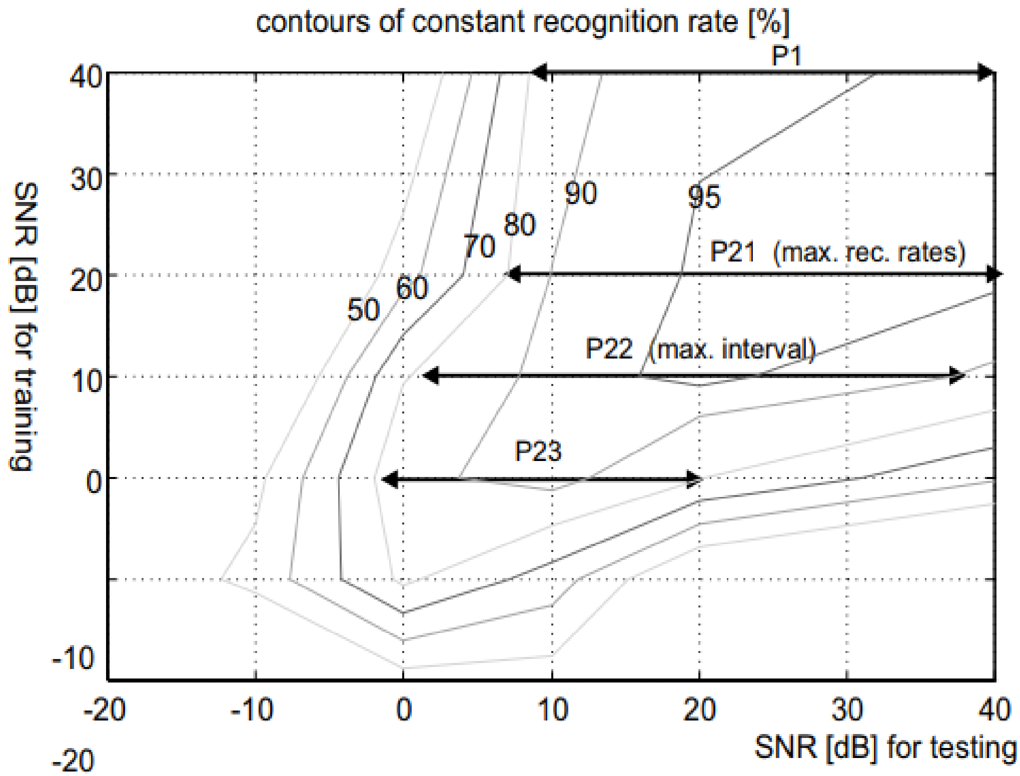

Figure 7- Consonant Recognition Rate Vs SNR of Input Speech during Testing & Training.

- By setting Compression factor or compression ratio, entire frequency range is compressed. Darker part is observed in spectrogram. In terms of pitch and intensity of input and processed consonant, it is observed that pitch variation occurs during compression. The larger variation is found in higher compression ratio.
- The change in nature of pitch frequency will leads to unnatural speech. It is harder to distinguish between male and female speech.
- The formant frequencies distribution F1, F2, F3 and F4 in both plots observed same.
- Frequency compression (M1) is beneficial for those patients who suffered from higher degree of high frequency hearing loss. Greater processing time is major challenge in this system, also gain to compression ratio relation to be controlled for effective implementation of algorithm.
- In terms of pitch and intensity of input and processed consonant, it is observed that pitch variation does not occurred during Transposition.
- Frequency Transposition (M2) is beneficial for those patients who suffered from Moderate degree of high frequency hearing loss.

V. CONCLUSION

In existing frequency compression (FC) and frequency transposition (FT) methods hearing aid users benefitted to limited extend. Both systems are designed by using FFT approach which creates higher processing time with considerable delay in between speakers and hearing aid user's communication. During speech processing through FFT, the speech signal become periodic and so in speech truncating process signal will tend to lose its original shape. Both systems needs higher gain during processing which leads to unequal intensity in processed speech, so random variation will degrades speech quality. Degradation in speech quality will leads to produce unnatural speech. Both systems has main drawback that they have less inability to preserve the spectral shape of the incoming signal. In frequency transposition, selection of dead region (DR) associated with patient is complex task and wrong selection will leads to unwanted processing. In frequency compression we need to control input output gain, compression ratio and threshold knee point within limiting range for intelligibility improvement towards hearing aid user's satisfaction. In both algorithm unwanted consonants with correlated noise are processed which degrades overall intelligence of system. To avoid this we need to classify wanted and unwanted alphabets with the help of neural classifier.

ACKNOWLEDGEMENT

This Work is granted & Funded by Vice Chancellor Research Promotional Scheme (VCRMS) under North Maharashtra University, Jalgaon.

REFERENCES

1. J Gou, J Smith, J Valero, I Rubio, “The Effect of Frequency Transposition on Speech Perception in Adolescents and Young Adults with Profound Hearing Loss”. Deafness & education international Vol. 13 No. 1, pp 17–33, March, 2011.

2. Yen-Teh Liu, Ronald Y. Chang, Yu Tsao, and Yi-ping Chang, “A New Frequency Lowering Technique For Mandarain-Speaking Hearing Aid Users” IEEE Global Conference on Signal and Information Processing, (Global SIP),pp 722-726,2015.

3. Joanna D. Robinson, Thomas Baer, et al, “Using transposition to improve consonant Discrimination and detection for listeners with severe high-frequency hearing loss”, International Journal of Audiology, Volume 46, pp 293-308, 2007.

4. Elisabeth Peltier, Cedric Peltier, Stephanie Tahar, “Long-Term Tinnitus Suppression with Linear Octave Frequency Transposition Hearing Aids”, PlosOne, Volume 8 Issue 1, 2013.

5. Hugh J. McDermott, Voula P. Dorkos, et. al, “Improvements in Speech Perception With Use Frequency-Transposing Hearing Aid”, Journal of Speech, Language, and Hearing Research, Volume 42, pp 1323-1335 , 1999.

6. Danielle Glista, “An Update on Modified Verification Approaches for Frequency Lowering Devices”, Journal of Audiology, Volume 67 Issue 7, pp 121-133, 2016.

7. Tobias Goehring, Federico Bolner, “Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users”, Hearing Research 344, pp 183-194, 2017.

8. Alexandre Enrique, Lucas Cuadra, Manuel Rosa, and Francisco Lopez-Ferreras. "Feature selection for sound classification in hearing aids through restricted search driven by genetic algorithms", IEEE Transactions on Audio, Speech, and Language Processing, Volume 15, Number 8, pp.2249-2256, 2007.

9. Veugen, Lidwien CE, Maartje ME Hendrikse, Marc M. van Wanrooij et al, "Horizontal sound localization in cochlear implant users with a contralateral hearing aid", Hearing research, Volume 336, pp. 72-82, 2016.

10. Finke, Mareike, Andreas Buchner, Esther Ruigendijk, Martin Meyer, and Pascale Sandmann "On the relationship between auditory cognition and speech intelligibility in cochlear implant users: an ERP study", Journal on Neuropsychologia, Volume 87, pp 169-181, 2016.

11. Siddala Vihari, A. Sreenivasa Murthy, et al. “Comparison of Speech Enhancement Algorithms”, Procedia Computer Science, Volume 89, pp 666 – 676, 2016.

12. Sang, Jinqiu, Hongmei Hu, Chengshi Zheng, Guoping Li, Mark E. Lutman, and Stefan Bleeck, "Speech quality evaluation of a sparse coding shrinkage noise reduction algorithm with normal hearing and hearing impaired listeners", Hearing research, Volume 327, pp 175-185, 2015.

13. James M. Kates and Kathryn H. Arehart, “The Hearing-Aid Audio Quality Index (HAAQI)” IEEE/ACM Transactions on audio, speech, and language processing, Volume 24, Number 2, pp 354-365, 2016.

14. Jorgen Schurz, Simone Wollermann, Hendrik Husted, “Classification of environmental sounds for future hearing aid applications”, 28th Conference on Electronic Speech Signal Processing, pp 294-299, 2017.

15. Zoghlami N, Lachiri Z, “Application of perceptual filtering models to noisy speech signals enhancement”, Journal of Electrical and Computer Engineering. Volume 12, pp 89-95, 2015.

16. Deepika, M. Vidya, “Background Noise Reduction using FFBPNNLM Network and Adaptive Filter”, International Journal of Innovative Research in Computer and Communication Engineering Volume 5, Issue 3,pp 107-117, April 2017.

17. Ing Yann Soona, Soongee Koha, Chai Kiat Yeo, “Noisy speech enhancement using discrete cosine transform”, Speech Communication, Volume 24, pp 249-257,2016.

18. Ferda Ernawan, Nur Azman Abu and Nanna Suryana, “Spectrum Analysis of Speech Recognition via discrete Cosine transform”, Proceedings. Of SPIE, Volume 7, pp 82-86L, doi: 10.1117/12.913491.

19. Zamanine zhad Ladan, Volker Hohmann, Andreas Buchner, Marc Ran Schaller, and Tim Jorgen, "A physiologically-inspired model reproducing the speech intelligibility benefit in cochlear implant listeners with residual acoustic hearing", Journal of Hearing Research, 2016.