A REVIEW ON SPEECH PERCEPTION IMPROVEMENT TECHNIQUE USING DICHTOTIC PRESENTATION TO SENSORINEURAL HEARING LOSS

1SAYLI PRADHAN, 2D. S. CHAUDHARI

1,2Department of Electronics and Telecommunication, Government college of Engineering, Amravati, Maharashtra, India
E-mail: 1saylipradhan@yahoo.com, 2ddscscc@yahoo.com

Abstract- Hearing loss has become very common today. Approximately 10% of the population suffering hearing impairment due to different reasons across the world. Among all most commonly observed sensorineural hearing loss caused due to poor cochlear hair cells function and damages to cochlear nerve. Sensorineural hearing loss is characterized by widening of auditory filter which leads to spectral masking and degrades speech perception. Different researches are incorporated to minimize this effect but from previous studies it is observed that splitting of signal into two different parts and providing them dichotically to both the ears reduce it to effectively. In this paper several technique are reviewed to improve speech perception along with binaural dichotic presentation of speech using Field Programmable Logic Array (FPGA) to be used in hearing aid.

Keywords- Binaural Hearing Aid, Dichotic Presentation, FPGA Implementation And Sensorineural Hearing Loss.

I. INTRODUCTION

Hearing process has been called complex because it includes capacity of both ear for detection of sound and the brain’s potential for interpretation of sound [19]. Hearing loss is classified based on the location of defect in an ear [5]. Conductive hearing loss occurs due to an abnormality in the middle ear leading to poor transmission of the sound to the inner ear. Sensorineural loss is caused by pathology in the cochlea and/or due to degeneration of the auditory nerves. Mixed hearing loss is combination of conductive and sensorineural hearing loss. Central loss occurs due to inability of the brain in decoding the neural firings into meaningful linguistic information.

Sensorineural hearing loss is widening of the auditory filter bandwidths thus the filter slope becomes steeper and the results in overlap of adjacent spectral bands called spectral masking. Due to spectral masking peaks and valleys of the speech spectrum are broadened affecting the speech perception. This leads to a decrease in frequency resolving capacity of auditory system [9]. Sensorineural loss results in elevated hearing thresholds, reduced dynamic range and intolerable loudness, and increased temporal and spectral masking, resulting in degraded speech perception [3]. Spectral contrasts, reduces it results in broadening of auditory filters. In general, most commonly found hearing loss is sensorineural loss and it progressively gets worse with time.

Hearing aids generally provide frequency-selective amplification to compensate for the elevated hearing thresholds. Automatic volume control and multichannel dynamic range compression are used to partially address the problems associated with the reduced dynamic range and loudness recruitment [3]. Increased in temporal masking, it results in poor detection of acoustic events. Increased spectral masking caused by widening of auditory filters results in poor discrimination of spectral features. The increased masking makes speech perception very difficult in the presence of noise. It also results in poor speech perception due to increased intra-speech masking. Improvement of consonant-to-vowel ratio (CVR) has been investigated for reducing the effects of increased temporal masking [12]. Techniques based on spectral contrast enhancement [15] and multiband frequency compression [14] have been used for reducing the effects of increased spectral masking. In this paper technological review is undertaken for speech perception improvement technique for sensorineural hearing loss and also explain hearing mechanism. This paper is organized into 4 sections. An introduction in first section is followed by hearing mechanism about transfer of information from auditory system to brain. Third section describes different researchers incorporated for improving speech perception. And last section provides Conclusion and future work.

II. ROLE OF BASILAR MEMBRANE IN HEARING MECHANISM

The cochlear is complex coiled structure. Cochlear region consist of long membrane called Basilar membrane. The basilar membrane inside the cochlea of the inner ear is a stiff structural element that separates two liquid-filled tubes that run along the cochlea, the scala media and the scala. Figure 1 represents schematic of Basilar membrane it shows detection of sound input varies from high frequency in base to apex detects lower frequency. There are approximately 15,000 hair cells in each human ear.
Basilar membrane acts like a series of linked resonators, much as a concatenated set of tuning forks. Each point on the basilar membrane was postulated to have a characteristic frequency at which it vibrated most efficiently; because it was physically linked to adjacent areas of the membrane, each point also vibrated (if somewhat less readily) at other frequencies, thus permitting propagation of the traveling wave. It is now clear, however, that the tuning of the auditory periphery, whether measured at the basilar membrane or recorded as the electrical activity of auditory nerve fibers, is too sharp to be explained by passive mechanics alone. At very low sound intensities, the basilar membrane vibrates much more than would be predicted by linear extrapolation from the motion measured at high intensities. Basilar membrane acts as frequency analyzer.

![Fig.1 Schematic of Basilar membrane, representing frequency variation from base to apex (logarithmic variation)](image)

### III. BINAURAL FUSION

Binaural fusion or binaural integration is a cognitive process that includes the "fusion" of different auditory information presented binaurally, or to each ear. In humans, this integration is essential integration in understanding speech because one ear may pick up more information about the speech stimuli than the other. Binaural fusion [5] is process through which brain compares sound information received by each ear and translates the difference into unified perception of sound coming from one source.

In binaural fusion, inputs from both ears integrate and fuse to create a complete auditory picture at the brainstem. Therefore, the signals sent to the central auditory nervous system are representative of this complete picture, integrated information from both ears instead of a single ear.

### IV. LITERATURE REVIEW

Researcher across the world employed speech processing methods for improving speech perception. A scheme proposed for speech discrimination in noise due to reduced frequency selectivity. Sum and difference with delayed analog signal [7] are employed in this scheme. The delay was adjusted for changing the bandwidth to 200Hz, 500Hz, and 800 Hz. These filters were designed for an efficient realization; with the magnitude responses had a large overlap between the pass bands and stop bands. Dichotic presentation did not improve word recognition scores and this was attributed to three possible factors: improper filtering, insufficient listening experience by the subjects and non-feasibility of binaural fusion of dichotic signals. Concluding there was no significant improvement observed with the scheme.

Later some of the researcher designed an 8-channel digital filter bank [6], realized using complementary interpolated linear phase filters with constant bandwidth of 700Hz, for spectral splitting. The filter gains were complementary on a linear scale and the magnitude responses had much better separation of pass and stop bands as compared to the filters used earlier by Lyregaard (1982)[7]. Dichotic presentation resulted in increase of 2 dB in signal-to-noise-ratio.

Chaudhari and Pandey investigated a scheme for splitting the speech into two complementary spectra each with nine bands for binaural dichotic presentation using psychophysical tuning curves[13]. The FIR comb filters were designed by frequency sampling method using linear optimization techniques. The investigated design was able to improve recognition scores, response time, speech quality and transmission of consonantal features as observed in the listening tests. Nonsense syllables using twelve English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in vowel-consonant-vowel (V CV) and consonant-vowel (CV) context with the vowel /a/ were used as testing sound and incorporated using antialiasing filter for removing noise components with cut-off frequency 4.8 kHz and 16-bit ADC at a rate of 10 k Samples/s. The process of filtering speech was done offline and real time. The input and processed speech signal with the said scheme was presented to two ears through DAC, antialiasing filter and power amplifier. Hearing test conducted on ten subject of age group 18-58 years. The subject had mild-to-very severe bilateral sensorineural hearing loss. The sound to be tested was presented binaurally at the individual subject’s most comfortable listening level. Listening tests were conducted with ten subjects in VCV and CV context. All the subjects have shown significant improvement in recognition scores and significant decrease in response time [2].

A study of integration of information in two different spectral regions in monaural presentation of speech signal by normal listeners [11], indicated that one of the two bands could be attenuated by up to 40 dB before intelligibility was affected. The result shows a ripple in the perceived loudness of spectral components during binaural dichotic presentation may
not adversely affect speech intelligibility. The ripple perceived is likely to adversely affect speech quality and source localization, particularly for the hearing-impaired listeners with a low dynamic range, and hence it should be minimized. The values of just noticeable difference for the level of the first formant, the level of the second formant, the overall vowel level, and irregularity in the spectrum of broadband noise are reported to be about 1.5 , 3 , 1.5 , and 1 dB, respectively.

In the other study[17], the effect of dichotic presentation on source localization, by dividing the speech spectrum into a lower and an upper band, for two inter-band crossovers at 800Hz and 1.6 kHz. Under diotic presentation(same sound presented to both ear), the normal-hearing subjects were able to localize the sounds accurately, but a large spread occurred for the hearing-impaired subjects. Under dichotic presentation, the source localization for normal-hearing subjects was poor and the hearing-impaired subjects localized the sound to the side of low frequency band. These studies indicate the need for assessing the effect of binaural dichotic presentation using the proposed set of comb filters on source localization.

Some of the researcher designed the comb filters with 256 coefficients for improving band separation and reducing the variation [16] in the loudness of spectral components with frequency. The filter magnitude responses had pass-band ripple less than 1 dB, stop-band attenuation of greater than 30 dB, transition bandwidth of 78 – 117 Hz, and gain of -4 to -6 dB at inter-band crossovers. Listening tests carried on subjects showed an improvement of 7 – 20% in consonant recognition. The improvement was mainly observed in the place feature, indicating that the processing reduced the effect of increased spectral masking. Chaudhari et al. proposed a scheme in which eighteen critical bands corresponding to auditory filters based on psychophysical tuning curves were used. In binaural dichotic presentation scheme, splitting of speech signal in real time [4] into two signals with complementary short time spectra using filters with magnitude based response on two complementary auditory filter banks with linear phase was implemented and evaluated. Filter banks corresponding to eighteen critical bands over 5kHz frequency range were used. Listening tests were conducted on subjects with mild to severe very several bilateral sensorineural hearing loss. The improvement possible with the scheme was better reception of spectral characteristics was evident as the results indicated improvement in speech quality, response time decrease, enhancement in recognition scores. Comb filters were designed having adjustable magnitude response at transition crossovers for minimizing any change in perception intensity. It also promoted pass band and increase attenuation in stop band. Listening tests were conducted on normal subjects with simulated hearing loss. The designed filter indicated better speech recognition scores and relative information transmission than the earlier filter. Therefore the designed comb filter can be used for binaural aids for persons with bilateral sensorineural hearing impairment, for reducing the effect of spectral masking [2]. The speech perception gets degraded due to increased spectral and temporal masking in persons with sensorineural hearing loss.

Abed and Neukar reported comb half-band FIR-FIR structure and conventional comb- FIR-FIR decimation filter for same specifications were implemented for comparative study [18]. The designed decimation filter of half-comb-band architecture contributed to a hardware saving of 69% as compared to the comb-FIR-FIR architecture; in addition, it reduced the power consumption by 83%, respectively. The resulting architecture was hardware efficient and consumed less power compared to conventional decimation filters. The resulting architecture was hardware efficient and consumed less power compared to conventional decimation filters.

A scheme proposed the use of three different types of bandwidths in the comb filter magnitude response: constant bandwidth filters with number of bands varying from 2 to 20, critical band based comb filters and 1/3 octave bandwidth filters [8]. In case of constant bandwidth filters, comb filters designed with up to 14 bands resulted in laterization of sounds presented. Constant bandwidth filters with 16 or more bands, critical band based filters, and 1/3 octave bandwidth based filters did not show this problem. In the presence of noise, all the three types of filters improved speech intelligibility. Improvements because of Critical Band based and 1/3 octave band based filters were similar for the same SNR, and better than constant bandwidth filters.

FPGA-based implementation of 513-coefficient filters using MatLab based offline processing [3] with sampling frequency of 10 k Sample/s was carried out for use in binaural hearing aids. Implementation using a 16- bit CODEC and 15-bit integer filter coefficients used 47, 34, and 53% of combinational functions, logic registers, and logic elements, respectively, available on FPGA kit. Binaural presentation through the headphones of stimuli used for listening tests did not show any distortion for the processed sound. The resulting magnitude responses have a close match to the offline floating-point implementation. The implementation has to be integrated with automatic gain control and filtering with the frequency selective response to compensate for increased hearing threshold. It needs to be tested on subject with sensorineural hearing loss.
V. PROPOSED METHOD

A scheme of binaural dichotic presentation is used to reduce the effect of spectral and temporal masking simultaneously for person with bilateral sensorineural hearing loss. Pair of time varying comb filter is used to split the speech signal into two for binaural dichotic presentation. The proposed method suggests the implementation of comb filter using field programmable logic array. Implementation of the comb filters using different architectures is investigated to compare the filter characteristics and the resource requirements in order to evaluate the feasibility for use in binaural hearing aids. Implementation of digital filter is carried out using delays, multiplier and adder connected in specific structure to realize comb filter.

In the proposed scheme, Data is acquired using data acquisition system. Splitting of speech carried out using even and odd Filter bank. Next step is provide digital output from FPGA to DAC (Digital to Analog Convertor) to get analog output. This analog output is applied dichotically to both the ear. In this scheme, prototype designing of ASIC(Application Specific Integrated Circuit) to reduce power and area constrains.

CONCLUSION

Several speech processing technique are discussed using dichotic presentation of speech signal. It was observed that binaural dichotic presentation of speech signal to sensorineural hearing impaired reduces spectral masking and improves speech recognition especially in terms of consonantal place feature. Also a scheme with time varying comb filter reduced temporal masking. Digital filter for use in hearing aid must provide gain at required frequency range, consumes less power, space and took least time to process signal. Earlier studies indicated that comb filter designed using Field Programmable Logic Array reduced the area and power constrains enhancing the speech perception for use in binaural hearing aid for bilateral hearing.

REFERENCES


