Chapter 5

MULTIBAND FIR FILTERS IN SPEECH PROCESSING SCHEME FOR SENSORINEURAL HEARING IMPAIRED

5.1 Introduction

This chapter provides an overview of the human auditory system, various types of hearing impairment and effects of sensorineural impairment on speech perception. Speech processing schemes for improving speech perception for sensorineural hearing impaired subjects are reviewed. FIR filters designed for speech processing schemes in literature are studied. Novel multiband filters are designed and implemented in a speech processing scheme to improve speech perception in sensorineural hearing impaired.

5.1.1 Human Auditory System

Fig. 5.1 shows the major components of the human ear. The sound waves received by external ear pass through ear canal and cause vibrations of the tympanic membrane or eardrum. The middle ear consists of a cavity with a delicate chain of three tiny bones, the malleus, the incus and the stapes. These bones couple the vibration of the tympanic membrane to the inner ear. The inner ear consists of a fluid filled bony spiral of two and a half turn called cochlea [41]. Fig. 5.2 shows the structure of the inner ear that is transverse section of the cochlea with its three chambers, scala vestibuli, scala media and scala tympani. At upper side, the scala media is separated from scala vestibuli by Reissner’s membrane. In turn, it is separated from scala tympani by basilar membrane and a bony shelf at lower side. The organ of corti sits on the basilar membrane and it is covered by tectorial membrane. The organ of corti contains about 30,000 hair cells. In our auditory system, there is one single row of inner hair cells and three rows of outer hair cells. The inner and outer hair cells play different roles in sound reception. Incidence of sound waves

139
Fig. 5.1. Major components of the ear.
on tympanic membrane causes it to vibrate. This energy reaches the oval window and gets transmitted to the fluid in the cochlear chambers i.e. perilymph and endolymph setting up traveling waves and causing basilar membrane to vibrate. The basilar membrane with stiffness highest near the oval window and progressively reducing along the length behaves like a dispersive transmission medium in which the traveling waves lose the high frequency energy while propagating towards apical end. Differential spatial activity takes place in response to these traveling waves of different frequencies. The highest frequency in audible range affects the region near the oval window whereas the lowest frequencies affect the far end, the helicotrema [42]. The low frequency sounds set relatively higher length of basilar membrane into vibration as compared to high frequency sounds. This might be the reason for high sensitivity of the ear to low frequency sounds. The upward and downward movement of basilar membrane results in upward-inward and downward-outward movement of reticular lamina respectively. The inward and outward bending of hair cell cause generation of electric potential difference, which stimulates the cochlear nerve endings resulting in nerve impulses. These impulses are transmitted over the cochlear nerve, a part of vestibulo-cochlear nerve to the higher processes of the brain. The information travels through the cochlear nucleus, the superior olivary complex, the inferior colliculus and the medical geniculate body ending at left and right hemisphere of the auditory cortex. The fibers in the pathways undergo considerable amount of convergence and divergence at many stages.

Hearing losses are classified as conductive loss, central loss, functional loss and sensorineural loss [2]. Conductive loss results due to defects of the ear canal, ear drum or middle ear cavity which results in less acoustic energy reaching the cochlea. It may be due to wax in the external auditory meatus, fluid, pus or infection in the middle ear. In conductive hearing loss the hearing threshold and discomfort levels get increased but
frequency selectivity does not degrade and losses are greater for frequencies below 1 KHz. Central impairment is usually caused by damage to auditory cortex, inflammation of the membrane covering the brain and spinal cord, skull trauma or congenital defect. It may result in decreased speech comprehension ability even though hearing thresholds may not increase. In central masking, the threshold of signal presented to one ear is raised due to a masking sound presented to another ear. The causes for the functional loss are psychological factors rather than physiological disorder. On the basis of hearing thresholds, the impairment can be graded into several categories like mild, moderate, severe and profound. In audiological clinics, the hearing threshold level as a function of pure tone frequency is routinely measured and plotted as audiogram. Frequency selective amplification is primary aim of most hearing aids for overcoming the problem of elevated hearing threshold.

5.1.2 Sensorineural Hearing Loss

Sensorineural loss is caused by defects in cochlea i.e. sensory and/or auditory nerve i.e. neural and is associated with decrease in frequency resolving capacity of the auditory system due to spread of spectral masking along the cochlear partition [43]. It is characterized by increase in threshold of hearing, reduction in dynamic range of hearing, degradation of temporal resolution, increase in temporal masking, degradation of frequency resolution or selectivity due to increase in spectral masking and an abnormal increase in perceived loudness with increase in sound level. The loss of frequency selectivity due to broadening of critical bands results in poor speech perception[65]. Sensorineural loss causes include congenital or hereditary factors, destruction of organ of corti, basilar membrane discontinuity, degeneration of neurons in the auditory nerve in ascending pathway, tumors, long term exposure to industrial noise, acoustic trauma, action
Fig: 5.2. Transverse section through the cochlea.
of the toxic agents, and degeneration of hair cells due to aging, interruption of blood supply to inner ear, viral infection spread from middle ear etc. In general, this loss is not medically curable and it becomes progressively worse with time. Two other phenomena associated with sensorineural loss are diplacusis and tinnitus. In case of diplacusis, a pure tone is perceived as sound with more than one pitch or as a harsh or buzzing sound. Also different pitches may be heard at the two ears for the same tone. Its causes may include mild injury to organ of corti or long term exposure to very intense sound. Tinnitus or ringing in ear is a commonly occurring auditory disorder. It may be caused by spontaneous discharge of hair cells or auditory nerves and may also be induced by exposure to intense sound. It occurs with many types of sensorineural impairment and can contribute significantly to the disruption of speech comprehension in severe cases. The cochlear loss could result due to a number of factors and typically deviated audiogram configurations indicate the various pathologies [44]. Flat audiogram is noticed due to salicylate poisoning. Damage in the organ of corti by ototoxic drug, stiffening of basilar membrane and damage to hair cells and supporting deiters cell manifest as audiogram with hearing threshold increasing with frequency. In sensorineural loss, it is quite common to find some hearing in low frequency up to 1 or 2 KHz. A very high threshold at a particular frequency between 3-6 KHz usually suggest an acoustic trauma i.e. loss of hair cell due to exposure to loud sound. The hearing threshold are relatively elevated for the frequencies in the 1-4 KHz range in the congenital sensorineural damage [45]. Various diseases like typhoid, meningitis, mumps etc indicate a peculiar pattern in audiogram e.g., moderate to severe initial bilateral sensorineural loss is seen in typhoid. Retro-cochlear impairment results due to damage of auditory nerve fibers. Its causes may include tumor and hemorrhage. Viral infection may results in primary degeneration of auditory neurons which is relatively
uncommon. Secondary degeneration involves the peripheral processes of auditory neurons constituting loss of hair cells mostly in the first half of basal turn.

5.1.3 Effects of Sensorineural Loss on Speech Perception

Speech signal generally can be described as having a dynamically varying broadband spectrum. The important acoustic characteristics include the amplitude, frequency, bandwidth of formants i.e. spectral resonances specific to the vocal tract configuration, nature of excitation i.e. voiced/unvoiced and pitch [46]. The degradation in temporal resolution and increase in temporal masking also severely affect speech perception. Proper perception of consonants requires adequate temporal resolution of subphonic segments like noise bursts and formant frequency transition. Further, in speech signals, vowels are generally more intense while the consonants carrying much of the information have lower intensities [47]. Therefore, there is a possibility of masking of consonantal segment by vowels. The loudness discomfort level is the level at which a tone becomes uncomfortably loud. The dynamic range is the difference between the loudness discomfort and hearing threshold levels. In case of sensorineural loss, the hearing threshold increases with no corresponding change in the discomfort levels and thus the dynamic range may get drastically reduced. This abnormal increase in perceived loudness with increase in sound level is known as recruitment. Increase in hearing threshold of one sound in the presence of another sound is known as ‘masking’. Intense sound has a masking effect on preceding and following weak sound known as backward and forward masking respectively and this phenomena gets very severe in case of sensorineural impairment. Most forward and backward masking occurs within 100ms either at onset or ending of [θθ3], [θ4], [θ5] masking sound. The backward masking at most extends over 20ms before the masker. Both the masking effect exists due to temporal overlap of cochlear responses. Sensorineural loss is associated with widening of the auditory filter bandwidth and the
filter slope becomes shallower. Researchers have tried to understand the characteristics of sensorineural hearing loss in terms of the role of the hair cells in the transduction mechanism and role of auditory filters for presenting the information to the higher brain processes [48]. It has been reported that basilar membrane does exhibit sharp tuning curves for tone stimuli of different frequencies. The basilar membrane vibrations at a particular location produce synchronized activity in the auditory nerve fibers innervating the corresponding hair cells. Inner hair cells act as the transducer for vibrations. The outer hair cells control the sensitivity of the inner hair cells in such a way that it is high at low levels of vibrations and progressively decreases for higher levels of vibrations. They also play a role in sharpening the tuning curves of the basilar membrane. Damage to inner hair cells reduces their transduction sensitivity for basilar membrane vibrations [49]. Hearing thresholds get increased, but the loudness growth curve and the dynamic range do not get much affected. Frequency selectivity does not get very much degraded. Damage to the outer hair cells drastically impairs the active control role played by them. The sensitivity at low sound levels get reduced resulting in recruitment i.e. loudness growth curve becomes more linear and consequently dynamic range get very much reduced. Further, the tuning curves become much broader resulting in severe spectral masking and reduces frequency selectivity. Damage to the auditory nerve fiber alters the loudness growth curve depending on the damage pattern.

5.2 Binaural Dichotic Presentation

As stated earlier, sensorineural loss is associated with widening of the auditory filter bandwidths and reduced frequency selectivity at the peripheral auditory system. The peripheral auditory system behaves like a bank of bandpass filters called auditory filters or critical bands with overlapping passbands. A tuning curve resembles the magnitude response of a band pass filter with a rounded top and sloping edges [50]. The effective
bandwidth of these tuning curves are known as critical bands. The tuning curve shapes and critical band estimates are obtained by various researchers using different type of masker and experimental technique. The shapes of auditory filters are nearly symmetric at moderate sound level and they become asymmetric at high level with shallower slope on the low frequency side.

Most of the speech processing techniques for hearing aids involve monaural listening which refers to sound presentation to one ear only whereas binaural listening involves both the ears. Binaural listening offers overall sound quality, clear speech intelligibility and more relaxed listening and it helps in source localization. Binaural listening could be "diotic" with same signals presented to both the ears or it could be "dichotic" with different signals presented to the two ears. One of the possible ways of improving speech perception degraded due to loss of frequency selectivity caused by masking at the peripheral level would be to split the speech signal on the basis of its short time spectrum into two complimentary spectra for presentation to the two ears dichotically. The main objective of spectral splitting for dichotic presentation should be to enhance the perception of spectral contrast of resonance peaks without adversely affecting the perception of features cued by amplitude and duration. Considering the perception of formants to be really important, a speech processing system can be devised such that the alternate formants are presented to different ears and therefore do not contribute to masking of each other. The splitting of speech should be done in such a way that the two adjacent strong spectral components that are likely to mask each other get presented to different ears. Several schemes are employed for splitting speech for binaural dichotic presentation. Listening tests using synthesized vowels in which first formant is presented to one ear and the second to the other was carried out by [50]. Research on binaural dichotic presentation
has also been carried out earlier by [1], [6] and [51]. Binaural hearing aids which can split the speech signal on the basis of critical bands can be helpful to hearing impaired persons.

5.2.1 Review of Speech Processing Schemes

Many investigations of the frequency resolving capabilities of the auditory system have evolved out of the concept of the critical band that originated with the auditory masking studies conducted. The subdivision into critical bands seems to be correlated very closely to the cochlear mechanics and to frequency discrimination. Studies of critical bands in hearing impaired with direct measurements provide evidence of an enlarged critical band. Zwicker's critical band model [3] has been used by many researchers.

One of the earliest reported studies of splitting speech on the basis of frequency, a scheme for dichotic presentation to improve speech intelligibility in noise due to reduced frequency selectivity was evaluated by Lunner et. al. [1]. They tested the use of an 8-channel digital filter bank in monaural, diotic and dichotic modes. The filter bank was designed to give eight parallel filtered outputs, which are added together with individually adjustable weighting factors in order to obtain a proper fit of the gain frequency response of the hearing aid as per the need of the individual hearing aid user. The scheme used comb filters with constant bandwidth of 700Hz and 40dB stopband attenuation so as to present signal containing different frequency components to the two ears. By combining alternate bands together, the filter bank was used for dichotic presentation. The filter bank was realized using complementary interpolated linear phase FIR filters so that most of the coefficients were zero valued in order to minimize the number of arithmetic computations. The scheme was implemented in real time using a digital signal processor and a computer. The signal-sampling rate was 11.6k samples/s and signal processing delay was about 4 ms. The test material consisted of a list of ten five word sentences. The processing system was later rebuilt using a digital signal processor and incorporated as part of a pocket type
hearing aid. The listening tests were conducted under three conditions i.e. dichotic presentation of odd numbered bands to the one ear and even numbered bands to the other ear and diotic presentation in which all the bands are presented to both the ears. Three subjects in age group of 39 to 69 years with bilateral and moderate sensorineural hearing loss participated. The gain of the filters was adjusted depending on the hearing loss of individual subject. The scheme achieved spectral splitting but the filters did not have the sharp transition bands. The results indicated an overall improvement in speech perception for dichotic conditions. Although use of constant bandwidth filters can be exploited for very efficient realization, a single bandwidth value may not be optimal for reducing the effect of spectral masking over the entire frequency range of speech signal. Too wide a band will not resolve the lower formants for presentation to different ears. If the bandwidth is too narrow, the pitch harmonics under the same formant peak are likely to be presented to different ears resulting in reduced perception of spectral contrast. Narrow bandwidths are also likely to smear timing related cues.

Lyregaard [50] conducted experiments with three values of comb filter bandwidths of 200 Hz, 500 Hz and 800 Hz. However improvements in the speech recognition scores for the dichotic over the diotic presentation were not statistically significant. No improvement was found in speech intelligibility indicating that the lack of improvement was possibly due to unsuitable filtering, lack of subject’s listening experience and lack of the binaural fusion of dichotic signals.

Chaudhari et. al. [6], [54] investigated a scheme for splitting the speech into two complementary spectra each with nine bands on the basis of critical bands for binaural dichotic presentation to lessen the effect of reduced auditory frequency selectivity and hence improve speech intelligibility. The reported scheme was aimed at reducing the effect of spectral masking due to loss of spectral resolution. This scheme used critical bands
corresponding to auditory filters based on psychophysical tuning curves as described by Zwicker [3]. Eighteen critical bands over 5 KHz frequency range were used. The critical bands formed the passbands of the FIR filters with linear phase response in order to maintain the timing related cues desirable to preserve the relative phases of the frequency components in speech signal. Initially, the scheme was implemented for off-line processing of digitized speech signals. The signal sampling rate was 10k samples/sec. For offline processing, the efficiency of filter realization is not critical. The listening tests were conducted under two conditions i.e. diotic presentation in which all the bands are presented to both the ears and dichotic presentation where alternate bands were presented to the left and right ear. The gains of the filter bands are adjustable as a way of partial matching of the filter response to the frequency characteristics of individual subject's hearing loss. The listening test were carried out on five normal hearing subjects with hearing loss simulated by mixing broadband noise as a masker. The listening tests in another experiment was conducted on ten subjects with bilateral sensorineural hearing loss. In offline processing, a cascade combination of nine FIR band reject filters each were used for the speech processing scheme. The scheme was implemented for real-time processing for use in a binaural hearing aid. Efficient realization for real time implementation is obtained by using two comb FIR filter with arbitrary frequency response for approximating the overall magnitude response coupled with linear phase for each channel. The filters used for real time speech processing were designed using frequency sampling method of FIR filter design. For analysis of speech signals, a spectrographic analysis package which can display the time-varying spectrum of speech was used. For ascertaining the improvement in the speech quality due to processing, a compilation of qualitative assessment by the subjects about the test stimuli under various listening conditions was carried out. The stimulus-response data was analyzed for obtaining recognition scores. The test stimuli consisted of
nonsense syllables formed with twelve English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ with vowel /a/ in vowel consonant vowel VCV and consonant vowel CV contexts. These were used for studying the reception of consonantal features of duration, frication, nasality, manner, place and voicing. A computer based setup was developed for off-line processing, real-time processing, signal acquisition, analysis and automated administration of listening tests. The scheme was able to improve speech quality, response time, recognition scores and transmission of consonantal features particularly the place feature, indicating the usefulness of the scheme for better reception of the spectral characteristics. The scheme does not adversely affect the reception of other consonantal features.

5.3 Design of Novel Sharp Transition Multiband FIR Filter

5.3.1 Introduction

This section deals with the review of filters used for the speech processing schemes. For the speech processing schemes for critical band filtering in literature [1], [6], [55] linear phase FIR filters with symmetric impulse response were used. Thus degradation in speech quality due to non-linear phase response is avoided and hence the relative phases of the frequency components in speech signal are preserved so that the timing related cues are not affected [52]. One of the widely used method for designing FIR filters is the windowing technique in which impulse response of an ideal filter is truncated to a finite duration sequence using a suitable window function. The window function should reduce the filter transition overshoot and reduce the side lobes in the stopband. Another approach for design of FIR filter is the frequency sampling technique. In this technique the desired magnitude and phase responses are specified at uniformly spaced N frequency samples for filter order N and the finite impulse response is obtained by taking the IDFT of the
nonsense syllables formed with twelve English consonants /p, b, t, d, k, g, m, n, s, z, f, v/ with vowel /a/ in vowel consonant vowel VCV and consonant vowel CV contexts. These were used for studying the reception of consonantal features of duration, frication, nasality, manner, place and voicing. A computer based setup was developed for off-line processing, real-time processing, signal acquisition, analysis and automated administration of listening tests. The scheme was able to improve speech quality, response time, recognition scores and transmission of consonantal features particularly the place feature, indicating the usefulness of the scheme for better reception of the spectral characteristics. The scheme does not adversely affect the reception of other consonantal features.

5.3 Design of Novel Sharp Transition Multiband FIR Filter

5.3.1 Introduction

This section deals with the review of filters used for the speech processing schemes. For the speech processing schemes for critical band filtering in literature [1], [6], [55] linear phase FIR filters with symmetric impulse response were used. Thus degradation in speech quality due to non-linear phase response is avoided and hence the relative phases of the frequency components in speech signal are preserved so that the timing related cues [66], [67] are not affected [52]. One of the widely used method for designing FIR filters is the windowing technique in which impulse response of an ideal filter is truncated to a finite duration sequence using a suitable window function. The window function should reduce the filter transition overshoot and reduce the side lobes in the stopband. Another approach for design of FIR filter is the frequency sampling technique. In this technique the desired magnitude and phase responses are specified at uniformly spaced N frequency samples for filter order N and the finite impulse response is obtained by taking the IDFT of the
frequency response. The frequency response is recalculated with a finer resolution and is compared with the desired response.

Chaudhari [54] for off-line processing of speech signals used two filter banks each with parallel combination of nine band pass filters with different gains. In this approach, there is a possibility of notches in the magnitude at crossover frequencies due to phase shifts in the adjacent filters, hence cascade combination of band reject filters were used. All the filters are linear phase FIR filters with 255 coefficients. The filter coefficients were obtained from the sampled magnitude response using rectangular window [53]. The FIR filters had transition bandwidths of 55 Hz with pass-band ripple of 2dB and stopband attenuation of 40 dB with each filter of 256 coefficients [54]. Rectangular window and filters of such a high order was selected in order to obtain sharp cut-off so that there is no significant overlap from the neighboring bands. For real time processing comb filter are used with passbands based on critical bands using frequency sampling FIR design method which laid emphasis on flat magnitude response in the passbands, sharp transition and linear phase response [55],[54].

In another similar binaural dichotic scheme with FIR comb filter designed using frequency sampling design method with linear optimization techniques [55], the filter transitions were kept the sharpest possible at 78 to 117 Hz for various bands which resulted in maximum pass-band ripple of 1dB and stopband attenuation of 38 dB with filters each of 256 coefficients. The signal sampling rate was 10k samples/sec.

Recently, a novel analytical method for the design of equiripple optimal comb FIR filters is presented [18]. The design algorithm provides a simple and robust tool for the evaluation of highly selective optimal equiripple comb FIR filters. Independent control is possible over number of notch bands, width of notch bands and passband attenuation. A digital cochlear filter [56],[57], which is an electronic model of cochlea was implemented
frequency response. The frequency response is recalculated with a finer resolution and is compared with the desired response.

Chaudhari [54] for off-line processing of speech signals used two filter banks each with parallel combination of nine band pass filters with different gains. In this approach, there is a possibility of notches in the magnitude at crossover frequencies due to phase shifts in the adjacent filters, hence cascade combination of band reject filters were used. All the filters are linear phase FIR filters with 255 coefficients. The filter coefficients were obtained from the sampled magnitude response using rectangular window [53]. The FIR filters had transition bandwidths of 55 Hz with pass-band ripple of 2dB and stopband attenuation of 40 dB with each filter of 256 coefficients [54]. Rectangular window and filters of such a high order was selected in order to obtain sharp cut-off so that there is no significant overlap from the neighboring bands. For real time processing comb filter are used with passbands based on critical bands using frequency sampling FIR design method which laid emphasis on flat magnitude response in the passbands, sharp transition and linear phase response [55],[54].

In another similar binaural dichotic scheme with FIR comb filter designed using frequency sampling design method with linear optimization techniques [55], the filter transitions were kept the sharpest possible at 78 to 117 Hz for various bands which resulted in maximum pass-band ripple of 1dB and stopband attenuation of 38 dB with filters each of 256 coefficients. The signal sampling rate was 10k samples/sec.

Recently, a novel analytical method for the design of equiripple optimal comb FIR filters is presented [18]. The design algorithm provides a simple and robust tool for the evaluation of highly selective optimal equiripple comb FIR filters. Independent control is possible over number of notch bands, width of notch bands and passband attenuation. A digital cochlear filter [56],[57], which is an electronic model of cochlea was implemented
and tested using IIR filters with a field programmable gate array. It gives a good fit to the cochlear data with efficiency of filter hardware usage.

For implementing the speech processing scheme as part of a binaural hearing aid, the factors like power consumption, FIR filter circuit complexity etc. are to be taken into consideration. Also a need was felt for better FIR filter design techniques to develop FIR filters with less passband ripple, good stopband attenuation and sharp transition and to vary the filter specification without the need of separate design [51]. The present work deals with the design of two FIR multiband filters with linear phase response to be employed in a speech processing scheme for splitting speech into two complementary short-time spectra signals. The filter passbands correspond to auditory filters as per Zwicker model to reduce the effect of spectral masking in sensorineural hearing impaired and hence increase speech intelligibility.

Filters having finite crossovers in magnitude response, the spectral components lying in the transition region are presented to both ears. With sharp transition band of filters, there is reduced possibility of spectral components presented to both ears at the transition region. Hence imbalance in loudness perceived at the crossovers between adjacent bands will reduce which increases speech intelligibility [55].

5.3.2 Multiband Filter Model and Design

In this section, the formulation of a linear phase, sharp transition, multiband FIR filter model with equiripple response, linear transition and variable density of ripple cycles in different regions of the filter is proposed. This design allows arbitrary center frequency and passband width without need of separate design. Filter model is such that large density of ripples is introduced in the regions where discontinuities are present to increase the sharpness of the transition and a non-ideal frequency response modeled without any abrupt discontinuities is used to reduce Gibb’s phenomenon.
The magnitude response \( H_{pm}(\omega) \) of the multiband filter model is as shown in Fig. 5.3. Index \( n \) refers to various bands which are nine in our application i.e. \( n = 1, 2, \ldots, 9 \) where \( BP_n \) is the \( n^{th} \) band pass component of the multiband filter. In the proposed multiband filter model, the various regions of the filter response are modeled using simple trigonometric functions of frequency as before.

In the stopband region \( BR_{n-1} \) preceding the bandpass section \( BP_n \) from \( \omega_{s_n-1} \leq \omega \leq \omega_{s_{2n-1}} \)

\[
H_{pm}(\omega) = \frac{\delta}{2} \cos k_{p2} (\omega - \omega_{s_{2n-1}}), \quad \omega_{s_n-1} \leq \omega \leq \omega_{s_{2n-1}}
\]  

(5.1)

where \( \omega \) is the frequency variable, \( H_{pm}(\omega) \) is the pseudo-magnitude of the bandpass filter response, \( \delta \) is passband ripple, \( \omega_{s_{n-1}} \) is the center frequency of the stopband preceding passband of bandpass response \( BP_n \), \( k_{p2} \) is the filter design parameter and \( \omega_{s_{2n-1}} \) the frequency at which second stopband region (preceding the passband) starts, \( \omega_{bn} \) is the center frequency of the bandpass section \( BP_n \) and \( \omega_{pn} \) is half passband width of \( BP_n \) and \( H_{pm}(\omega_{bn} \pm \omega_{pn}) = 1 \).

In the stopband region from \( \omega_{s_{2n-1}} \leq \omega \leq (\omega_{bn} - \omega_{zn}) \) the frequency response is

\[
H_{pm}(\omega) = \frac{\delta}{2} \cos k_{p} (\omega - \omega_{s_{2n-1}}), \quad \omega_{s_{2n-1}} \leq \omega \leq (\omega_{bn} - \omega_{zn})
\]  

(5.2)

where, \( (\omega_{bn} \pm \omega_{zn}) \) is the frequency at which magnitude response is zero, \( \omega_{bn} \) is the center frequency of passband \( BP_n \) and \( k_{p} \) is the filter design parameter.

In the transition region from \( (\omega_{bn} - \omega_{zn}) \leq \omega \leq (\omega_{bn} - \omega_{pn}) \) the frequency response is

\[
H_{pm}(\omega) = \frac{1}{(\omega_{zn} - \omega_{pn})} \left[ (\omega - (\omega_{bn} - \omega_{zn})) \right], \quad (\omega_{bn} - \omega_{zn}) \leq \omega \leq (\omega_{bn} - \omega_{pn})
\]  

(5.3)

where \( (\omega_{bn} \pm \omega_{pn}) \) are passband edge frequencies.
Fig. 5.3. Illustration of proposed Multiband filter (n\textsuperscript{th} band shown) for speech processing scheme with equiripple magnitude response, linear transition and variable density ripple cycles.
In the passband region \( (\omega_{bn} - \omega_{pn}) \leq \omega \leq (\omega_{on} - \omega_{n2}) \) the frequency response is

\[
H_{pm}(\omega) = 1 + \frac{\delta}{2} \text{sinc} \left[ \frac{\omega - (\omega_{bn} - \omega_{pn})}{p} \right], \quad (\omega_{bn} - \omega_{pn}) \leq \omega \leq (\omega_{bn} - \omega_{n2})
\]  

(5.4)

where \( (\omega_{bn} - \omega_{n2}) \) is the frequency at which second passband region starts.

In the passband region from \( (\omega_{bn} - \omega_{n2}) \leq \omega \leq (\omega_{bn} + \omega_{n2}) \) the frequency response is

\[
H_{pm}(\omega) = 1 + \frac{\delta}{2} \text{cosh} \left[ \frac{\omega - (\omega_{bn} - \omega_{n2})}{p1} \right], \quad (\omega_{bn} - \omega_{n2}) \leq \omega \leq (\omega_{bn} + \omega_{n2})
\]  

(5.5)

where \( (\omega_{bn} + \omega_{n2}) \) is the frequency at which third passband region starts and \( k_{p1} \) is the filter design parameter.

In the passband region from \( (\omega_{bn} + \omega_{n2}) \leq \omega \leq (\omega_{bn} + \omega_{pn}) \) the frequency response is

\[
H_{pm}(\omega) = 1 + \frac{\delta}{2} \text{cosh} \left[ \frac{\omega - (\omega_{bn} + \omega_{n2})}{p1} \right], \quad (\omega_{bn} + \omega_{n2}) \leq \omega \leq (\omega_{bn} + \omega_{pn})
\]  

(5.6)

In the transition region \( (\omega_{bn} + \omega_{pn}) \leq \omega \leq (\omega_{bn} + \omega_{zn}) \) from the frequency response is

\[
H_{pm}(\omega) = 1 - \frac{1}{(\omega_{zn} - \omega_{pn})} \left[ \omega - (\omega_{bn} + \omega_{pn}) \right], \quad (\omega_{bn} + \omega_{pn}) \leq \omega \leq (\omega_{bn} + \omega_{zn})
\]  

(5.7)

In the stopband region from \( (\omega_{bn} + \omega_{zn}) \leq \omega \leq \omega_{s2,n+1} \) the frequency response is

\[
H_{pm}(\omega) = -\frac{\delta}{2} \text{sinc} \left[ \frac{\omega - (\omega_{bn} + \omega_{zn})}{p} \right], \quad (\omega_{bn} + \omega_{zn}) \leq \omega \leq \omega_{s2,n+1}
\]  

(5.8)

where \( \delta \) is the stopband attenuation, \( \omega_{s2,n+1} \) the frequency at which third region of the stopband starts (succeeding the passband).

In the stopband region from \( \omega_{s2,n+1} \leq \omega \leq \omega_{s,n+1} \) the frequency response is

\[
H_{pm}(\omega) = -\frac{\delta}{2} \text{sinc} \left[ \frac{\omega - \omega_{s2,n+1}}{p3} \right], \quad \omega_{s2,n+1} \leq \omega \leq \omega_{s,n+1}
\]  

(5.9)
\( \omega_{s,n+1} \) is the center frequency of the stopband succeeding the band \( B_P \) and \( k_{p3} \) is a filter design parameter.

In the multiband filter design, in the formulation and design of \( n^{th} \) component bandpass filter section the magnitude response has three regions in the passband \( B_P \) and possesses a total of \((n+1/2)\) cycles of ripple in the entire passband. In the proposed model, the first region possesses \((m+1/4)\) cycles of ripple where \( m \) is an integer and ranges from \((\omega_{bn} - \omega_{pn})\) to \((\omega_{bn} - \omega_{n2})\). It is characterized by the filter design parameter \( k_p \). The third region possesses \((m+1/4)\) cycles of ripple and this region is in the frequency range from \((\omega_{bn} + \omega_{n2})\) to \((\omega_{bn} + \omega_{pn})\). It is characterized also by filter design parameter \( k_p \). The second region possesses \((n-2m)\) cycles of ripples (\( n \) is even for symmetry) and spans the frequency range from \((\omega_{bn} - \omega_{n2})\) to \((\omega_{bn} + \omega_{n2})\). This region is characterized by the filter design parameter \( k_{p1} \).

Therefore, for multiband design we have, \( \omega_{n2} = \frac{(n - m)}{2} \frac{2\pi}{k_{p1}} \), \( (5.10) \)

From \((5.10)\), \( k_{p1} = \frac{(n - m)}{2} \frac{2\pi}{\omega_{n2}} \), \( (5.11) \)

Also, \( \omega_{n2} = \omega_{pn} - \left( \frac{m + \frac{1}{4}}{4} \right) \frac{2\pi}{k_p} \), \( (5.12) \)

The stopband regions \( BR_{n-1} \) preceding \( B_P \) and possesses a total of \((n0+3/4)\) cycles of ripple in the frequency range of \( \omega_{s,n-1} \) to \((\omega_{bn} - \omega_{zn})\). The first region possesses \((n0-p)\) cycles of ripple where \( m \) is an integer, in the frequency range from \( \omega_{s,n-1} \) to \( \omega_{s2,n-1} \) characterized by \( k_{p2} \). The second region possesses \((p+3/4)\) cycles of ripple in the frequency range from \( \omega_{s2,n-1} \) to \((\omega_{bn} - \omega_{zn}) \) characterized by \( k_p \).
Therefore, \( (\omega_{bn} - \omega_{zn}) \cdot \omega_{s, n - 1} = \left( p + \frac{3}{4} \right) \frac{2\pi}{k_p} + (n_0 - p) \frac{2\pi}{k_{p2}} \) \hspace{1cm} (5.13)

Also, \( \omega_{s2, n - 1} = \omega_{s, n - 1} + (n_0 - p) \frac{2\pi}{k_{p2}} \) \hspace{1cm} (5.14)

using (5.13) in (5.14) we obtain,

\( \omega_{s2, n - 1} = (\omega_{bn} - \omega_{zn}) \cdot \left( p + \frac{3}{4} \right) \frac{2\pi}{k_p} \) \hspace{1cm} (5.15)

The parameter \( k_{p2} \) can be computed from (5.13) for any specified value of \( p \) and \( n_0 \). The bandpass filter has stopband region \( BR_{n+1} \) succeeding \( BP_n \) and possesses a total of \( (n2+3/4) \) cycles of ripple in the frequency range of \( (\omega_{bn} + \omega_{zn}) \) to \( \omega_{s, n + 1} \) which is split up into two regions characterized by parameters \( k_p \) and \( k_{p3} \) respectively.

The first region possesses \( (q+3/4) \) cycles of ripple from \( (\omega_{on} + \omega_{zn}) \) to \( \omega_{s2, n + 1} \) characterized by \( k_p \) and \( q \) is any integer. The second region possesses \( (n2-q) \) cycles of ripple where \( m \) is an integer from \( \omega_{s2, n + 1} \) to \( \omega_{s, n + 1} \) characterized by \( k_{p3} \).

\( \omega_{s, n + 1} - (\omega_{on} + \omega_{zn}) = \left( q + \frac{3}{4} \right) \frac{2\pi}{k_p} + (n_2 - q) \frac{2\pi}{k_{p3}} \) \hspace{1cm} (5.16)

\( k_{p3} \) can be computed from (5.16) for specified values of \( q \) and \( k_p \).

Using \( k_{p3} \) we obtain,

\( \omega_{s2, n + 1} = (\omega_{on} + \omega_{zn}) + \left( q + \frac{3}{4} \right) \frac{2\pi}{k_p} \) \hspace{1cm} (5.17)

For the multiband filter model we have from (5.3),

at \( \omega = (\omega_{bn} - \omega_{cn}) \), \( H_{pm}(\omega) = \left[ 1 - \frac{\delta_p}{2} \right] = \left[ \begin{array}{c} (\omega_{zn} - \omega_{cn}) \\ (\omega_{zn} - \omega_{pn}) \end{array} \right] \) \hspace{1cm} (5.18)
Simplifying, \( \omega_{cn} = \frac{\delta_p}{2} (\omega_{zn} - \omega_{pn}) + \omega_{pn} \) \hspace{1cm} (5.19)

From (5.3) we also have,

\[ H_p(\omega) = \frac{\delta_p}{2} \left( \frac{\omega_{zn} - \omega_{sn}}{\omega_{zn} - \omega_{pn}} \right) \] \hspace{1cm} (5.20)

Simplifying, \( \omega_{sn} = \omega_{zn} - (\delta_p / 2)(\omega_{zn} - \omega_{pn}) \)

Using (5.19) and (5.21) we obtain

\[ (\omega_{zn} - \omega_{cn}) = (\omega_{zn} - \omega_{pn}) \left( 1 - \frac{\delta_x + \delta_p}{2} \right) \] \hspace{1cm} (5.22)

### 5.3.3 Slope Equalization

For the multiband design, the slope equalization technique applied to bandpass filters of section 3.3.1 holds good because each bandpass section of the multiband filter is designed using the same technique. It can be shown using (3.39) that the slope of the frequency response at \( \omega_{bn} \pm \omega_{pn} \) all the bands of the multiband filter is

\[ \frac{k_p \delta_p}{2} \] \hspace{1cm} (5.23)

Similarly, slope in the transition regions of all component bands of the multiband filter obtained using (3.40), i.e. slope at \( \omega_{bn} \pm \omega_{zn} = \pm \left( -\frac{1}{(\omega_{zn} - \omega_{pn})} \right) \) \hspace{1cm} (5.24)

Equating the slopes at \( (\omega_{bn} + \omega_{pn}) \) and \( (\omega_{bn} + \omega_{zn}) \) using (5.23) and (5.24) we obtain,

\[ k_p = \frac{2}{\delta_p (\omega_{zn} - \omega_{pn})} \] \hspace{1cm} (5.25)

Also, substituting (5.22) in (5.25) we get,
\[ k_p = \frac{2\left[1 - \left(\frac{\delta_s + \delta_p}{2}\right)\right]}{\delta_p (\omega_{sn} - \omega_{cn})} \] 

(5.26)

Note that equalization of slopes leads to \( \delta_s = \delta_p \).

### 5.3.4 Expressions for Impulse Response Coefficients

Referring to filter design theory of section 2.4, the impulse response coefficients \( h(n) \) for the lowpass filter are obtained by evaluating the integrals below to obtain the expressions for the impulse response coefficients \( h(n) \) for the multiband filter. Thus,

\[
h(n) = \frac{1}{\pi} \int_{0}^{\pi} \left[ \frac{\omega_m - \omega_m}{\omega_{s1,s-1}} H_{pm}(\omega) \cos k\omega \, d\omega \right]
\]

(5.27)

\[
h(n) = \frac{1}{\pi} \int_{\omega_{s1,s-1}}^{\omega_{s1,s-1}} H_{pm}(\omega) \cos k\omega \, d\omega + \int_{\omega_{s1,s-1}}^{\omega_{s1,s-1} + \delta_p} H_{pm}(\omega) \cos k\omega \, d\omega + \int_{\omega_{s1,s-1} + \delta_p}^{\omega_{s1,s-1}} H_{pm}(\omega) \cos k\omega \, d\omega
\]

(5.28)
\[
\begin{align*}
&+ \int_{\omega_{bn} + \omega_{pn}}^{\omega_{bn} + \omega_{zn}} 1 - \left( \frac{1}{\Delta \omega} (\omega - (\omega_{bn} + \omega_{pn})) \right) \cos k \omega \, d\omega \\
&+ \int_{\omega_{bn} + \omega_{zn}}^{\omega_{2,n+1}} \left[ -\frac{\delta}{2} \sin k_p (\omega - (\omega_{bn} + \omega_{zn})) \right] \cos k \omega \, d\omega \\
&+ \int_{\omega_{2,n+1}}^{\omega_{2,p+1}} \left[ \frac{\delta}{2} \sin k_p (\omega - \omega_{2,n+1}) \right] \cos k \omega \, d\omega \\
\end{align*}
\]

(5.29)

where \( \Delta \omega = (\omega_{zn} - \omega_{pn}) \) and \( \delta_p = \delta \)

\[n = 0, 1, \ldots, \frac{N - 1}{2} \] for \( N \) odd,

\[n = 0, 1, 2, \ldots, \frac{N}{2} - 1 \] for \( N \) even and \( k = \left( \frac{N - 1}{2} \right) - n \).

Evaluating (5.29), the expressions obtained for the impulse response coefficients \( h(n) \) for the bandpass filter are

\[
\begin{align*}
\h(n) &= \frac{k_0}{2\pi} \left[ \sin k_{\omega_{s,2,n-1}} - \sin k_{\omega_{s,2,n-1}} \right] + \frac{k_0}{2\pi} \left[ \sin k_{\omega_{s,2,n+1}} - \sin k_{\omega_{s,2,n+1}} \right] \\
&\quad - \frac{k_0}{2\pi} \left[ 2 \cos k_{\omega_{bn}} \sin k_{\omega_{n2}} \right] \\
&\quad + \frac{2 \cos k_{\omega_{bn}} \left( \cos k_{\omega_{pn}} - \cos k_{\omega_{zn}} \right)}{\pi k^2 (\omega_{zn} - \omega_{pn})} + \frac{\delta_p \sin k_{\omega_{bn}} \left( k \cos k_{\omega_{n2}} - k_p \sin k_{\omega_{pn}} \right)}{\pi (k_p^2 - k^2)} \\
&\quad + \frac{\delta_p \left( k \sin k_{\omega_{s,2,n+1}} + 2 k_p \cos k_{\omega_{bn}} \cos k_{\omega_{zn}} - k \sin k_{\omega_{s,2,n+1}} \right)}{2\pi (k_p^2 - k^2)} \\
\end{align*}
\]

(5.30)

Eq. (5.30) is valid for \( N \) even where \( k \) is a non-integer, for \( N \) odd (5.30) is valid except for \( k = 0 \).

For \( N \) odd, \( k = 0 \),

\[
h \left( \frac{N - 1}{2} \right) = \left( \frac{\omega_{zn} + \omega_{pn}}{\pi} \right)
\]

(5.31)
5.3.5 Filter Synthesis Results and Discussions

Design Example: Two multiband linear phase, sharp transition FIR filters are designed each for the following desired specifications using the proposed multiband filter design approach.

Number of passbands: Nine

Passband-width: As per Zwicker’s model in Table A5.1.

Transition bandwidth (constant for all bands) is 35Hz

\[ i.e. \ 2 \times 35/11025 = 0.0063 \ (normalized). \]

Maximum passband ripple (for all bands) is \( \pm 0.15 \)dB

Minimum stopband attenuation (for all bands) is 40dB

Sampling Frequency is 11025 Hz

The filter was designed using MATLAB with program MML and measurements of various performance specifications of the designed filter done using MATLAB’s Signal Processing Toolbox. Results approximate the desired multiband filter specifications closely. For the two multiband filters, the eighteen desired passbands are obtained as per specifications. The passband ripple for each band does not exceed 0.15dB, stopband attenuation for each band is more than 40dB and the transition bandwidth of each band is 35 Hz (normalized 0.0063) for a sampling frequency is 11025 Hz obtained with a low filter order of 1025. For the multiband filters designed for the speech processing scheme the magnitude, phase and impulse response are shown in Fig.5.4 and Fig.5.5. Filter synthesis and design steps are given in Appendix A 6.1.8.

Multiband filters designed using our approach have linear phase, sharp transition, least passband ripple and good stopband attenuation with low arithmetic complexity to avoid distortion of processed speech fed to the impaired ears. The performance of our proposed multiband filters are superior compared to the most efficient filters used for the
same speech processing scheme. These filters [55] are comb filters with transition bandwidth of 70 Hz to 120 Hz, stopband attenuation of 38 dB and passband ripple of 1 dB with a sampling frequency is 10 KHz. The transition width, passband ripple, stopband attenuation, center frequency and passband width can be easily varied in our multiband filter design. The proposed multiband filter design yields closed form expression for impulse response coefficients.
Fig. 5.4. (a) Magnitude response of the proposed multiband filter used in speech processing scheme (b) Magnified view of the passband (dB) (c) Phase response (d) Impulse response sequence.
Fig. 5.5. (a) Magnitude response of the proposed complementary multiband filter used in speech processing scheme (b) Magnified view of the passband (c) Phase response (d) Impulse response sequence.
5.4 Conclusions

The proposed multiband filters possess large stopband attenuation which aids in better band separation, sharp transition that leads to proper separation of formants to be fed to the ears and low passband ripple that prevents deterioration of processed speech quality. All these features when combined lead to improvement in speech recognition for bilateral hearing impaired as observed in an experimental study conducted. It indicates that the processing scheme reduces the effect of spectral masking at the cochlear level.

Earlier schemes used conventional methods for design of FIR filters with filters in parallel, cascade and comb filters [55] were used. The use of the proposed multiband filters in the speech processing scheme leads to large savings in arithmetic hardware over conventional FIR filters and hence suitable for VLSI implementation in a digital hearing aid. The processed speech sample spectrographic analysis showed that the formants of most of the syllables were presented dichotically thus aided in higher speech intelligibility and verifies the usefulness of the scheme [58]. Listening tests were carried out on five sensorineural hearing impaired subjects and statements made by the subjects are recorded[59]. Subjects reported that there is an improvement in quality of speech after processing and they preferred the processed binaural dichotic presentation to unprocessed diotic presentation. For all the subjects the recognition scores for processed speech were higher than those for unprocessed speech.