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

1SAYLI PRADHAN, 2D.S. CHAUDHARI
1,2Electronic and Telecommunication Engineering Department, Government college of Engineering, Amravati.
E-mail: 1saylipradhan@yahoo.com, 2ddssc@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 in corporated to minimize this effect but from previous studies it is observed that splitting of signal in to two different parts and providing them dichotically to both the ears reduced 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.

INTRODUCTION

Hearing process has been called complex because it includes capacity of both ear for detection of sound and the brain potential for interpretation of sound [19]. Hearing loss is classified based on the location of defectinarer [5]. Conductive hearing loss occurs due to an abnormality in the middle ear leading to poor transmission of the sound to the innerear. Sensorineural loss is cause due to degeneration of the auditory nerves. Mixed hearing loss is combination of conductive and sensorineural hearing loss. Central loss occurs due to failure of the brain in decoding the neural information into meaningful linguistic information.

Sensorineural hearing loss is widening of the auditory filter band widths thus the filters lop becomes gli band results in overlap of adjacent spectral bands called spectral masking. Due to spectral mas king peak sand valleys of the speech spectrum are broadened affecting the speech perception. This leads to a decrease in frequency resolving capacity of auditory system [9]. Sensorineurallloss results in elevated hearing thres holds, reduced dynamic range and intolerable loudness, increased temporal and spectral masking, results in degraded speech perception [3]. Spectral contrasts, reduces ir 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 increased hearing thresholds. The Automatic volume control and multichannel dynamic range compression are used to deal 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 researches incorporated for improving speech perception. And last section provides Conclusion and future work.

II. ROLE OF BASILAR MEMBRANE IN HEARING MECHANISM

The cochlea is complex coiled structure. Cochlea region consist of long membrane called Basilar membrane. The basilar membrane inside the cochlea of the inner ear is an 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 represents
a characteristic frequency at which it vibrated most effectively; because it is physically linked to contiguous areas of the membrane, each point also vibrated (less freely) at other frequencies, thus permitting propagation of the traveling wave. It can be seen, however, that the tuning of the auditory edge, 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.

Figure 1 represents at very low sound frequencies the basilar membrane vibrates almost completely till the apex and at much higher sound frequencies only small region near base vibrates depending on intensity of sound. Therefore, it can be said that Basilar membrane acts as frequency analyzer.

III. BINAURAL FUSION

Binaural fusion or binaural integration is a cognitive process which performs 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 the process of binaural fusion, inputs from different ears integrate and fuse to create a complete auditory picture at the brainstem. Therefore, the signals deliver to the central auditory nervous system are descriptive of this complete information, integrated information from both ears instead of a single ear.

IV. SPEECH PROCESSING TECHNIQUE REVIEW

Researchers 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. An efficient realization is possible with these filter; with the magnitude responses had a large overlap between the stop bands and pass bands. When it is experimented with dichotic presentation did not improve word recognition scores and this was attributed to three possible factors: inappropriate filtering, inadequate 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 researchers designed an 8-channel digital filter bank [6], realized using complementary interpolated linear phase filters with 700Hz of constant bandwidth, used 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 Lyregaad (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 in each with nine bands for binaural dichotic presentation using psychophysical tuning curves[13]. Frequency sampling method were used for designing a FIR comb filter design 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. Twelve Different English Nonsense consonants /p, b, t, d, k, g, m, n, s, z, f, v/ in vowel-consonant-vowel (VCV) 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 carried out using offline and in real time. The input and processed speech signal with the said scheme was presented to both the ears using 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 on different subject’s most comfortable listening level. Listening test was 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 separate spectral regions in monaural presentation of speech signal by normal listeners [11], indicated that one of the two bands attenuated by up to 40 dB before fluency was affected. The result shows a ripple in the apparent loudness of spectral components in binaural dichotic presentation[11] may not adversely affect speech intelligibility. The ripple perceived is likely to undesirably affect speech quality and source localization, particularly for the hearing-impaired of...
listeners with a low dynamic range, and therefore it should be minimized. The values of just noticeable difference for different levels like 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 observed to be about 1.5 \text{ dB}, 3 \text{ dB}, 1.5 \text{ dB}, and 1 \text{ dB}, respectively.

In the other study[17], source localization which is the effect of dichotic presentation, by dividing the speech spectrum in to a lower and an upper band, for two inter-band crossovers at 800 Hz and 1.6 kHz.Underdaptive presentation (same sound presented to both ears), the normal-hearing subjects were capable 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 confined the sound to the side of low frequency band. These studies indicate the need for assessing the effect of binaural dichotic presentation with the given 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 response shad pass-band rippleless than 1dB, stop-band at attenuation of greater than 30 dB, transition band width of 78–117Hz, and gain of 4to-6-d Batinter-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 constructed on psychophysical tuning curves were used. In binaural dichotic presentation scheme, splitting of speech signal in real time [4] in to two signals with complementary short time spectra using filters with magnitude based response on two complementary auditory filterbanks with linear phase was implemented and evaluated. Filter banks corresponding to the eighteen critical bands over 5kHz frequency range were used. Listening tests were conducted on subjects with mild to severe very severe bilateral sensorineural hearing loss. The improvement possible with the scheme was better response of spectral characteristics was marked as the results indicated improvement in speech quality, response time dropped, improvement in recognition scores. Comb filters were designed having adjustable magnitude response at transition cross overs for reducing any change in perception intensity. It also promoted pass band and increase at tenuation in stop band. Listening tests were conducted on normal subjects with simulated hearing loss. The designed filter showed better speech recognition scores and relative information transmission than the earlier filter. Hence the design of comb filter can be use for binaural aids for persons with bilateral sensorineural hearing impairment, for decreasing the effect of spectral masking[2]. The speech perception gets degraded due to improved spectral masking in persons with sensor in eural hearing loss. Abed and Neurkar 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 effective 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 band widths in the comb filter magnitude response: constant band width filters with number of bands varying from 2to20, critical band based comb filters and 1/3 octave band width filters [8]. In case of constant bandwidth filters, comb filters designed with up to 14 bands result edinlateralization of sounds presented. Constant band width filters with 16 or more bands, critical band based filters, and 1/3 octave band width 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 band width filters.

FPGA-based implementation of 513-coefficient filters using MatLab based offline processing [3] with sampling frequency of 10 kSample/s was carried out for use in binaural hearing aids. Implementation using a 16-bit CODEC Cand 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 an earmatch 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.

Filter bank is designed using FPGA (Field Programmable Logic Array) with Verilog coding. FPGA implementation of filter is beneficial for fast prototype designing of ASIC (Application Specific Integrated Circuit) to reduce power and area constrains.

CONCLUSIONS

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


★★★