



Speech Processing for Sensorineural Hearing Impairment: A Review

Shraddha D. Sharma

Electronics and telecommunication department
Government College of engineering Amravati

Devendra S. Chaudhari

Electronics and telecommunication department
Government College of engineering Amravati

Abstract— *Sensorineural hearing impairment is associated with widening of auditory filters, leading to increased spectral masking and degraded speech perception. The hearing aids, based on speech processing for binaural dichotic presentation can reduce the effect of spectral masking and can help in improving speech perception by person suffering from sensorineural hearing loss. This paper presents the techniques of binaural dichotic presentation using auditory filters with complementary frequency response.*

Keywords- *Sensorineural hearing loss, Auditory filter, Binaural dichotic presentation, Spectral masking*

I. Introduction

The characteristics of sensorineural hearing loss, which occurs due to the damage of hair cells in the cochlea or degeneration of auditory nerve fibre, are frequency dependent increase in hearing threshold, reduction in dynamic range, reduced frequency resolution masking and decrease in temporal resolution. Sensorineural hearing loss exhibits widening of auditory filters due to increased spectral masking, resulting in severe degradation of spectral envelop. Normally vowels are characterized by formant frequency cues, which are widely separated from each other, hence the perception of vowels are not much affected. However the perception of consonant is severely degraded, since it requires discrimination of sub-phonemic segments like formant transition and noise burst. Splitting of speech into two signals, such that the frequency components that are likely to get masked are separated and presented to different ears has helped in reducing the effect of spectral masking for people with bilateral sensorineural hearing impairment, with residual hearing in both ears. Since the masking takes place primarily at the peripheral level and integration of information takes place at higher level in auditory perception process.

II. Binaural Dichotic Presentation Based On Auditory Filter

The processing of speech was carried out by filter banks for separating the even and odd frequency band for dichotic presentation, the processing was carried out off line, using implemented using two digital signal processors (TI/TMS320C50) with 14-bit Analog to Digital convertor (ADC) and Digital to Analog (DAC). The filter coefficients were loaded on the RAM of DSP kit. The two DSP operated at 10 k samples/s. The DSP with two audio channels can be used as in TI/TMS320C25. For listening test, a computerised administered set up was used, since the repetitiveness and time consuming nature of the test. The test administered for both processed and unprocessed speech. The subject is already briefed about the test procedure [5]. The stimuli used, were for minimising the linguistic factors and maximising the acoustic factor. Therefore non-sense syllables were used in vowel-consonant-vowel context and for intelligibility test and perceived sound quality judgement the standard words (defined by Central Institute of Deaf Word List) were used and correct response noted. The subjects were normal hearing with simulated hearing impairment by adding broadband noise (Gaussian white noise) with varying degree of signal to noise ratios to simulate mild to severe sensorineural hearing impairment.

To improve speech perception by person suffering from bilateral hearing impairment, several studies have investigated the dichotic presentation, by spectrally splitting the speech signal using comb filters with complementary magnitude responses. Lyregaard (1981) realized pair of comb filter with fixed band width for splitting the speech signal into two complementary bands for dichotic presentation. The comb filter realized had a large overlap in magnitude response between the pass band and stop band. Lunner *et al.* used fixed bandwidth (700 Hz) 8-channel digital filter bank, using interpolated linear phase filters, for spectral splitting which had better improvement over Lyregaard (1962). Chaudhari *et al.* (1997), used 18-band comb filter with complementary magnitude response based on auditory critical band described by Zwicker [1]. The critical bands are constant at 100 Hz to the 500 Hz and 15–17% of the centre frequency to the range of 770 Hz to 15 KHz. They realized filter with 128 coefficient FIR linear phase filter, with sampling frequency 10k samples/s. The filters had pass-band ripple of 4 dB, stop band attenuation of 11dB, transition width of 78 Hz. The listening test was conducted on normal hearing subjects with simulated hearing impairment. They included Gaussian white noise band limited to the band of speech that helped for simulating aspects of sensorineural hearing impairment and also varied the signal to noise ratios for varying the severity of simulated loss. The scores of

processed speech over unprocessed speech was found 5 %, 7 %, 8.9 % and 8.9 % higher for signal to noise ratios 6, 3, 0 and -3 dB respectively but the scores generally decrease as the masking noise level increases.

Some researchers investigated to improve the comb filters based on auditory critical bands with objective to adjust the magnitude response at the transition crossovers to minimise the perceived spectral distortion, pass band ripples and increase stop band attenuation. Comb filters were designed with 256 coefficients unlike earlier, it was 128 coefficients and found that the transition width halved, it was 78 Hz in Chaudhari *et al.* (1997) and it became 38 Hz but there is no significant improvement in pass band ripples and stop band attenuation was observed. The component lying between the pass band and stop band i.e. transition region presented to both side of even and odd frequency bands, was presented to both the ears, hence perceived loudness for binaural was louder than the monaural, so the magnitude response in transition region was adjusted by loudness evaluation test to determine the difference in intensity for the same perception in binaural and monaural. The result found that perceived level matched when binaural was 4–9 dB lower than the monaural level. To adjust the perceptual balance listening test were conducted; auditory filters were designed with varying crossover between -3 dB to 9 dB. They found that between -4 dB to -6 dB changes in intensity perception was not noticeable and overall improvement of processed speech was 15.5 % over unprocessed speech [6].

Later research was carried out on time varying comb filters, since being advantageous for temporal as well as spectral masking. The comb filters were 256 coefficient linear phase filter FIR filter corresponding to auditory critical bands. At an instant of time, two comb filters had magnitude response complementary to each other. The coefficients selected in cyclic sweeping of magnitude response and each time varying comb filter were constituted of numbers of comb filter depending on the shift. The pre-calculated set of coefficients were cyclically swept with m shiftings (2, 4, 8 or 16) with time period of 20 ms. For further development Cheeran *et al.* (2002) investigated to optimize the sweep cycle of the time varying comb filter and shifts to select the pre calculated coefficients of filters i.e. 2, 4, 8 and 16 shiftings and sweep cycle of 10, 20, 40, 80 and 100 ms. The result found after listening test that 16 shiftings with sweep cycle time period of 50 ms had maximum recognition score. For low sweep of 20 a ringing sound perceived along with processed speech stimuli and for sweep cycle of 100 ms an echo was observed [7].

Further researchers investigated the four binaural dichotic processing schemes i.e. adjustable gain filtering filter, spectral splitting with perceptually balanced comb filters, and temporal splitting by trapezoidal fading and combined splitting with cyclically swept comb filter. For adjustable gain filtering, gain varied between -3 dB to +3 dB for each ear separately. Each filter was constituted of 256 coefficients linear phase filter using frequency sampling technique. For temporal splitting fading function was used for inter aural switching with 70 % duty cycle and switching of 20 ms interval. For combined splitting, the sweep cycle of 20, 40 and 80 ms with 8 and 16 shiftings were considered. The result obtained were an average 22 % and 21 % for spectral and temporal splitting respectively. Maximum improvement over unprocessed speech were obtained was 138 % for 40 ms of sweep cycle and 16 shiftings in combined splitting scheme [9]. For further improvement investigation was carried out on splitting of speech signal on basis of using different frequency bands other than critical bands, hence they made comparisons by taking three different frequency banks i.e. constant bandwidth filter bank with 100 Hz, auditory critical band and one-third octave filter bank. The one-third octave band is divided into 19 bands with 70 Hz to 5 kHz. In critical band the bandwidth is nearly constant and then 15–17 % of centre frequency at higher ends. They used 512 coefficients FIR comb filter, designed using frequency sampling technique. Filter had pass band ripples less than 1 dB and stop band attenuation of 64 dB for constant width filter band, 29 dB for one-third octave filter and 22 dB for critical band filters. The crossover gain ranged from -5 dB to -6 dB for perceptual balance as observed by Cheeran *et al.* (2002). With sampling rate of speech kept at 10 k samples/s and 16 bit quantisation. Listening tests were conducted on normal hearing subjects with simulated hearing loss. They conducted experiments, first to investigate the perceptual balance when the signal switched between two ears, for each three filters the swept sine wave with low frequency 50 Hz to 5 kHz for over 40 ms duration and high frequency from 3 kHz to 3.5 kHz for 30 s duration was used and subjects were asked to notice the change in the perceived loudness during the switching between two ears, second to study the effect of different band in understanding the speech with broadband noise. The added noise is changed with different signal to noise ratios to vary its severity. In first listening test the subject did not observe any changes in loudness, when switched between two ears and in second test small distortion was detected for constant bandwidth filter band but not in critical bands and one-third octave. With different signal to noise ratios, it was found that the intelligibility of processed speech was better than the unprocessed speech equally in case of one-third octave filter and critical band filter and relatively better than the constant bandwidth filter. Many investigations have been reported on binaural hearing impairment which has shown mixed result from no advantage to improved recognition score with signal to noise ratio advantage of 2–9 dB [10].

III. Discussion And Conclusion

The study was carried out to evaluate the effectiveness of binaural dichotic presentation using comb filters with complementary magnitude response in improving speech perception by person with sensorineural hearing loss. The comb filters in earlier studies have used linear phase filters but with different bandwidth and magnitude responses. Lyregaard (1982) and Lunner *et al.* (1993) used fixed bandwidth based filters; Lunner had better improvement over Lyregaard and found 2 dB dichotic improvements over diotic presentation. While Chaudhari *et al.* (1999) and Cheeran *et al.* (2004) used auditory critical bandwidth based filters. The filters designed also varied in gain of the transition bands, pass band ripple, stop band attenuation and transition width. Chaudhari *et al.* made significant step by considering the auditory critical band described by the Zwicker [1] and used 18 band auditory filter, the test was conducted on normal hearing subjects with simulated hearing loss. The result obtained was 15% higher recognition score over unprocessed speech. Presenting

speech to impaired listener and improving recognition score was not the main goal but presenting to the most comfort level, so the listener perceived loudness remains distortion less, so further improvement was done to adjust the crossover gain which was perceived with different loudness. A test was conducted in which it was found that for -4dB to -6 dB, the change in perceived intensity was minimised. To eliminate temporal masking as well as spectral masking, time varying comb filter were designed. In the scheme of spectral splitting sensory cells corresponding to alternate bands of the basilar membrane get periodic relaxation from stimulation was investigated. Each time varying comb filter constitute a number of comb filters depend on number of shiftings. The schemes helped in reducing spectral and temporal masking and also load on the perception process for the subject. The investigation carried out to optimize the sweep cycle of time varying comb filters by considering the different sweep cycles. It was found that for 50 ms of sweep cycle the perceptual quality was highest. For further improvement, the comb filters were optimized on the basis of number of bands and bandwidth. Three filters were designed i.e. constant bandwidth filter, critical bandwidth filter and one-third octave bandwidth filter. The test result found that the perceived loudness during transfer of tones from one ear to another was perceptually balanced for all three filters. For constant bandwidth and one-third octave filter, the subject did not perceive any distortion for the speech signals but for constant bandwidth filter, a small distortion was perceived. All the results reported shows significant improvement. The result indicates that the processing may be useful in binaural hearing aid by the listener with bilateral sensorineural hearing impairment with residual hearing.

References

- [1] E. Zwicker, "Subdivision of the audible range into Critical bands (Frequenzgruppen)," J. Acoust. Soc. Am., Vol. 33, pp. 248, 1961.
- [2] Lyregaard P.E, "Frequency selectivity & Speech intelligibility in noise", Scand Audiol Suppl.15, pp.113 – 122, 1981.
- [3] L.T. Arlinger S. & J. Hellgren, "8-channel digital filter bank for hearing aid us: Preliminary results in monaural, diotic, and dichotic modes", Scand Audiol Suppl, 38, pp. 75– 81, 1993.
- [4] D. S. Chaudhari and P. C. Pandey, "Real Time Speech Processing For Dichotic Presentation for Binaural Hearing Impairment" *proc. NCBME*, Paper II, 1998.
- [5] D. S. Chaudhari and P. C. Pandey, "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, October 1997.
- [6] A. N. Cheeran, P. C. Pandey and D. S. Jangamashetti, "Design Of Comb Filters Based on Auditory Filter Bandwidths for Binaural Dichotic Presentation for Persons with Sensorineural Hearing Impairment" *Proc. 14th Int. Nat. Conf. on Digital Signal Processing (DSP 2002)*, Santorni, Greece, July 2002.
- [7] D. S. Jangamashetti, P. C. Pandey, and A. N. Cheeran, "Time Varying Comb Filters to Reduce Spectral and Temporal Masking in Sensorineural Hearing Impairment" *Bio Vision 2001 Intl Conf. Biomed. Engg. Bangalore*, India, December 2001.
- [8] A. N. Cheeran and P. C. Pandey, "Optimizing the Sweep Cycle of Time-Varying Comb Filters for Binaural Dichotic Presentation in Sensorineural Hearing Impairment" *Proc. 14th Int. Nat. Conf. on Digital Signal Processing (DSP 2002)*, Santorni, Greece, July 2002.
- [9] A. N. Cheeran and P. C. Pandey, "Speech Processing for Hearing Aids for Moderate Bilateral Sensorineural Hearing Loss" International Conference, Cybernetics and Informatics, February 2004.
- [10] P. N. Kulkarni and P. C. Pandey, "Optimizing the Comb Filters for Spectral Splitting of Speech to Reduce the Effect of Spectral Masking" *IEEE-International Conference on Signal processing, Communications and Networking Madras Institute of Technology*, Anna University Chennai India, January 2008.

Shraddha D. Sharma is a student of M.Tech in Electronics System and Communication at Government College of Engineering, Amravati. She received her B.E in Electronics and Telecommunication Engineering from Nagpur University in 2010. Her area of research includes Speech Processing for Sensorineural Hearing Impairment.

Devendra S. Chaudhari obtained BE, ME, from Marathwada University, Aurangabad and PhD from Indian Institute of Technology, Bombay, Powai, Mumbai. He has been engaged in teaching, research for period of about 25 years and worked on DST-SERC sponsored Fast Track Project for Young Scientists. He has worked as Head Electronics and Telecommunication, Instrumentation, Electrical, Research and incharge Principal at Government Engineering Colleges. Presently he is working as Head, Department of Electronics and Telecommunication Engineering at Government College of Engineering, Amravati.