Design of Computationally Efficient Digital FIR Filters and Filter Banks

Wei Ying

(M.Sc)

A THESIS SUBMITTED FOR THE DEGREE OF PH. D.

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2007

Acknowledgements

First of all, thanks to my supervisor, Dr. Lian Yong, for being a great mentor. His encouragement and advice are greatly appreciated. Without him, I would not accomplish my study successfully.

Thanks to my parents, Wei Chentang and Zhou Cuiying, my sister Wei Li and my brother in law Yin Jixing, for their love, trust and support.

Thanks to the students and staff in Signal Processing and VLSI Lab, especially to Mr. Yu Jianghong, for so many enlightening discussions, to Ms. Zheng Huanqun and Mr. Teo Seow Miang, for their technical support.

Thanks to my dear friends for their accompany and support.

Finally, thanks to the National University of Singapore for the financial support.

Contents

Summary

Finite impulse response (FIR) filters and filter banks have attractive properties that the stability can be guaranteed and linear-phase can be easily achieved. Therefore, they are popular in many applications such as communication systems, audio signal processing, biomedical instruments and so on. Unfortunately, due to the longer filter length, the cost of VLSI implementation of a FIR filter is generally higher than that of an infinite impulse response (IIR) filter which meets the same specifications. It is well known that the filter length of a FIR filter is inversely proportional to its transition bandwidth. Therefore the drawback becomes acute when the objective filter has a narrow transition band. The main purpose of this study is to develop computationally efficient techniques to design sharp FIR filters and filter banks.

The thesis consists of two parts. In the first part, computationally efficient methods are proposed to design filter banks suitable for hearing amplification. First, a 8-band non-uniformly spaced digital FIR filter bank with low complexity is proposed. It improves the matching between audiograms and the outputs of the filter bank due to the non-uniform allocation of frequency bands. The use of two half-band FIR filters as prototype filters and the combination of frequency-response masking (FRM) technique lead to significant savings in terms of number of multipliers. Then a 16-band non-uniformly spaced digital FIR filter bank with low group delay is proposed. The overall delay is significantly reduced as the result of novel filter structure which reduces the interpolation factor for the prototype filters.

In the second part of the thesis, efficient synthesis structures are proposed to design sharp filters. First, two low complexity designs based on frequency response masking technique are proposed. The first design uses a filter with non-periodical frequency response instead of an interpolated filter as the band-edge shaping filter. The multipliers of the sub-filters which synthesizes the band-ege shaping filter are shared efficiently. The second design uses two-step serial masking instead of parallel masking to mask the band-edge shaping filter and its complement. The first-step masking filter can be an interpolated finite impulse response (IFIR) filter which contributes to the reduction of the complexity. Secondly, a high speed decimation filter is proposed. It employs the polyphase structure to minimize the power consumption.

Abbreviation

ADC	analog-to-digital converter
BTE	behind the ear
CIC	cascaded integrator-comb
DAC	digital-to-analog converter
DFT	discrete fourier transform
DSP	digital signal processing
FFB	fast filter bank
FIR	finite impulse response
FRM	frequency response masking
HT	hearing threshold
IFIR	interpolated finite impulse response
IIR	infinite impulse response
LS	least square
MCL	most comfortable loudness
PFOM	power figure of merit

SFFM	single filter frequency masking
SQP	sequential quadratic programming
SRT	speech-recognition threshold
UCL	uncomfortable loudness
WLS	weighted least square

List of Figures

List of Tables

Chapter 1

Introduction

Digital Signal Processing (DSP) has become the focus of attention in new product design and technical literature for decades of years. The fields which adopt DSP include multimedia systems, communication systems, imaging processing, radar, medical and etc. Nowadays digital signal processors can be found at the heart of digital cameras, cell phones, hearing aid devices, audio and video players, satellites, and even biometric security equipment.

In DSP, the main function of digital filters is to extract the desired components or to remove the undesired components of the input signal. From an mathematical view, a digital filter computes the convolution of the sampled input and the weighting function of the filter. There are two types of digital filters, namely, finite impulse response (FIR) filter and infinite impulse response (IIR) filter. They are quite different in the structure and the way they work. The structure of a FIR filter is non-recursive while the structure of an IIR filter is recursive. IIR filters can achieve given filtering specifications using less memory and calculations than similar FIR filters. However they have poor stability. It is well known that FIR filters have some desirable features like stability, low coefficient sensitivity and linear phase response if the coefficients are symmetric. The drawback of an FIR filter is relatively high computational cost due to the involvement of large amount of multipliers. For a non-linear phase filter, the number of multipliers is equal to the length of the filter. For a linear phase filter, the number of multipliers is about half of the filter length.

The complexity of a digital FIR filter is inversely proportional to its transition bandwidth [1]. Therefore, the drawback of FIR filters becomes acute when the filters have sharp transition bands. The same problem occurs in the design of FIR filter banks. It is attractive to find ways to reduce the implementation complexity of sharp filters. Much effort has been invested into efficient implementation of sharp filters and filter banks. Section 1.1 and 1.2 give a brief review of these work.

1.1 Literature Review I - Approaches of Designing Sharp FIR Filters

Let a lowpass FIR filter be designed with the following specifications:

passband edge: ω_p

stopband edge: ω_s

maximum passband ripple: δ_p

maximum stopband ripple: δ_s

The length of the filter, L, can be estimated as [1].

$$L = \frac{-20\log_{10}\sqrt{\delta_p\delta_s} - 13}{14.6(\omega_s - \omega_p)/2\pi} + 1. \tag{1.1}$$

To implement the filter, the number of multipliers needed is about half of the length and the number of adders needed is about the same as the length. The complexity of such a filter is inversely proportional to $(\omega_s - \omega_p)$, which leads to a high computational cost if the transition band is narrow. Much effort has been invested into synthesizing sharp filters with low complexity. Interpolated finite impulse response (IFIR) method and frequency response masking (FRM) technique are the most efficient approaches developed so far.

1.1.1 Interpolated Finite Impulse Response (IFIR) Filters

One approach to reduce the complexity of sharp filters is interpolated finite impulse response (IFIR) method, proposed by Neuvo et al in 1984 [2]. The structure consists of two cascaded FIR filters, as shown in Fig. 1.1. $H(z^M)$ is the band-edge shaping filter obtained by replacing each delay element of the prototype filter $H(z)$ with M delay elements. It has a periodical frequency response with period of $2\pi/M$. $G(z)$ is the masking filter which is used to eliminate the unwanted passbands caused by the interpolation. The process is illustrated in Fig. 1.2. The transition bandwidth of the prototype filter is M times of that of the overall filter, which contributes to a reduced filter length. When the transition band of the desired filter is very narrow, the number of arithmetic operations using IFIR filter is much less than that of the direct design. The IFIR method is also suitable for designing highpass sharp filters.

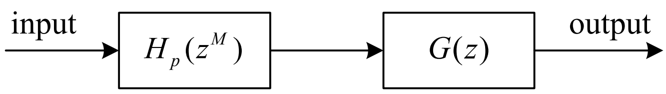


Figure 1.1: Structure of IFIR filters.

The complexity of IFIR filters can be further reduced by employing efficient algorithms in the design. A low complexity design was realized by using Remez multiple exchange algorithm iteratively to design the band-edge shaping filter and the masking filter and employing recursive running sum (RRS) method to design the masking filter [3]. In [4], an efficient family of interpolators were proposed to

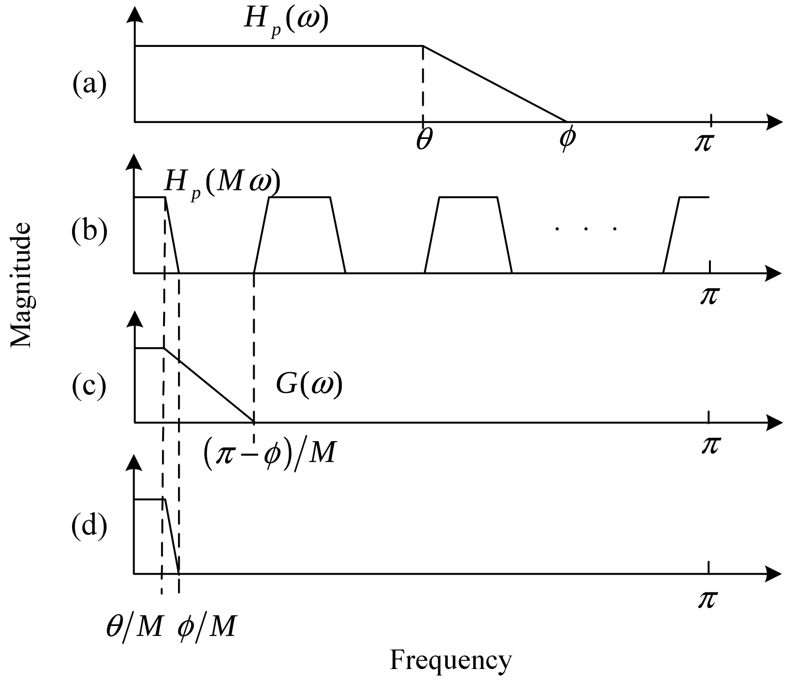


Figure 1.2: The process of synthesizing an IFIR filter.

gain further savings based on Cyclotomic Polynomials which is multiplication-free and recursion-free. In [5] IFIR filters were designed using the uniform B-spline function as an interpolator and solving the optimal Chehyshev approximation problem on the appropriate subinterval. This program nearly always provides a substantial reduction when compared to Parks-MeClellan designs. Another important development based on IFIR is the single filter frequency masking (SFFM) technique [6][7]. This approach employs several interpolated filters which come from the same prototype filter. The overall filter is obtained by cascading these filters together. SFFM results in savings in terms of the number of multipliers and adders at the cost of delay.

IFIR filters have many applications in the design of filters [8–12] and filter banks [13] [14] especially in the areas of communication and audio signal processing. However, the structure of conventional IFIR filters limits their validity to narrow-band filters. To design filters with wide passbands, a modified structure was proposed in [15] where the overall filter is decomposed into several sub-filters with less stringent constraints. Later, an extremely efficient technique, frequency response masking, was proposed by Lim [16]. Using FRM technique, sharp filters with arbitrary passbands can be designed with low complexity. The idea of FRM is using two masking filters to mask the prototype filter and its complement respectively. The outputs of the masking filters are combined to produce the desired output. The complexity is effectively reduced when the transition band is very narrow.

1.1.2 Frequency Response Masking (FRM) Technique

Instead of designing a sharp filter with transition bandwidth $\triangle B$ directly, a prototype filter with transition bandwidth $M \cdot \triangle B$ is firstly produced. The interpolation factor M is properly chosen to obtain the transition band. Then the interpolated filter and its complement are masked by the masking filters respectively. The overall filter is obtained by combining the results of masking together. This technique produces filters with very sparse coefficients which leads to very low arithmetic complexity. The structure of FRM is shown in Fig. 1.3, where $H_a(z^M)$ is the band-edge shaping filter. $H_{ma}(z)$ and $H_{mc}(z)$ are the two masking filters. N_{ma} and N_{mc}

are the length of the masking filters $H_{ma}(z)$ and $H_{mc}(z)$ respectively, which should be both even or odd. N_a is the length of $H_a(z)$, which should be odd. D_1 is the group delay of $H_a(z^M)$,

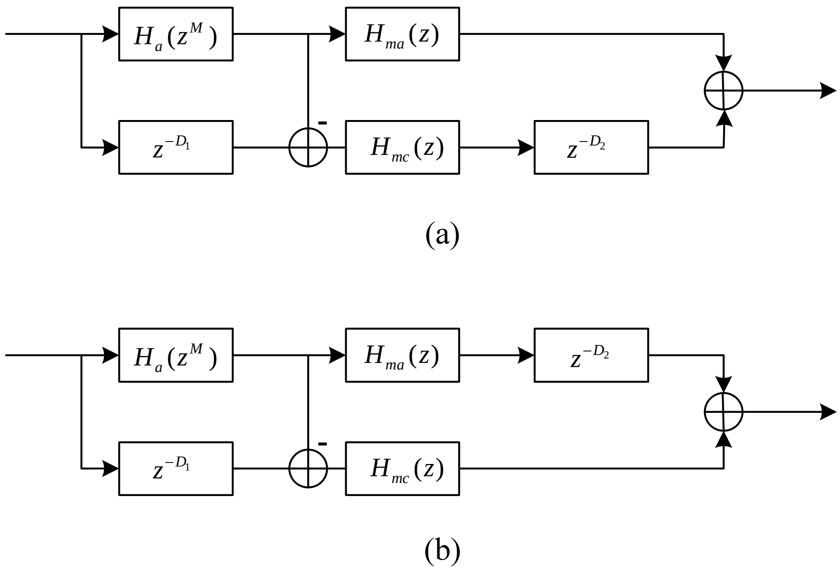


Figure 1.3: Structure of FRM filters (a) $N_{ma} \geq N_{mc}$; (b) $N_{ma} < N_{mc}$.

$$D_1 = \frac{M(N_a - 1)}{2}, \tag{1.2}$$

and D_2 is the difference of the group delay between the two masking filters.

$$D_2 = \frac{|N_{ma} - N_{mc}|}{2}. \tag{1.3}$$

The z-transform transfer function of the overall filter is given by Equation (1.4) when $N_{ma} \geq N_{mc}$ and Equation (1.5) when $N_{ma} < N_{mc}$.

$$H(z) = H_a(z^M)H_{ma}(z) + z^{-D_2}\left[z^{-D_1} - H_a(z^M)\right]H_{mc}(z), \tag{1.4}$$

$$H(z) = z^{-D_2}H_a(z^M)H_{ma}(z) + \left[z^{-D_1} - H_a(z^M)\right]H_{mc}(z). \tag{1.5}$$

After FRM technique was proposed, Lim and Lian presented their further findings in [17]. An expression for the optimal interpolation factor M was derived. It was proved that as the number of FRM stages increases, M approaches e (the base of the natural logarithm). They also proved that the complexity of a FRM filter in a K-stage design is inversely proportional to the $(K+1)^{th}$ root of the transition bandwidth. The FRM technique is effective if the normalized transition bandwidth is less than 1/16 and more efficient than the IFIR technique if the square root of the normalized transition bandwidth is less than the arithmetic mean of the normalized passband edge and stopband edge.

Much study has been carried on to obtain better performance by modifying the conventional structure. One approach is to implement the masking filters using a cascade of a common sub-filter and a pair of equalizers. Three methods based on that approach were proposed to reduce the complexity [18]. An efficient pre-filter was formed in [19], which yielded savings of 20 percent in terms of the number of multipliers compared to the original FRM approach. Furthermore, the sub-filters can be realized by IFIR technique [20][21]. Another modified structure which can achieve considerable savings introduced one more masking filter between the band-edge shaping filter and the original masking filters [22]. Additionally, new structures combining FRM and the SFFM techniques were presented in [23]. The introduced SFFM-FRM structure reduces the number of masking filters from two to one and leads to more than 35 percent savings in terms of the number of multipliers compared with the original single-stage FRM approach. In [24], the band-edge shaping filter was replaced with a Cyclotomic Polynomial pre-filters based on IFIR filters which significantly reduces the arithmetic operations.

The conventional FRM structure and the modified structures require the band-edge shaping filter to be an odd-length filter. In [25] the design of FRM filters using an even-length filter as prototype filter was presented. The optimization of the sub-filters is carried out by the Sequential Quadratic Programming (SQP) technique. It was proven that the FRM filters with even-length band-edge shaping filter leads to designs comparable to the original FRM filters.

One drawback in the synthesis techniques is that the sub-filters in the overall implementation are designed separately and iteratively. In order to improve the performance of the overall filter, Saramaki and Johanssona proposed a two-step solution which designs the sub-filters jointly. The first step uses a simple iterative algorithm to obtain a good suboptimal solution. In the second step, the suboptimal solution is used as a start-up solution for further optimization carried out by using the second algorithm of Dutta and Vidyasagar. Examples showed that the savings in terms of number of multipliers can be as much as 20 percent of the original design [27].

Besides the algorithms mentioned above, many other algorithms have been used to jointly optimize the sub-filters. One optimization technique was proposed in [29]. The algorithm uses a sequence of linear updates for the design variables. Each update is carried out by semi-definite programming. This method provides a unified design framework for a variety of FRM filters. The update can also be carried out by second-order cone programming [30]. Weighted least square (WLS) approach is also used to optimize the FRM design. The original least square (LS) problem is decomposed into two LS problems, each of that can be solved analytically. The design problem is solved iteratively [31]. In [32] the original least square problem is decomposed into four LS problems. A sequential quadratic programming (SQP) algorithm based method for FRM filters was propose in [33]. The complexity reduction results from the complementarity conditions in the SQP algorithm. This reduces the amount of computation required to update the Lagrange multipliers in a significant manner. Another efficient algorithm is the genetic algorithm (GA) applied to optimize the discrete coefficient values of the sub-filters simultaneously. It is proven that if the GA starts from the continuous solution obtained by using nonlinear joint optimization, the overall ripple of the discrete solution is very close to the continuous one [28]. Additionally, FRM technique combining with extrapolated impulse response band-edge shaping filter was introduced in [34]. An iterative optimization is used to design the sub-filters.

FRM technique has a wide spectrum of applications. It is used to synthesize filters such as diamond-shaped filters [35–37], half-band filters [38–40], IIR filters [41][42] and complex filters [43]. It also applies to the design of Hilbert transformers [44][45], noises reduction filters in ECG [46], intermediate frequency filters in CDMA and wide-band GSM modules [47] and etc. FRM technique is also efficient in designing filter banks, such as cosine-modulated filter banks [48–51], and filter banks with rational sampling factors [52][53].

1.2 Literature Review II - Filter Bank Overview

A filter bank is an array of bandpass filters. An analysis filter bank separates the input signal into several components, with each one of the sub-filters carrying a single frequency subband of the original signal. On the contrary, a synthesis filter bank combines the outputs of subbands to recover the original input signal. In most applications there are certain frequencies more important than others. Filter banks can isolate different frequency components in a signal. Therefore we can put more effort to process the more important components and put less effort to process the less important components. The subband filters can combine with down-sampling or up-sampling to form a multi-rate filter bank. The analysis and synthesis filter-array filter banks are show in Fig. 1.4, where $H_k(z), 0 \leq k \leq M-1$, is the analysis bandpass filter and $G_k(z), 0 \leq k \leq M-1$, is the synthesis bandpass filter.

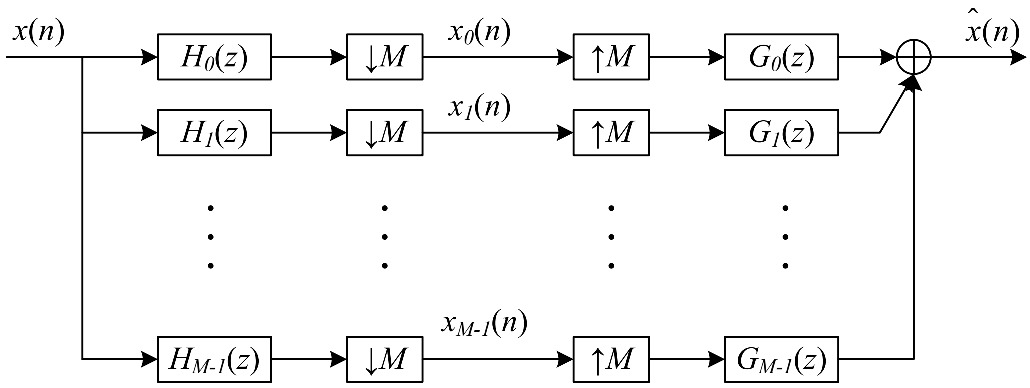


Figure 1.4: Analysis and synthesis filter bank pair.

If the frequency responses of the subbands have equal bandwidth and equal passband and stopband ripples, the filter bank is called uniform filter bank. Otherwise, it is called non-uniform filter bank. Particularly, if all the $M(M > 1)$ subband filters are derived from $H_0(z)$, such an analysis filter bank is a uniform discrete fourier transform (DFT) filter bank. $H_0(z)$ is the prototype filter. $H_k(z)$ is produced according to (1.6).

$$H_k(z) = H_0(e^{-j2\pi k/M}z), \ 0 \leq k \leq M-1. \tag{1.6}$$

If the passbands of the subbands have center frequency at $2\pi m/M$ as shown in

Fig. 1.5(a), the filter bank is even-stacked. If the passbands of the subbands have center frequency at $2\pi(m + 0.5)/M$ as shown in Fig. 1.5(b), the filter bank is odd-stacked.

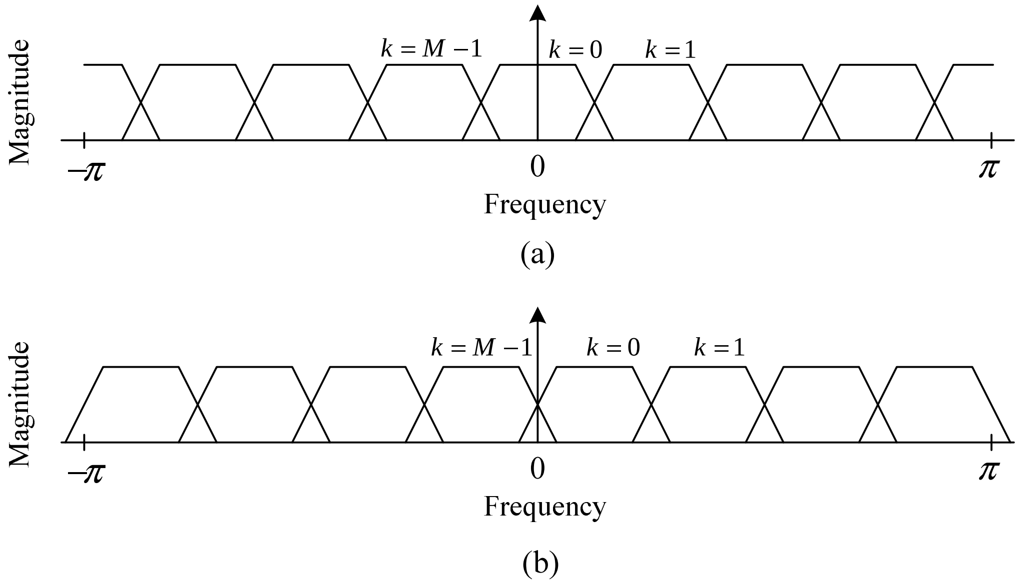


Figure 1.5: (a) Even stacked filter bank ; (b) Odd stacked filter bank.

Filter banks were originally proposed for application in speech compression [54]. They are also widely used in speech recognition [55] and speech enhancement [56]. Nowadays filter banks have extended their applications to video processing [57] [58] and image processing [59–62]. Additionally, filter banks are very useful in communication systems including digital receivers and transmitters [63], filter bank precoding for channel equalization [64–66], discrete multi-tone modulation [67] and blind channel equalization [68] [69].

Some efficient implementations of filter banks will be discussed in the following sections.

1.2.1 Polyphase Filter Bank

The invention of the polyphase representation is an important advancement in multi-rate signal processing, which results in great simplification in the implementation of filter banks as well as decimation/interpolation filters [70]. Polyphase filter banks have been widely used because of its efficiency [63] [71–73].

Considering a filter $H(z)$ and a given integer M, we can always decompose it as Type 1 polyphase

$$H(z) = \sum_{l=0}^{M-1} z^{-l} E_l(z^M), \tag{1.7}$$

where

$$E_l(z) = \sum_{n=-\infty}^{\infty} e_l(n) z^{-n}, \tag{1.8}$$

with

$$e_l(n) = h(Mn + l), 0 \leq l \leq M - 1, \tag{1.9}$$

or as Type 2 polyphase

$$H(z) = \sum_{l=0}^{M-1} z^{-M-1-l} R_l(z^M), \tag{1.10}$$

where $R_l(z)$ are permutations of $E_l(z)$

$$R_l(z) = E_{M-1-l}(z). \tag{1.11}$$

Suppose $M = 2$. If $H(z)$ is decomposed as (1.7), it is suitable to synthesize a decimation filter, shown in Fig. 1.6(a). If $H(z)$ is decomposed as (1.10), it is suitable to synthesize a interpolation filter, shown in Fig. 1.6(b).

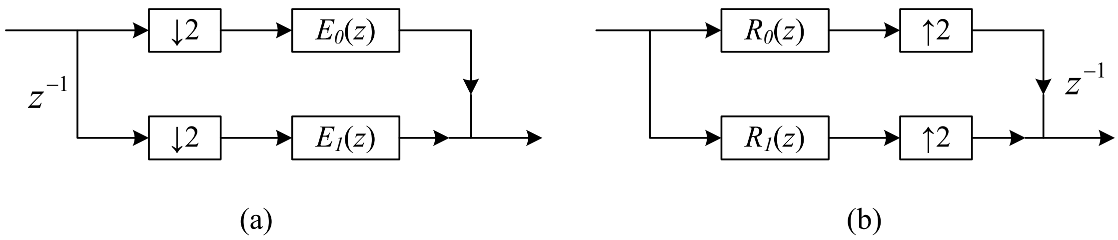


Figure 1.6: The polyphase implementation of (a) a decimation filter and (b) an interpolation filter.

Recall that a DFT filter bank is given by (1.6). With $W = e^{-j2\pi k/M}$, the k^{th} filter can now be expressed as

$$H_k(z) = H_0(zW^k) = \sum_{l=0}^{M-1} z^{-l} W^{-k} E_l(z^M), \ 0 \le k \le M-1. \tag{1.12}$$

With $X_k(z)$ denote the outputs of the DFT filter bank, $X_k(z)$ can be expressed as

$$X_k(z) = \sum_{l=0}^{M-1} W^{-kl} \left(z^{-l} E_l(z^M) X(z) \right). \tag{1.13}$$

Therefore, the polyphase uniform DFT filter bank can be illustrated in Fig. 1.7.

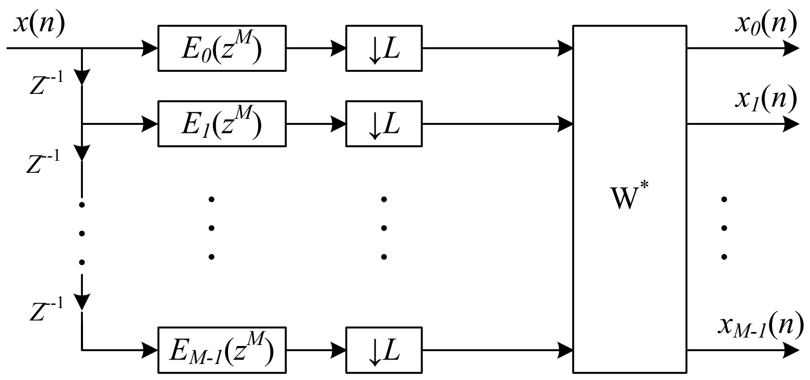


Figure 1.7: The polyphase implementation of uniform DFT filter bank.

The complexity of the polyphase filter bank is given by (1.14)

$$\Gamma_F = N + 2M \log_2(M/L), \tag{1.14}$$

where N is the length of the prototype filter.

Each of the outputs has a bandwidth approximately M times narrower than that of the original signal. It is rational to decimate the outputs by a factor of M. When $L = M$, with noble identity, the decimation block can be brought to in front of the sub-filters. Such a structure requires M times fewer multiplications and additions per unit time.

1.2.2 Fast Filter Bank

Another efficient implementation of filter banks is the fast filter bank (FFB). It was proposed in [74] and developed in [75] [76]. A fast filter bank is shown in Fig. 1.8, where $K = \log_2 M$.

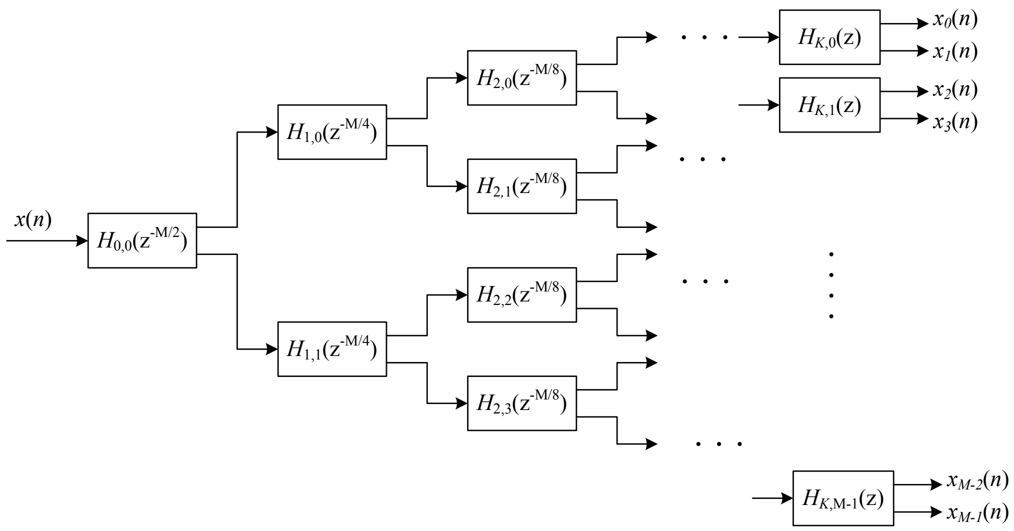


Figure 1.8: The structure of fast filter bank.

Fast filter bank is tree-structured. For a desired M-band filter bank, the filter $H_{0,0}(z)$ is first interpolated by a factor of $M/2$ to produce multiple passbands. Then the other sub-filters remove the unwanted passbands so that the outputs of FFB have a single passband for each subband. The lower branch of each sub-filter produces the complementary output.

Using $M = 4$ as an example, the frequency responses of the sub-filters and the outputs are shown in Fig. 1.9. The transition bandwidths of $H_{0,0}(\omega)$ and $H_{1,0}(\omega)$ are much wider than that of the desired subbands. This is where the efficiency of the FFB comes. The complexity of FFB can be generally expressed as

$$\Gamma_F = N_{0,0} + 4\sum_{i=1}^{K-1} 2^i N_{i,0}, \tag{1.15}$$

where $N_{i,0}$ is the minimum filter length of the sub-filters.

1.2.3 Octave Filter Bank

An Octave filter bank has a structure that at each stage, the input signal is split into two complementary parts and then decimated by 2 [77] [78]. It is also tree-structured though the basic idea is very different from the FFB. The structure of the octave filter bank is shown in Fig. 1.10.

The complexity of the sub-filters $H_{i,j}(z)$ increases as i decreases. If we want to design a filter bank with equal bandwidth, $H_{i,j}(z)$ shall be selected to be equal to $H_{0,j}(z)$. The overall complexity of the octave filter bank can be expressed as

$$\Gamma_O = \sum_{i=1}^{K-1} 2^i N_{i,0}. \tag{1.16}$$

For some applications, such as subband coding of speech and audio signals, most of the spectral energy is concentrated at the lower frequencies. In this situation,

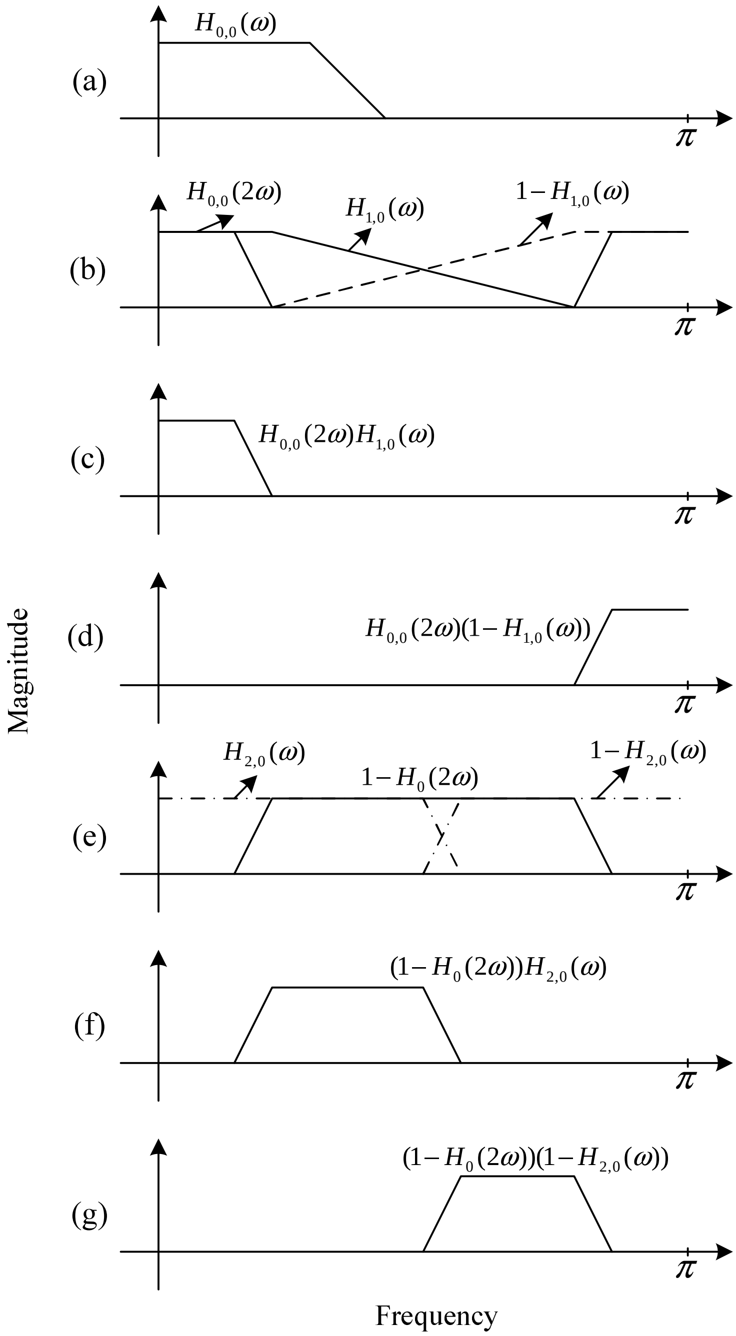


Figure 1.9: The frequency response of a fast filter bank with $M = 4$.

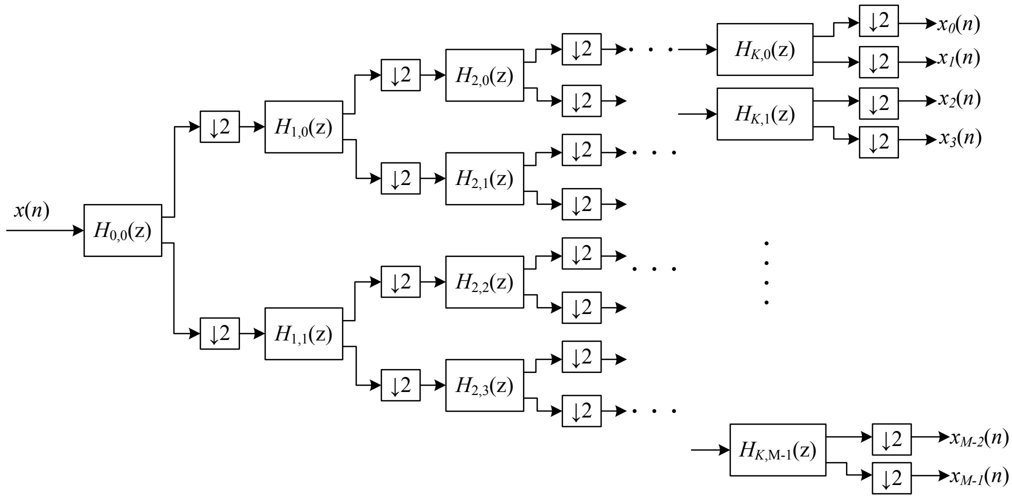


Figure 1.10: The structure of octave filter bank.

the decomposition had better meet the critical bands of human hearing. Fig. 1.11 shows a modified octave filter bank structure which can be used in audio signal processing. The lower outputs of every subbands in Fig. 1.10 are removed. The small subbands are located at the lower frequencies where human ear is more sensitive to noise, and the larger subbands are located at the higher frequencies where human ear is less sensitive to noise. By non-uniformly allocating the subbands, satisfactory performance can be obtained with less subbands compared with allocating the subbands uniformly. Further discussion about filter banks in hearing aid will be given in Chapters 2, 3 and 4.

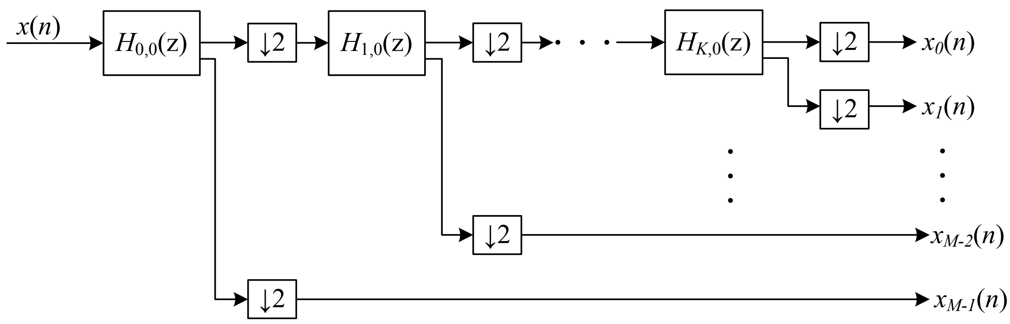


Figure 1.11: Octave filter bank with non-uniform subbands.

1.3 Research Objectives

The main purpose of this study is to design computationally efficient filters and filter banks. Non-uniform filter banks based on FRM technique and the characteristics of human hearing will be synthesized to meet the requirement of hearing amplification. Additionally, efficient synthesis structures for sharp FIR filters will be proposed.

1.4 Thesis Overview

Literature review is presented in Chapter 1. The basic concepts, the efficient implementation approaches and the applications of filters and filter banks are briefly described. The following chapters are divided into two parts. one part covers the design of filter banks for hearing aid applications. The other part covers the design of efficient sharp FIR filters.

In Chapter 2, hearing amplification is introduced. The basic concepts of audiograms, typical impairment of cochlear function, requirements of hearing aid device and modern hearing aid techniques are presented.

In Chapter 3, a non-uniform 8-band computationally efficient filter bank is proposed for hearing amplification. The background of the work is briefly introduced. Then the structure of the proposed filter bank is presented. The impacts of the transition bandwidth and the optimization for the gains of each subband are discussed. The proposed structure is verified on various audiograms.

In Chapter 4, a novel low delay non-uniform 16-band filter bank for hearing aid is proposed. The background of the work is first introduced. Then the structure and the implementation of the proposed filter bank are presented. A design example is given to illustrate the effectiveness. Estimation of the complexity and delay of the filter bank is presented in the last part of the chapter.

In Chapter 5, a low complexity design for sharp FIR filter based on FRM approach is proposed. The modified structure and the design procedure are described in detail. Implementation of the band-edge shaping filter is also discussed. Ripple analysis is done to facilitate the design. Design examples are given to test the effectiveness of the approach.

In Chapter 6, a low complexity serial masking scheme based on FRM approach is proposed. Four different structures are presented according to the implementation case and the stopband edge of the desired filter. Simulations and analysis are also presented.

In Chapter 7, a 1GHz decimation filter for sigma-delta ADC is introduced. The decimation filter is designed to minimize the power consumption.

Conclusions are given in Chapter 8.

1.5 List of Publications

[1] Ying Wei, and Yong Lian, "A computationally efficient non-uniform digital filter bank for hearing aid," *2004 IEEE International Workshop on Biomedical Circuits and Systems*, pp. S1.3.INV-17 - 20, Singapore, Dec 2- 4, 2004.

[2] Yong Lian, and Ying Wei, "A computationally efficient non-uniform digital filter bank for hearing aid," *IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 52, pp. 2754-2762, Dec. 2005.

[3] Ying Wei, and Yong Lian, "A 16-band non-uniform FIR digital filter bank for

hearing aid," *2006 IEEE International Conference on Biomedical Circuits and Systems*, London, U.K., Nov. 29 - Dec. 1, 2006.

[4] Yong Lian, Ying Wei and Chandrasekaran Rajasekaran, "A 1 GHz decimation filter for Sigma-Delta ADC," *The* 50^{th} *IEEE Int'l Midwest Symposium on Circuits and Systems*, pp. 401-404, Montreal Canada, Aug. 5 - 8, 2007.

[5] Ying Wei, and Yong Lian, "Low complexity serial masking scheme based on frequency-response masking technique," *2008 IEEE International Symposium on Circuits and Systems*, pp. 2438-2441, 18-21 May 2008, Seattle, Washington, USA.

[6] Ying Wei, and Yong Lian, "A low delay design of 16-band non-uniform FIR filter bank for hearing amplification," under preparation.

[7] Ying Wei, and Yong Lian, "Low complexity design using non-periodical band-edge shaping filter in frequency-response masking technique," under preparation.

Chapter 2

A General Introduction to Hearing Amplification

The auditory system is a sensitive and complex network which transfers sound waves to neuroelectrical signals towards the brain [79–81]. Problems in auditory system will cause hearing difficulties or deafness. According to the statistic provided by the Nemours Foundation's Center for Children's Health, about three in one thousand babies are born with hearing impairment, making it one of the most common birth defects. Hearing problem can also develop in later life because of aging or physical damages of the ears. To help the hearing-impaired people improve the quality of life, assistive technology has been developed for quite a long time.

The first hearing aid appeared at the end of 19^{th} century following the invention of telephone. These devices are based on the carbon granule microphones. Distortion and noise deteriorate the performances heavily and they are able to amplify only a narrow band of frequencies. Then hearing aids using electronic amplification appeared in 1921. The devices provide satisfactory performance for conductive hearing loss. However they are large and consume a huge amount of power. Later, the invention of transistors contributed a lot to the reduction of size and power consumption. In 1969, the first hearing aid with a directional microphone became commercially available. In 1972, non-linear compression was adopted in an integrated-circuit hearing aid. Digitally controlled analogue processing instruments, called hybrid hearing aids, were developed in the late of 1980s. Early body worn digital devices appeared in 1985. True DSP behind the ear (BTE) hearing aids became commercially available during the 1990s. Then in 1997, completely in canal (CIC) hearing aid DSP device was developed. Further advances include cochlea implant for profound hearing loss, middle-ear implant, etc. [82]

In this chapter, hearing impairment is first presented. Then audiograms which represent hearing sensitivity for different frequencies at different intensities are introduced. Next, the requirements of ideal hearing aids are clarified followed by the introduction to modern hearing aid techniques. In the last part, the necessity of using filter banks in hearing aid is discussed.

2.1 Basic Understanding of Hearing Impairment

When sound enters the cochlea, the partition vibrates, which produces travelling waves to be propagated from base to apex. As the wave propagates, the amplitude grows exponentially. After reaching the maximum amplitude at certain position along the cochlea, the wave dissipates rapidly as it continues to travel towards the apex. The position where the wave reaches its maximum amplitude is a function of frequency. Studies show that higher frequencies reach maximum amplitude closer to the base and lower frequencies reach maximum amplitude closer to the apex [83]. More damage to hearing occurs at high frequencies, near the base, where all the sound energy passes, than at low frequencies, near the apex, which is reached only by the low frequency components of the signal.

The most common hearing loss is presbycusis. Body aging process and long-term cumulative exposure to sound energy change the ear gradually. As one grows older, the ear begins losing sensitivity to sound. The ability of the ear to analyze sound and to process speech also degenerates. The first change may not be obvious. As time passes by, one begins to have difficulty following a conversation in a noisy environment. By the time these changes are manifested, it is estimated that approximately 30 % to 50 % or more of the sensory cells in the inner ear have suffered structural damage or are missing.

The main types of hearing loss are categorized as conductive, sensorineural, mixed and central [82]. The damage to outer and middle ear results in conductive hearing loss. The problem of inner ear leads to the sensorineural hearing loss. This can be further divided into sensory hearing loss due to impairment of cochlea and neural hearing loss due to impairment of auditory nerve. Mixed hearing loss is caused by the problem of both inner ear or auditory nerve and the middle or outer ears. The central hearing loss is due to lesions, disorders, or dysfunctions within the brainstem and the cortical auditory pathway. The problems, causes and its rehabilitations of the main types of hearing loss are shown in Table 2.1. The degree of hearing impairment can be classified using the hearing thresholds. Table 2.2 shows the categories of degree of hearing loss.

2.2 Audiograms

Loss of sensitivity to sound energy can be measured with a simple hearing test and represented using an audiogram. An audiogram is a graph which represents one's hearing sensitivity to different frequencies at different intensities (at different pitches and different volumes). Fig. 2.1 demonstrates different sounds and where they would be represented in an audiogram.

The horizonal axis represents pitch or frequency. The vertical axis represents loudness or intensity. Sounds with intensities less than 35dB are considered to be soft

Table 2.1: Types of hearing loss.

Type	Problem	Causes	Rehabilitation
conductive	reduction of conductive ability in the outer or middle ear.	ear infection, cerumen accumulation, otosclerosis, eardrum rupture, etc.	medical and surgical treatment, hearing aids fitting.
sensorineural	ability to hear some frequencies more than others.	damage to the cochlea or nerves of hearing. head trauma, meningitis, presbycusis, etc.	hearing aids fitting, speech therapy.
mixed	a mixture of problems in conductive and sensorineural hearing loss.	a mixture of causes in conductive and sensorineural hearing loss.	hearing aids fitting, speech therapy.
central	losing the ability of interpreting speech.	lesions, disorders, or dysfunctions within the brainstem and the cortical auditory pathway.	medical and surgical treatment, auditory processing.

Table 2.2: Degree of hearing loss.

Degree	Hearing loss range	Communication problems
mild	20-40 dB	unable to perceive some parts of daily conversations.
moderate	41-55 dB	able to perceive only 50% of daily conversations.
moderate-severe	56-70 dB	able to perceive only 30-40% of daily conversations.
severe	71-90 dB	almost unable to perceive daily conversations. distorted speech sounds. Low-pitched voice.
profound	90 dB	unable to perceive daily conversations. low speech intelligibility. low-pitched voice.

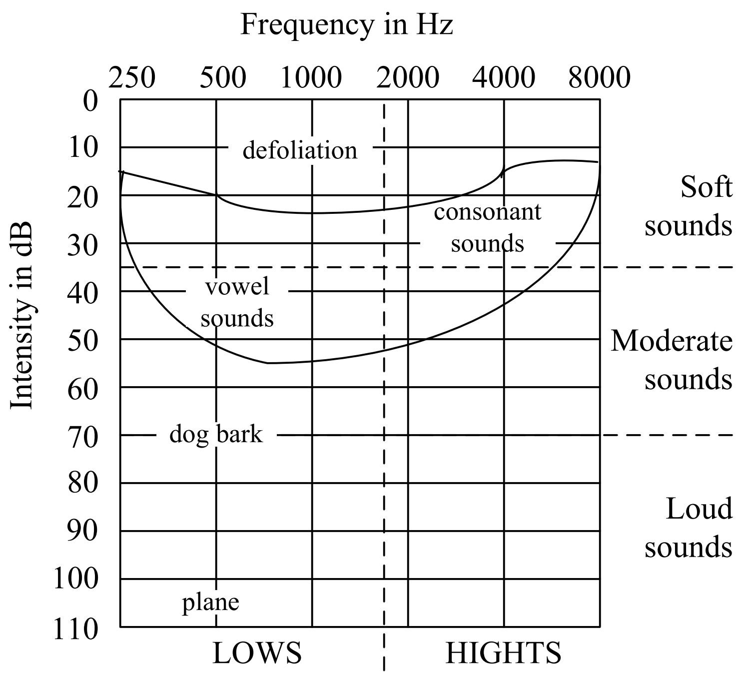


Figure 2.1: Different sounds represented in an audiogram.

sounds, such as the sound of defoliation. Sounds with intensities between 35dB and 70dB are considered to be moderate sounds, such as dogs' bark. Sounds with intensities greater than 70dB are considered to be loud sounds, such as the noise of a plane. Normal speech is either soft or moderate sounds.

Hearing thresholds are defined as the softest sounds one can hear [84]. For an adult, thresholds of 0-20 dB are considered normal. Fig. 2.2 demonstrates the normal hearing ability. In such an audiogram, the intensities where speech normally occurs are higher than the hearing thresholds ('X' represents the thresholds for the left ear and 'O' represents the thresholds for the right ear).

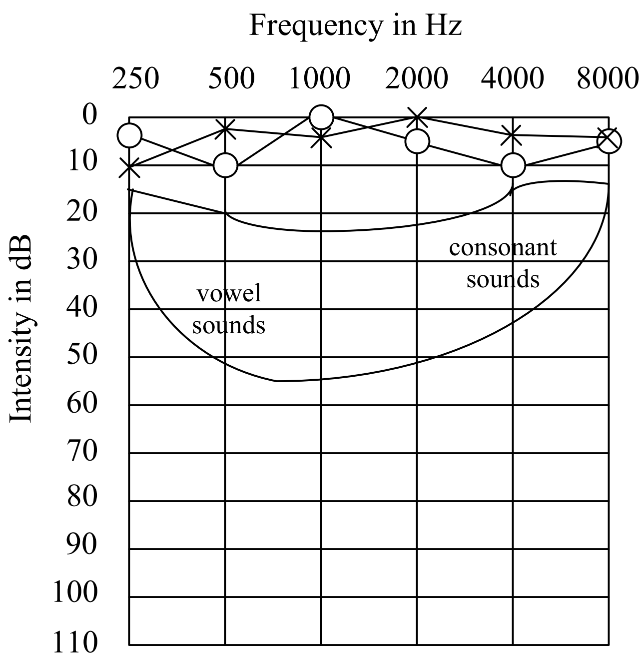


Figure 2.2: Audiogram for fine hearing.

Fig. 2.3 represents the most common hearing loss (due to aging) provided by Independent Hearing Aid Information, a public service of Hearing Alliance of America. Note that at the lower frequencies, the hearing sensitivity is much better than is needed to hear vowels. A person with such an audiogram will have no difficulty knowing that someone is talking. However, the thresholds are below the consonant areas. Because too many consonants are missed, he/she may not able to distinguish one word from another.

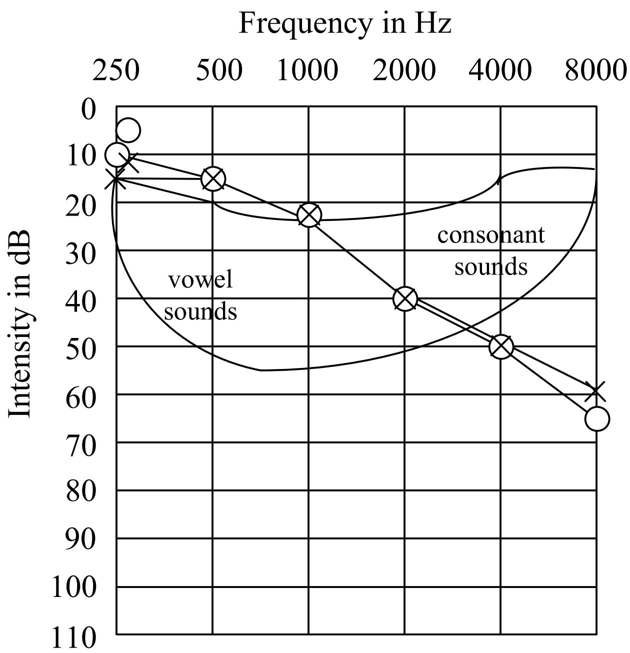


Figure 2.3: Audiogram for the most common hearing loss.

2.3 Requirements of Ideal Hearing Aids

Ideal hearing aid devices should satisfy several requirements [82], as listed below.

1. The ability of adjusting the magnitude response on arbitrary frequencies.
2. Low power consumption, given reasonable battery life.
3. Low delay and real time signal processing algorithms.
4. Small in size, light weighted, easy to wear.

2.4 Modern Hearing Aid Techniques

Modern hearing aid techniques include analog, hybrid and digital approaches. Analog hearing aids use the conventional approach [85–87]. The electronics in the hearing aids are analog. They are the least expensive category of hearing aids. However, with fixed frequency response, analog hearing aids amplify signals and background noise equally. Additionally, problems such as internal self-generated electrical noise levels, current consumption, size of battery, and user control limit the further improvement of analog hearing aid devices.

The hybrid hearing aid devices adopt digital programming during fitting while its construction is analog. Fig. 2.4 shows the structure of such a digital-programming hearing aid device. This type of devices use a computer interface to make the "analog" filtering properties controllable. This is actually different with a truly digital hearing aid. It is basically an analog device. Comparing with the pure analog hearing aids, programmable hearing aid devices provide better fitting flexibility and are able to amplify soft sounds while not to over-amplify loud sounds.

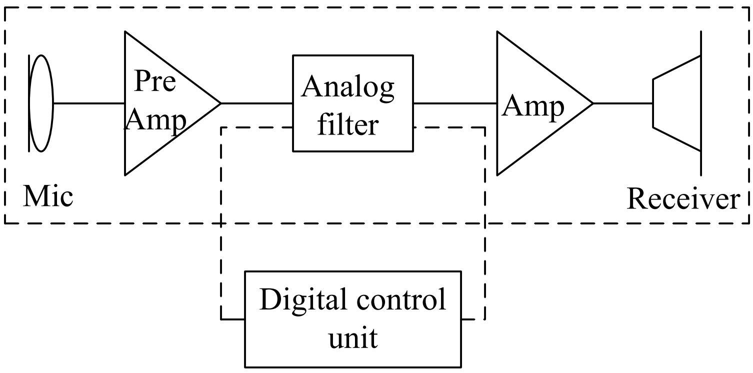


Figure 2.4: Model of digitally-programmable hearing aids.

Digital signal processing approach [88–90] uses digital filters to get arbitrary frequency responses. Linear phase is easily achieved if FIR filters are used. Additionally, noise cancellation and acoustical feedback cancellation are possible by applying digital signal processing techniques. The basic structure of a digital hearing aid is shown in Fig. 2.5. The system consists of the microphone, the analog-to-digital converter (ADC), the digital signal processor (DSP), the digital-to-analog converter (DAC), the receiver, and a memory.

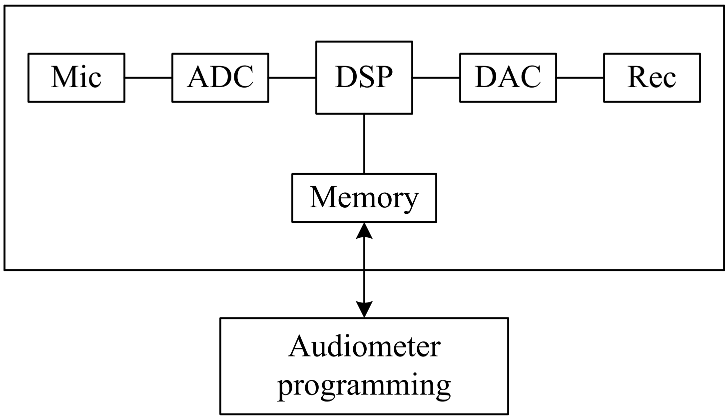


Figure 2.5: Model of digital hearing aids.

2.5 Necessity of Using Filter Banks in Digital Hearing Aid

The audiometric data used to describe the performance of a hearing aid device are hearing threshold (HT), most comfortable loudness (MCL) level and the uncomfortable loudness (UCL) level. MCL is the intensity level of speech that is most comfortably loud. For most people with normal hearing, speech is most comfortable at 40-50 dB above speech-recognition threshold (SRT). The ratio of UCL to HT is known as dynamic range, which sets a limit on the maximum output of hearing aid devices.

One way to compensate the hearing loss is to raise the speech by a wide-band gain. Fig. 2.6 shows the effect of raising a speech by 20dB gain. Before amplification, in high frequency range the speech intensities are smaller than the hearing thresholds, which means speech in this range cannot be heard. After amplifying by 20dB, all the speech can be heard. In this case the intensities of the amplified speech are smaller than the discomfort threshold.

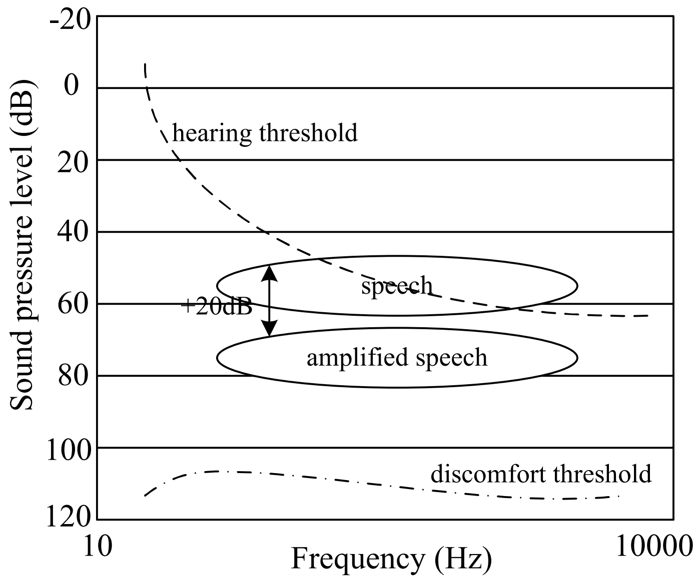


Figure 2.6: Effect of raising speech for 20dB.

However, sometimes the method of wide-band gain causes problems. In Fig. 2.7, the intensities of some speech sounds are larger than the discomfort thresholds after amplification because the gain is applied to the whole frequency. One way to solve this problem is to use filter bank based algorithm. The whole frequency range is divided into several subbands and each subband has its own amplification coefficient. The filter bank based method avoids making the low frequency sounds larger than the discomfort thresholds, as shown in Fig. 2.8.

Filter bank based algorithm permits an easy adjustment of speech amplification. Within the considered speech spectrum, the adjustment is fully programmable and is able to suit patients' comfort [91] [92]. A schematic diagram for digital signal processing method using filter bank is shown in Fig. 2.9.

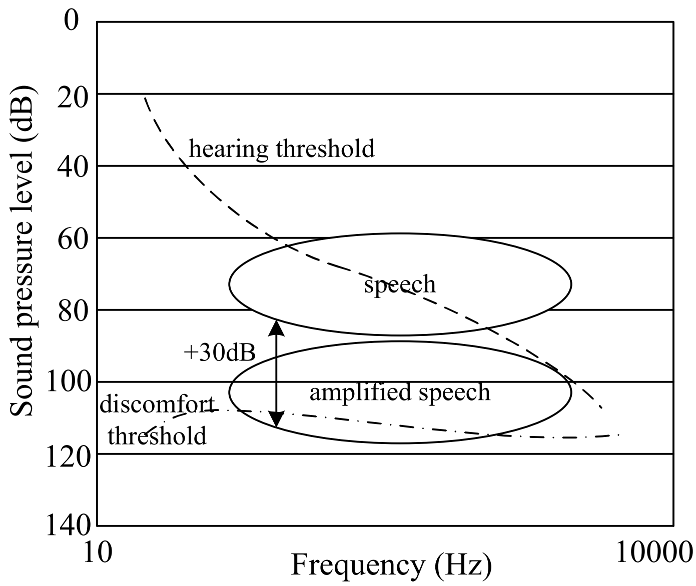


Figure 2.7: Problem caused by wide band gain method.

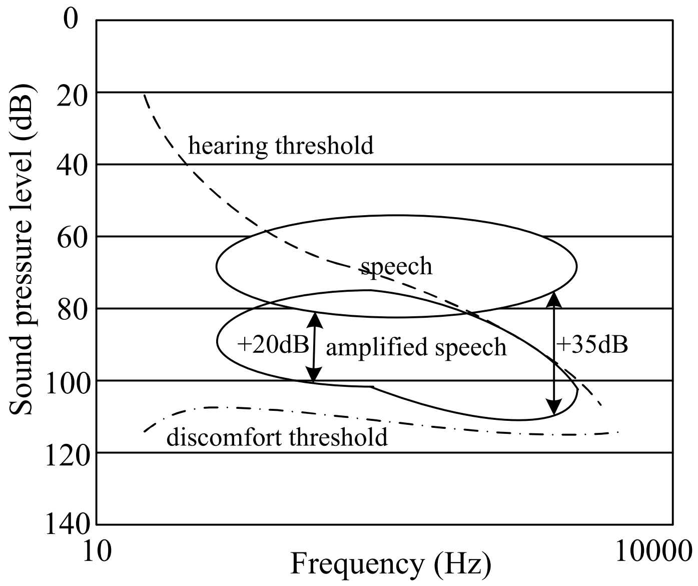


Figure 2.8: Using different gains for different bands.

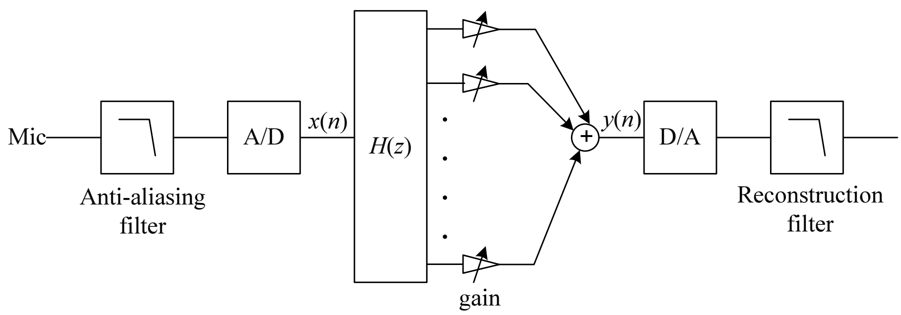


Figure 2.9: Schematic diagram for digital signal processing method.

Chapter 3

An 8-band Non-uniform Computationally Efficient Filter Bank for Hearing Aid

3.1 Introduction

Hearing thresholds which are the softest sounds one can hear are usually represented by a typical pure tone audiogram. The audiogram for normal hearing is shown in Fig. 2.2. The two marked curves demonstrate the fine hearing ability of the left and right ears where the hearing thresholds are lower than the intensities where speech normally occurs. A person with impaired hearing tends to have a low sensitivity towards certain frequencies, shown in Fig. 2.3, where hearing thresholds for high frequency sounds are higher than the normal speech intensities. The main task of hearing aid is to selectively amplify the audio sounds such that the processed sound matches one's audiogram [93–96]. To achieve this goal, ideal hearing aid should be able to adjust sound levels at arbitrary frequencies within a given spectrum. In practice, this is achieved by passing the input signal through a filter bank that divides them into different frequency bands. The gains for each subband are adjustable to suit the needs of hearing impaired people, i.e. the amplitude response of filter bank should equalize or 'match' the audiogram.

Two types of techniques are commonly used to implement the filter bank in hearing aid devices, i.e. analog and digital techniques. The analog approach manipulates the audio signal in a continuous fashion using the analog circuits. The digital method converts the sound signals into digital domain and deals with them using digital signal processing techniques. Such an approach provides the programmable capabilities and is very flexible towards the needs of hearing impaired people.

Much effort has been invested into the design of uniform digital filter banks for hearing aid applications [91] [97–100]. However, hearing level measurements are done at each octave: 250 / 500 / 1k / 2k / 4k / 8k in a standard audiogram, which suggests that the uniform filter banks may face difficulties matching the audiogram in all frequencies. It is worth noting that the typical hearing loss, especially for the cases caused by aging, occurs at high frequencies. To achieve a better compensation, narrower bands need to be allocated at high frequencies. Therefore a non-uniform spaced digital filter bank becomes very attractive.

Both FIR filters and IIR filters are widely used in audio applications. Audition is more sensitive to amplitude than to phase. However, when the outputs of sub-filters in a filter bank are summed up, a phase distortion in each of the sub-filters leads to an amplitude distortion in the output. Therefore in this paper, a non-uniform FIR filter bank is proposed to meet the hearing requirement. The filter bank is based on frequency-response masking technique and provides better matches at both low and high frequencies.

This chapter is organized as follows. In Section 3.2, details on how to form an 8-band non-uniform filter bank using two half-band prototype filters are presented. In Section 3.3, the relationship among transition bandwidth, filter complexity and matching errors is discussed. Design examples are given in Section 3.4. The optimization of gains for each subband is covered in Section 3.5. The effectiveness of the proposed filter bank is evaluated in Section 3.6. Conclusion is drawn in Section 3.7.

3.2 Structure of Proposed Filter Bank

It is well known facts that an FIR filter is always stable and possesses a linear phase response if its coefficients are symmetric. Such properties are most welcome by the hearing aid devices due to the requirement of arbitrary magnitude adjustment in the different frequency bands. The drawback of an FIR filter is relatively high computational cost due to the involvement of large amount of multipliers. In order to reduce the filter complexity, the proposed non-uniform FIR filter bank uses two simple half-band filters, $H(z)$ and $F_m(z)$, as prototypes and the subbands are designed with symmetry at the mid frequency point, as shown in Fig. 3.1 and Fig. 3.2, respectively.

Note in Fig. 3.1 the right most filters in each branch provide a pair of outputs, e.g. $P_1(z)$ and $P_8(z)$, $P_2(z)$ and $P_7(z)$, etc. These are complementary output pairs as illustrated in Fig. 3.3. $F_{mc}(z)$ is the complement of an original filter $F_m(z)$ and is formed by (3.1). Where N_F is the filter length of $F_m(z)$. The hardware cost for producing the complementary output is minimized because the required delays can be obtained from the original filter as shown in Fig. 3.3.

$$F_{mc}(z) = z^{-(N_F-1)/2} - F_m(z). \tag{3.1}$$

The outputs of the subbands are termed as $B_i(z)$, $i = 1, \cdots, 8$, as shown in Fig. 3.1, where $B_1(z)$ to $B_4(z)$ are formed by outputs $P_1(z)$ to $P_4(z)$ from the original filter

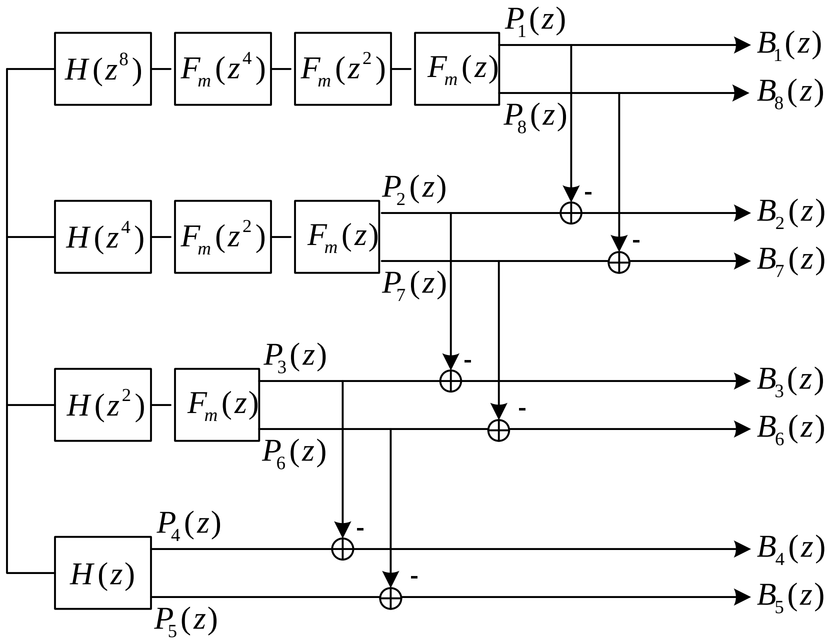


Figure 3.1: Structure of the proposed non-uniform filter bank.

$F_m(z)$, and $B_5(z)$ to $B_8(z)$ are based on the complementary outputs $P_5(z)$ to $P_8(z)$ from $F_{mc}(z)$. In order to achieve the desired frequency response and avoid frequency dependant delay [101], Leading delays should be added to each branch except for the top one to ensure that all branches have the same phase shift. To understand how subbands $B_1(z)$ to $B_8(z)$ are created, the formation of two mid bands, $B_4(z)$ and $B_5(z)$, will be illustrated. First, a half-band prototype filter $H(z)$ is formed. Its output is defiend as $P_4(z)$ and its complement output is defined as $P_5(z)$ as shown in Fig. 3.2(a). Next, two passbands at low and high frequencies are produced using $H(z^2)$, i.e. interpolating $H(z)$ by a factor of 2 as shown in Fig. 3.2(b). A masking filter $F_m(z)$ is designed in such a way that it is able to remove the passband at high frequency , as shown in Fig. 3.2(b). $P_3(z)$ is produced when $F_m(z)$ is cascaded with

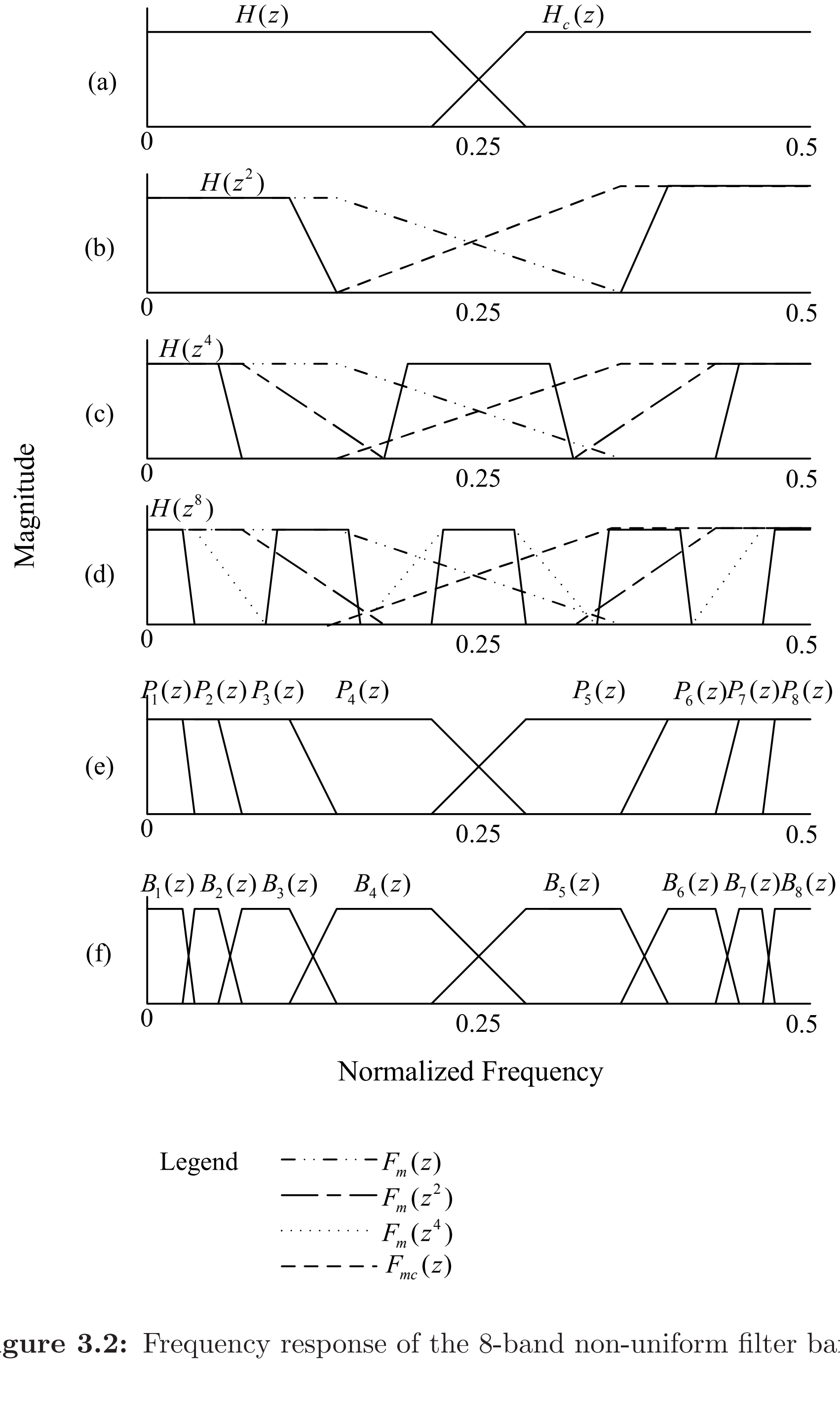


Figure 3.2: Frequency response of the 8-band non-uniform filter bank.

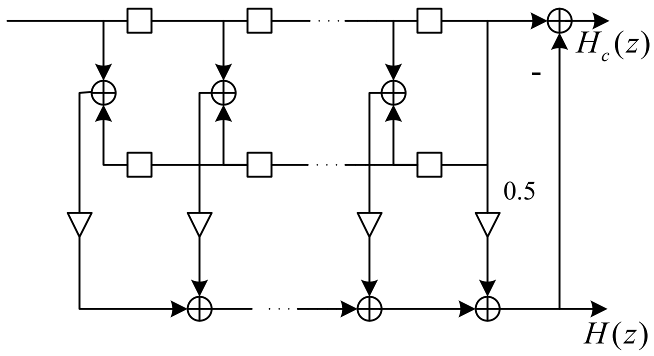


Figure 3.3: A pair of complementary filters with delay sharing.

$H(z^2)$. Similarly, $P_6(z)$ is a result of connecting $H(z^2)$ with $F_{mc}(z)$, the complement of $F_{mc}(z)$. The mid bands $B_4(z)$ and $B_5(z)$ are formed by subtracting $P_3(z)$ from $P_4(z)$ and $P_6(z)$ from $P_5(z)$, respectively. The z-transform transfer function for the lower 4 bands can be written as

$$B_i(z) = \begin{cases} P_i(z) & i = 1 \\ P_i(z) - P_{i-1}(z) & i = 2, 3, 4, \end{cases} \tag{3.2}$$

where $P_i(z)$ is defined as

$$P_i(z) = \begin{cases} H(z^{2^{4-i}}) \prod_{k=1}^{4-i} F_m(z^{2^{k-1}}) & i = 1, 2, 3 \\ H(z) & i = 4. \end{cases} \tag{3.3}$$

For the higher 4 bands, the z-transform transfer function is given by

$$B_i(z) = \begin{cases} P_i(z) & i = 8 \\ P_i(z) - P_{i+1}(z) & i = 5, 6, 7, \end{cases} \tag{3.4}$$

where $P_i(z)$ is defined as

$$P_i(z) = \begin{cases} H(z^{2^{i-5}})F_{mc}(z)\prod_{k=2}^{i-5} F_m(z^{2^{k-1}}) & i = 7, 8 \\ H(z^2)F_{mc}(z) & i = 6 \\ H_c(z) & i = 5. \end{cases} \tag{3.5}$$

The implementation of an 8-band filter bank requires 10 sub-filters. However, the multipliers can be shared among interpolated $H(z)$ and $F_m(z)$. One suggested implementation of the interpolated $H(z)$ is given in Fig. 3.4 where multipliers are shared among $H(z)$, $H(z^2)$, $H(z^4)$ and $H(z^8)$. The frequency response masking is achieved by repeated use of $F_m(z)$.

The implementation of cascaded filters such as $F_m(z^2)F_m(z)$ can be effectively done through a hardware sharing scheme. Fig. 3.5 shows an example of the implementation of $F_m(z)F_m(z^2)$, where T is a latch used to store previous signal for the next round and multipliers are shared between $F_m(z)$ and $F_m(z^2)$.

3.3 Impacts of the Transition Bandwidth

The complexity of the proposed filter bank largely depends on the lengths of two prototype half-band filters. The prototype $H(z)$ determines transition bandwidth of each subband and directly affects the length of masking filter $F_m(z)$. This can

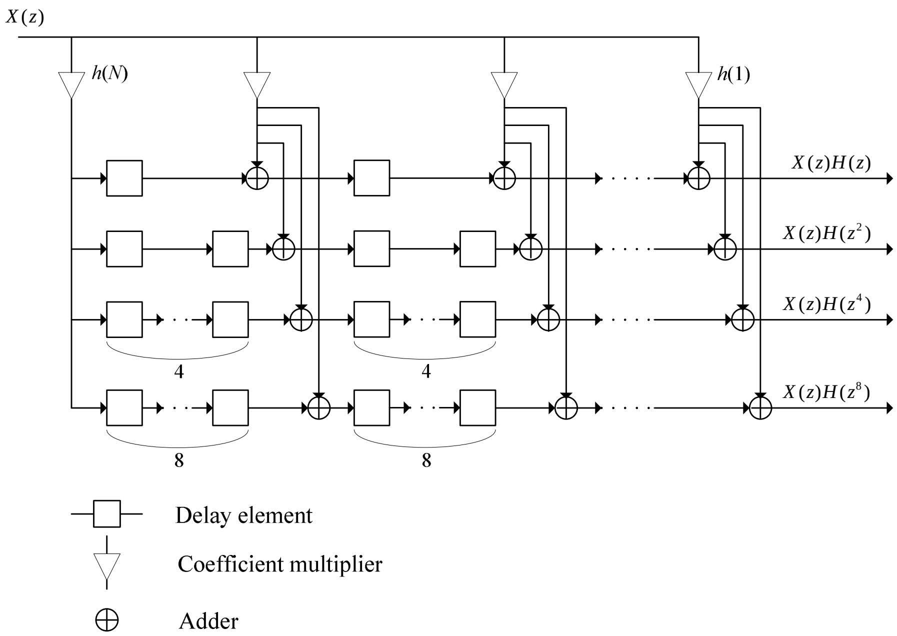


Figure 3.4: Sharing multipliers among $H(z)$, $H(z^2)$, $H(z^4)$ and $H(z^8)$.

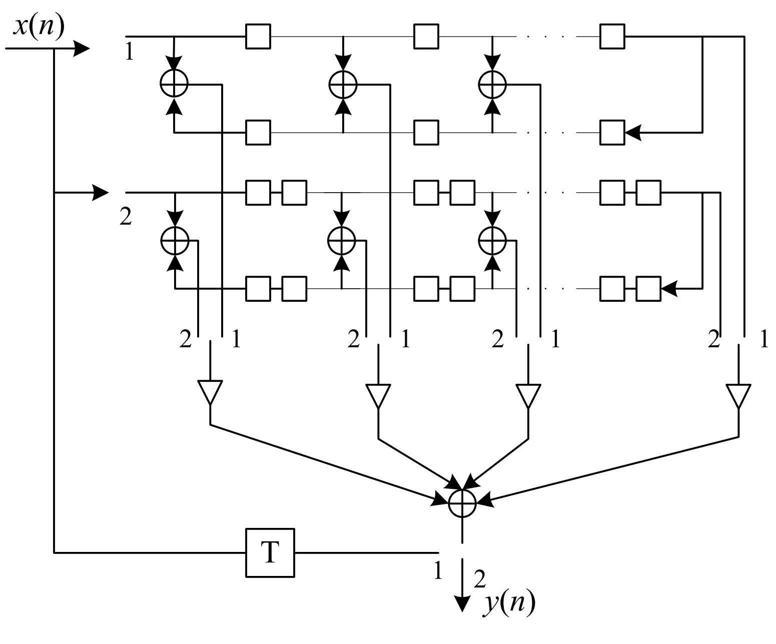


Figure 3.5: An example of the implementation of $F_m(z)F_m(z^2)$.

be seen from Fig. 3.2. An increase in transition bandwidth of $H(z)$ reduces its length, but narrows stopband width of $H(z^2)$ as shown in Fig. 3.2(b). As a result, the transition bandwidth of $F_m(z)$ decreases which contributes to the increases of its length. Furthermore, the matching between audiogram and magnitude response of the filter bank is closely related to transition bandwidth. To observe the influence of the transition bandwidth, several non-uniform filter banks with normalized transition bandwidths ranging from 0.05 to 0.4 were designed. The specifications of these filters are listed in Table 3.1.

Table 3.1: Filter bank's specifications.

Sampling frequency	16 kHz
Maximum passband ripple	0.0001
Minimum stopband attenuation	80dB
Number of bands	8

In order to observe the matching errors, the audiogram of presbycusis is selected as the objective curve, as shown in Fig. 2.3. Such a hearing loss is caused by aging and long-term cumulative exposure to sound energy. At the lower frequencies, the hearing sensitivity is good. Therefore, person with such an audiogram will have no problem knowing that someone is talking. However, at the high frequencies, the thresholds are below the consonant areas. He/She may not be able to distinguish one word from another due to the loss of too many consonants. The main feature of such an audiogram is that the hearing loss is severer at the high frequencies. The gain for each subband is chosen to be the mean of the two thresholds at the passband edges of the subband (Recruitment phenomenon [102] is not considered here).

The matching errors between presbycusis audiogram and magnitude responses of the filter banks with different transition bandwidth is shown in Fig. 3.6. Table 3.2 lists the lengths of prototype filter $H(z)$ and masking filter $F_m(z)$ at different transition bandwidths together with maximum matching errors. It is obvious that matching error decreases with increase of transition bandwidths till 0.25. Further increase in transition bandwidth worsens the matching errors due to large overlaps among different bands, especially at low and high frequencies where the subbands are narrow. Taking into consideration matching error and complexity, 0.25 is a reasonable choice for transition bandwidth.

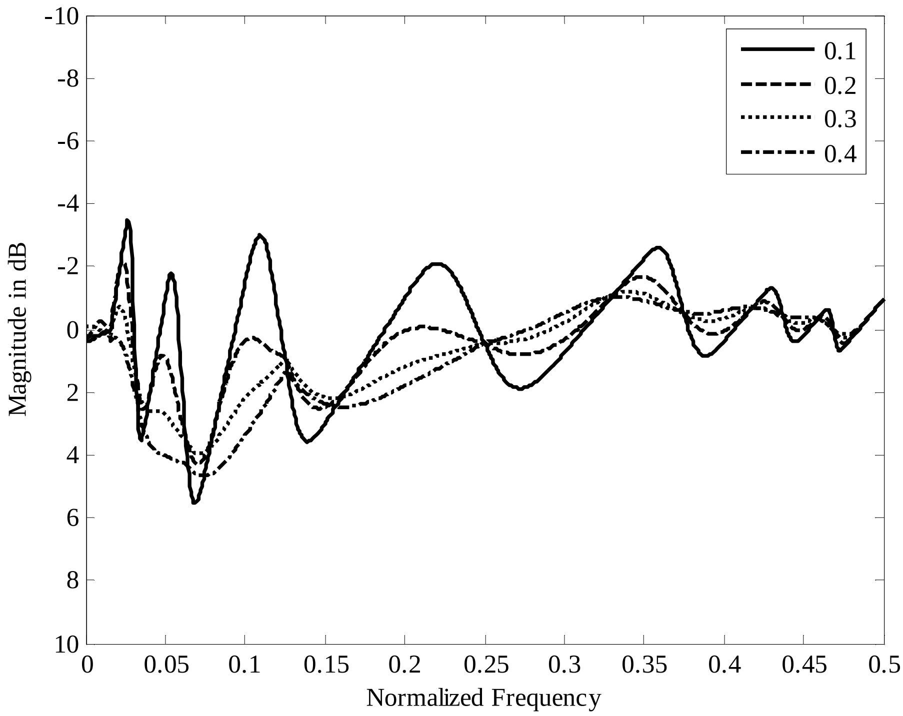


Figure 3.6: Matching errors of different transition bandwidths

Table 3.2: Impacts of the transition bandwidth.

Normalized transition bandwidth	Total length of ($H(z)$ and $F_m(z)$)	Maximum matching error(dB)
0.05	103+23=126	6.3747
0.10	51+27=78	5.5322
0.15	35+27=62	4.8981
0.20	27+35=62	4.4247
0.25	19+39=58	3.9635
0.30	15+51=66	3.9758
0.35	15+67=82	4.1059
0.40	11+95=106	4.6519

3.4 Design Examples

With the normalized transition bandwidth of 0.25 and sampling frequency of 16kHz, the normalized passband and stopband edges of $H(z)$ are at 0.125 and 0.375, respectively. The normalized passband and stopband edges of $F_m(z)$ are at 0.1875 and 0.3125 accordingly. For a minimum stopband attenuation of 40dB, the lengths of $H(z)$ and $F_m(z)$ are 11 and 19, respectively. 8 multiplications are needed to implement the filter bank. This is about 58% savings in terms of number of multipliers while the maximum matching error maintains the same, compared to that of the design in [98] with 8 uniform bands.

However, 40dB attenuation in stopband is not sufficient for matching the audiograms of severe hearing loss. For a minimum of 80dB stopband attenuation and normalized transition bandwidth of 0.25, the filter lengths of $H(z)$ and $F_m(z)$ are 19 and 39, respectively. 15 multiplications are needed to implement the filter bank.

The frequency response of the filter bank is shown in Fig. 3.7. The overall magnitude response that matches to the audiogram of presbycusis is shown in Fig. 3.8. It is clear that the match at high frequency is very close to the audiogram. The errors between the filter outputs and the audiogram are within +4dB to -2dB range.

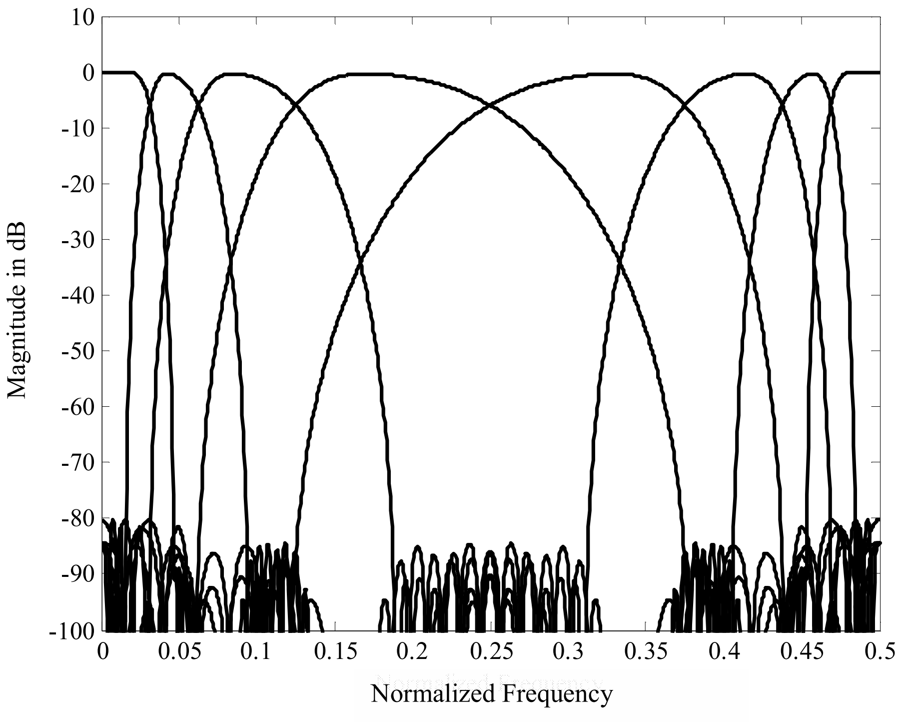


Figure 3.7: Frequency response of the proposed filter bank.

3.5 Optimization of the Gains

Till now, the gain of each subband is chosen to be the mean of the two thresholds at the passband edges of the subband. It is interesting to see whether there are

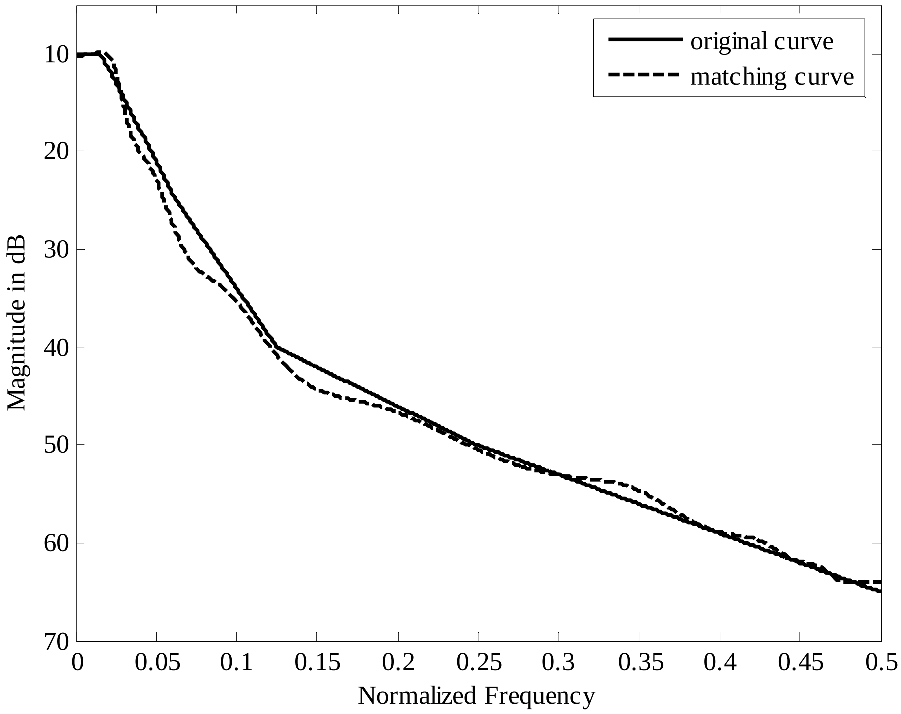


Figure 3.8: Matching result for audiogram of presbycusis.

better choices for the gains to improve the matching error. Suppose an audiogram is sampled and N data points are collected. a_i represents the intensity of the i^{th} sample among N data points. $b_{k,i}$ is defined as the intensity of the i^{th} sample within the k^{th} subband. The gain for each subband is x_k, $k = 1, \cdots, 8$. The gains should be chosen in such a way that they minimize the following least square error in (3.6). The optimized gain for each subband that minimizes $f(X)$ can be found using a Matlab routine.

$$f(X) = \sum_{i=1}^{N} \left(a_i - \sum_{k=1}^{8} b_{k,i} x_k \right)^2 \tag{3.6}$$

The maximum matching errors for different normalized transition bandwidths before and after gain optimization are listed in Table 3.3. All errors are found between presbycusis audiogram and the magnitude response of the proposed filter bank. The specifications of the filter banks are the same as that in Table 3.1.

It is clear that optimized gains help to reduce the maximum matching errors considerably. The best matching is achieved at the transition bandwidth of 0.27 and the lengths of $H(z)$ and $F_m(z)$ are 19 and 43, respectively. However, considering both the complexity and performance, 0.25 is chosen as the normalized transition bandwidth for the filter bank. For 0.25, the matching curves before/after optimization and matching errors are shown in Figs. 3.9 and 3.10, respectively.

Table 3.3: Comparison of maximum matching errors.

Normalized transition bandwidth	Maximum matching error before optimization(dB)	Maximum matching error after optimization (dB)
0.10	5.5322	4.3684
0.15	4.8981	2.9809
0.20	4.4247	1.9015
0.25	3.9635	1.1188
0.27	3.9011	1.0431
0.29	3.9072	1.1564
0.30	3.9758	1.2159

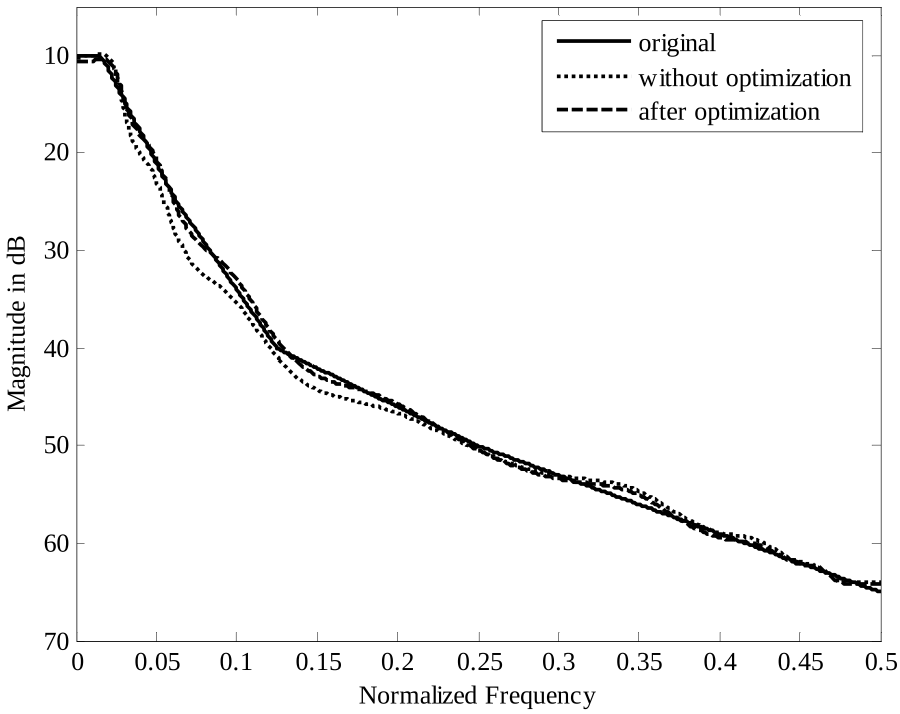


Figure 3.9: Matching curves before and after gain optimization.

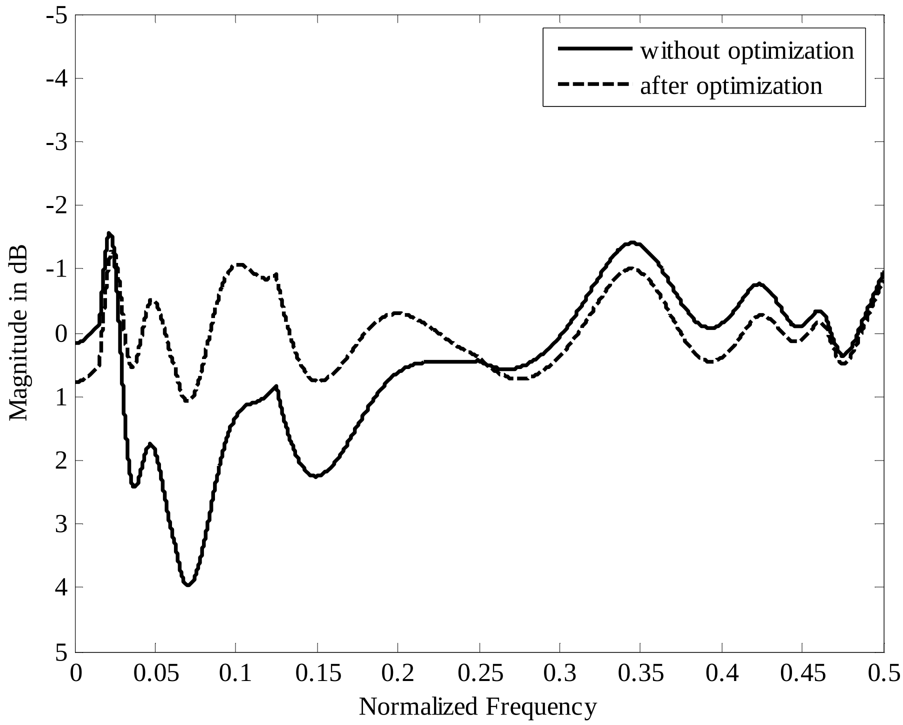


Figure 3.10: Matching errors before and after gain optimization.

3.6 Verification on Various Audiograms

Various audiograms are used to evaluate the effectiveness of the filter bank. These audiograms were downloaded from the Independent Hearing Aid Information, a public service by Hearing Alliance of America. The right ear thresholds (represented by 'O') were matched. The normalized transition bandwidth is set to 0.25 and other specifications are the same as that in Table 3.1. To gauge the impacts of optimization, define η as

$$\eta = \frac{f(X_0) - f(X_{opt})}{f(X_0)}, \tag{3.7}$$

where $f(X_0)$ is the least square error before gain optimization. $f(X_{opt})$ is the least square error after gain optimization. If the audiogram of presbycusis is used as the objective curve, the optimization reduces least square errors by 83%.

A. Mild to moderate hearing loss in high frequencies

People with such kind of hearing loss cannot hear s's, z's, th's, v's, and other soft, high frequency consonants. The audiogram is shown in Fig. 3.11(a). Figs. 3.11(b) and (c) show the magnitude responses of the filter bank and matching errors before and after optimization, respectively. ($\eta = 32.61\%$)

B. Mild hearing losses in the whole frequencies

In this case, the communication distance is reduced to 1 or 2 meters, whereas the distance is up to 12 meters for normal hearing. The audiogram is shown in Fig. 3.12(a). Figs. 3.12(b) and (c) show the matching results and matching errors, respectively. ($\eta = 9.97\%$)

C. Mild to moderate hearing loss in low frequencies

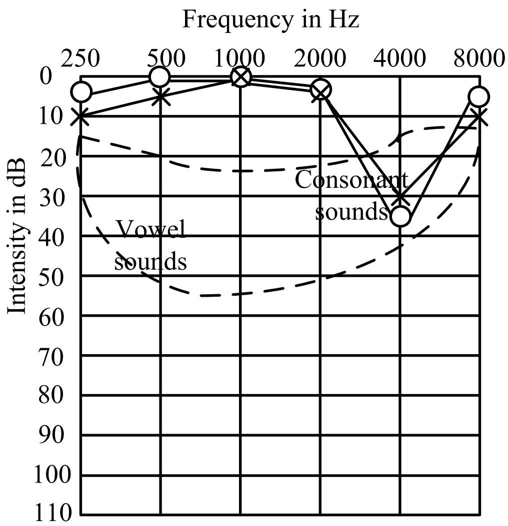
This hearing loss is unusual, but not rare. The primary effect will be a loss of overall loudness. The audiogram is shown in Fig. 3.13 (a). Figs. 3.13(b) and (c) show the matching results and matching errors, respectively. ($\eta = 22.42\%$)

D. Severe hearing loss in the middle to high frequencies

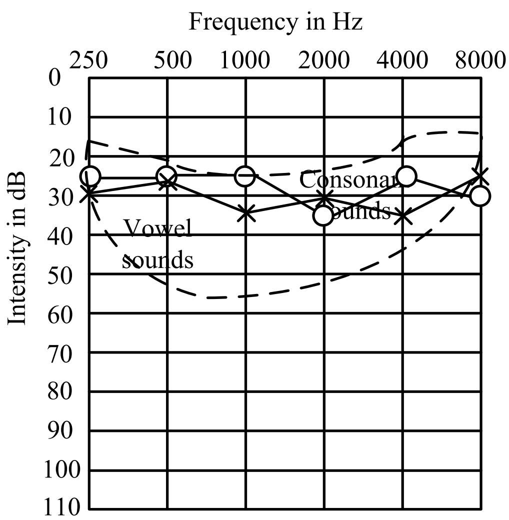
Older workers who are exposed to high level of noise for a long period in noisy industry usually exhibit such an audiogram, as shown in Fig. 3.14(a). Figs. 3.14(b) and (c) show the matching results and matching errors, respectively. Note that recruitment phenomenon is not considered here. ($\eta = 64.34\%$)

E. Profound hearing loss

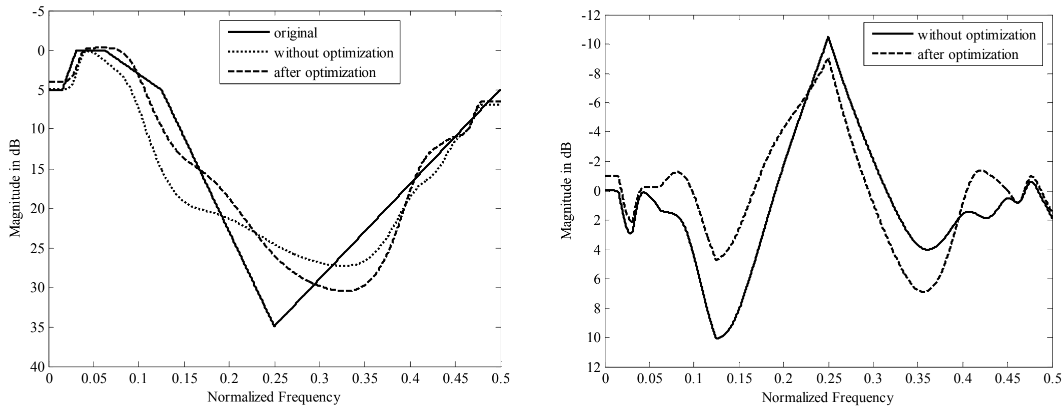
In this case the thresholds over most frequencies are under 90dB. The audiogram is shown in Fig. 3.15(a). Figs. 3.15(b) and (c) show the matching results and matching errors, respectively. It is important to note that although the proposed filter bank is able to compensate the profound hearing loss but the recruitment-phenomenon must be taken into consideration in a real application. ($\eta = 11.11\%$).



(a)

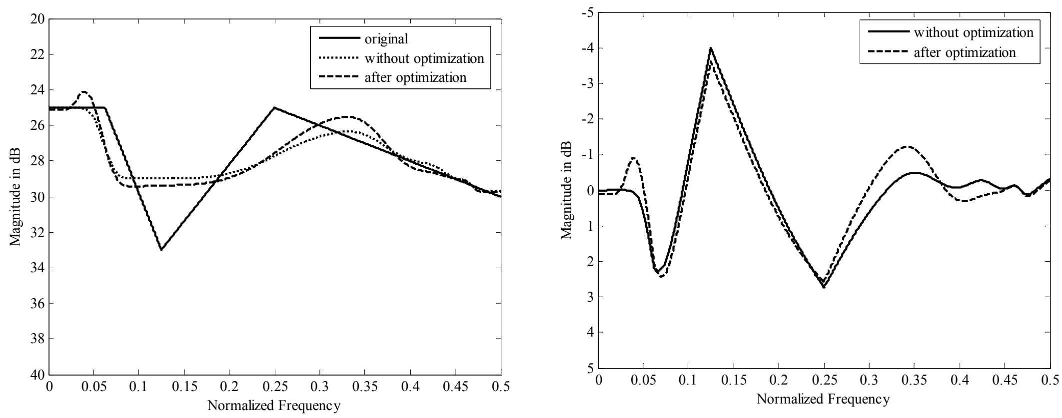


(a)



(b) (c)

Figure 3.11: (a) Audiogram for mild hearing loss in the high frequencies; (b) matching results; (c) matching error.



(b) (c)

Figure 3.12: (a) Audiogram for mild hearing losses in all frequencies; (b) matching results; (c) matching error.

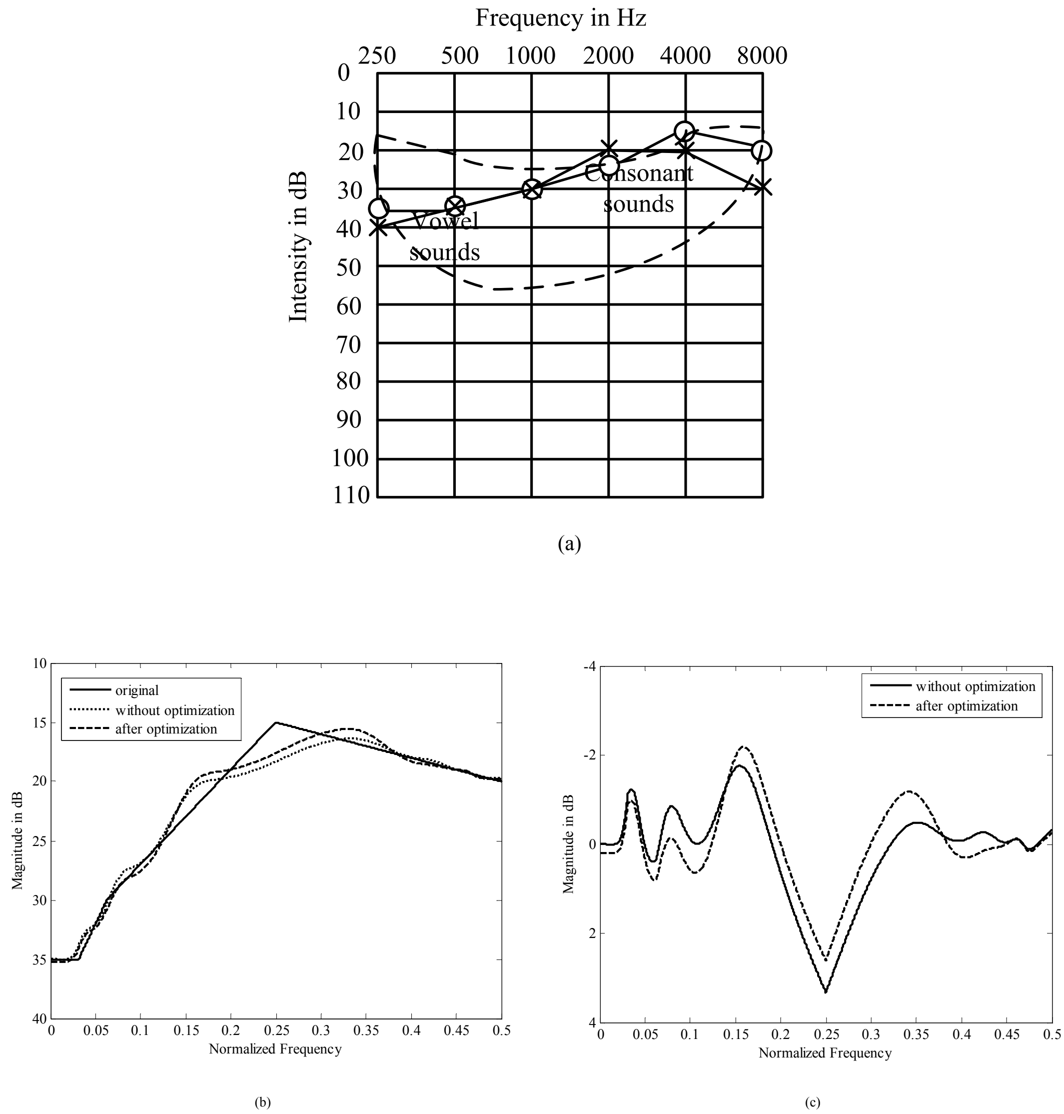


Figure 3.13: (a) Audiogram for mild to moderate hearing loss in low frequencies; (b) matching results; (c) matching error.

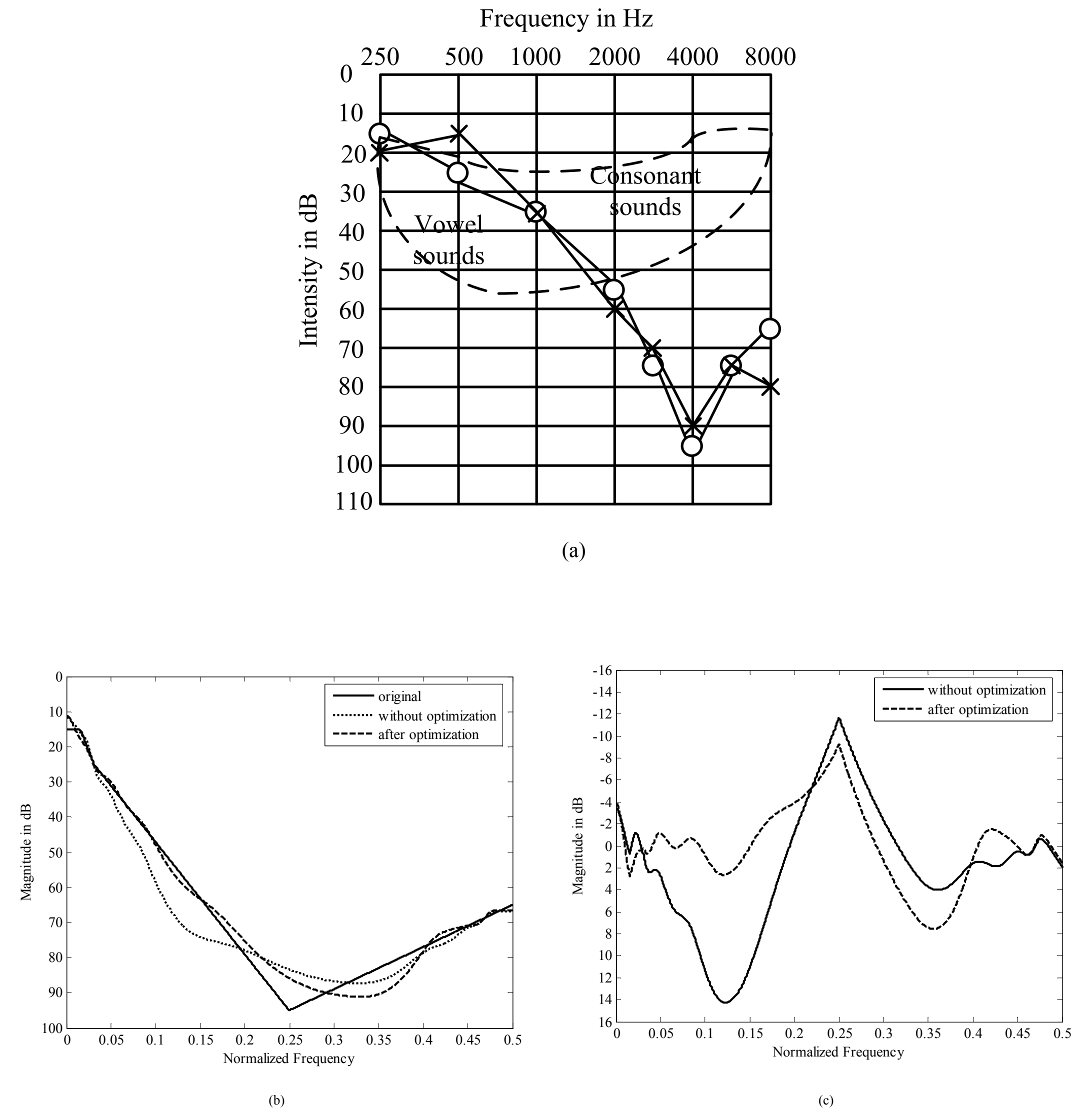


Figure 3.14: (a) Audiogram for Severe hearing loss in the middle to high frequencies; (b) matching results; (c) matching error.

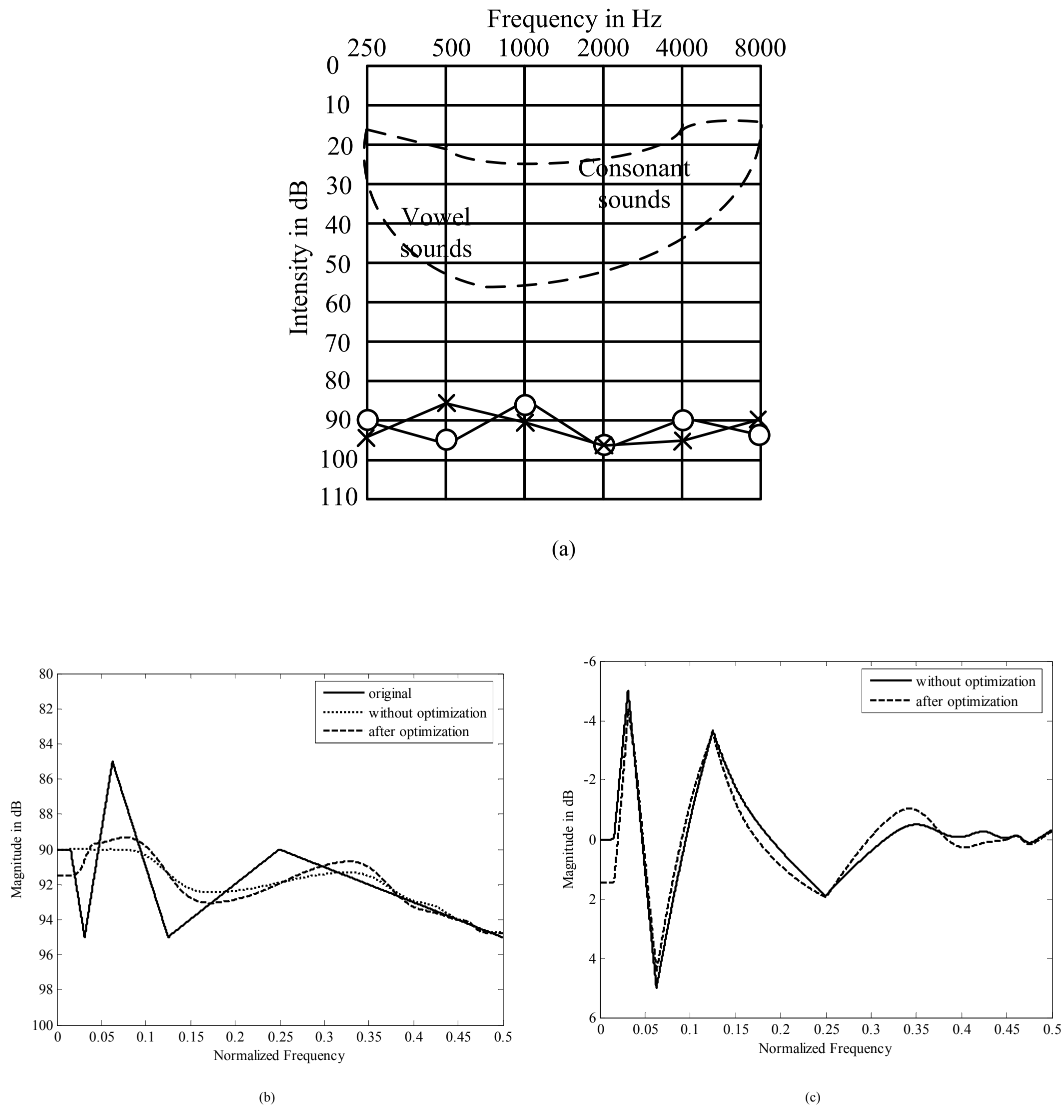


Figure 3.15: (a) Audiogram for profound hearing loss; (b) matching results; (c) matching error.

From Figs. 3.11 to 3.15, we can see that the proposed filter bank provides reasonable matching between various audiograms and magnitude responses of filter bank. The matching errors are very small if hearing loss occurs at high frequencies. The worst compensation happens when the hearing loss occurs in middle frequency range.

It would be interesting to see how the performance of the proposed non-uniform filter bank compares to that of a uniform filter bank. A uniform 8-band filter bank with subband width of 1 kHz and attenuation 80dB was designed. Both filter banks are used to match the audiogram of presbycusis as shown in Fig. 2.3. The non-uniform filter bank performs better at low and high frequencies, which benefits from the narrower subbands at low and high frequencies. Fig. 3.16 shows the matching errors.

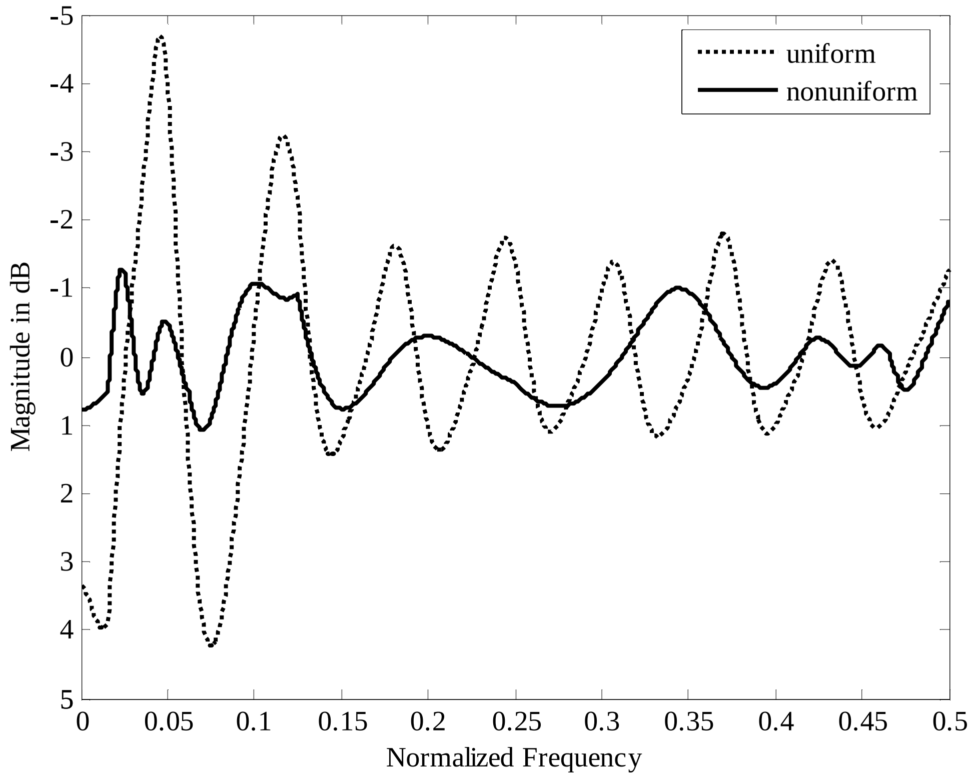


Figure 3.16: Comparison between uniform and non-uniform filter banks.

3.7 Summary

In this chapter, a low complexity 8-band non-uniformly spaced digital FIR filter bank has been proposed. It improves the matching between audiograms and the outputs of the filter bank due to non-uniform allocation of frequency bands. The use of two half-band FIR filters as prototype filters and the combination of frequency-response masking technique lead to significant savings in terms of number of multipliers, e.g. a minimum 80dB stopband attenuation is achieved using only 15 coefficients. The performance of the filter bank is enhanced by an optimal gain allocation process that helps to minimize the least square error between the objective audiogram and magnitude response of the filter bank.

Chapter 4

A 16-Band Non-uniform Low Delay Filter Bank for Hearing Aid

4.1 Introduction

As mentioned in Chapter 3, The main function of hearing aid is to amplify sound selectively and transfer it to the ear. This can be done by a filter bank that separates input signal into different frequency bands. Compared with analog approach, digital approach has the advantages including programmable capabilities, which are very flexible to the needs of hearing impaired people. An ideal hearing aid device should include features such as adjustable magnitude response, low power consumption and low processing delay. The effectiveness of compensation is relative to not only the quality of hearing aid devices but also the degree of hearing loss. Study has indicated that for hearing loss below 45dB, audibility is the most important problem, but for hearing loss beyond 45dB, besides audibility, discrimination, which cannot be compensated by amplification, plays a major part in hearing [103]. Though amplification is not the complete solution, it is a significant part for compensating hearing loss.

Much study has been invested into the design of efficient digital filter banks for selective amplification. Uniform filter banks have been widely used. One method is to design each sub-filter separately. This method is simple but seldom used because of high complexity. A more efficient way is to use fast modulation technique [104]. Subbands are realized by modulating a lowpass prototype filter to different frequencies. Since hearing loss could be more than 60dB, the length of the prototype filter can be few hundreds long. Complexity can be further reduced by IFIR filter design technique [97] [98]. A prototype filter is interpolated and then the undesired bands are removed by the masking filters. The transition bandwidths of the prototype filter and masking filters are usually wide. Thus the number of the taps is reduced effectively. However the IFIR filter bank design obtains its computational efficiency at the cost of a large group delay.

Non-uniform filter banks in hearing application are mostly used to mimic the bark scale which is from 1 to 24 barks, consisting with the first 24 critical bands of hearing [105]. The published schemes for implementing bark scale analysis can be categorized into several main approaches: direct analog or digital implementation [106][107], constant bandwidth [108] [109], constant-Q transform(CQT) [110], wavelet packet transform (WPT) [111][112] and combination of constant bandwidth and constant-Q transform [113][114]. The first several bands of bark scale are very narrow and the last several bands are extremely wide. However, hearing loss often occurs at middle and high frequencies. Non-uniform filter banks with wide subbands at high frequencies are not suitable for selective amplification. New scheme needs to be developed.

Since human ear has higher resolution at low frequencies than at high frequencies and the most common hearing loss occurs at high frequencies, if more subbands are allocated at both low and high frequencies, better compensation can be expected. For non-uniform filter banks, fast modulation technology is not directly applicable because the modulation results in identical response shape. In Chapter 3, an 8-band non-uniform FIR filter bank was proposed based on FRM technique [115]. However, the bandwidths of two mid bands are close to 2 kHz under a 16 kHz sampling frequency. This is not very attractive for certain hearing loss cases. Additionally, the filter bank has relatively long processing delay due to the use of large interpolation factor for complexity reduction. This is not good because long delays can cause mismatch in speech and lip-reading [100]. A difference between 4 to 5 msec will start to become noticeable to listeners with normal hearing. Differences between 10 to 15 msec may become objectionable [116]. In this chapter, a novel filter bank which has low complexity yet keeps a comparatively low delay was proposed.

The chapter is organized as follows. In Section 4.2, the details on how to form the 16-band non-uniform filter bank are presented. Then in Section 4.3, the implementation of the proposed structure is discussed. In Section 4.4, a design example is given. The influence of the transition bandwidth is discussed in Section 4.5. In Section 4.6, the effectiveness of the proposed filter bank and the delay are evaluated. Finally, conclusion is drawn in Section 4.7.

4.2 The Proposed 16-Band Non-uniform Filter Bank

To boost the computational efficiency, the construction of the proposed filter bank is based on the FRM technique. The proposed filter bank is designed in such a way that the lower and upper 8 subbands are symmetric. In order to improve the matching performance at low frequencies, we allocate more bands compared to a uniform filter bank. Since hearing level measurements are done at each octave: 250 / 500 / 1k / 2k / 4k / 8k in a standard audiogram, the objective band-edges of subbands are shown in Table 4.1.

Table 4.1: 3dB frequencies of the subbands.

Band	Lower 3dB frequency	Upper 3dB frequency
1	-	250
2	250	500
3	500	750
4	750	1000
5	1000	1500
6	1500	2000
7	2000	3000
8	3000	4000
9	4000	5000
10	5000	6000
11	6000	6500
12	6500	7000
13	7000	7250
14	7250	7500
15	7500	7750
16	7750	-

The lower 8 subbands can be obtained by subtracting the frequency response of one lowpass filter from that of another, as shown in Equation (4.1) and Fig. 4.1.

$$B_i(z) = \begin{cases} P_i(z) & i = 1 \\ P_i(z) - P_{i-1}(z) & i = 2, \cdots, 8. \end{cases} \tag{4.1}$$

The lowpass filters $P_i(z), i = 1, \cdots, 8$ can be divided into three groups based on their upper 3dB frequencies. The first group consists of $P_4(z)$, $P_6(z)$ and $P_8(z)$ whose 3dB frequencies are 1000Hz, 2000Hz and 4000Hz respectively. The second group consists of $P_3(z)$, $P_5(z)$ and $P_7(z)$ whose 3dB frequencies are 750Hz, 1500Hz and 3000Hz respectively. And the third group consists of $P_1(z)$ and $P_2(z)$ whose 3dB frequencies are 250Hz and 500Hz respectively. Since the 3dB frequencies in

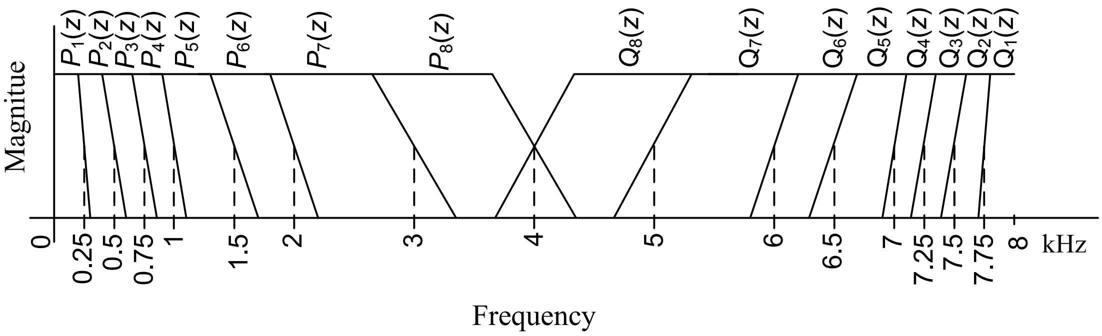


Figure 4.1: The frequency response of lowpass filters $P_i(z)$ and highpass filters $Q_i(z)$, $i = 1, \cdots, 8$.

each group are at octaves, the filters in the same group can be obtained by first interpolating a prototype filter and then removing the undesired passbands using one or several masking filters. In order to reduce the complexity, one of the prototype filters can also be used as a masking filter by careful design. According to the 3dB frequencies of the three groups, three prototype filters $H_1(z)$, $H_2(z)$ and $H_3(z)$ with 3dB frequencies at 4000Hz, 3000Hz and 1000Hz are employed to produce the subbands. The z-transform transfer functions of the lowpass filters $P_i(z), i = 1, \cdots, 8$ are shown in Table 4.2. $H_1(z)$ is used as the masking filter to remove the unwanted passbands.

The upper 8 subbands can be obtained by subtracting the frequency response of one highpass filter from that of another, as shown in Equation (4.2) and Fig. 4.1.

$$B_i(z) = \begin{cases} Q_{16-i+1}(z) & i = 16 \\ Q_{16-i+1}(z) - Q_{16-i}(z) & i = 9, \cdots, 15. \end{cases} \tag{4.2}$$

Table 4.2: z-transform transfer functions of lowpass filters $P_i(z), i = 1, \cdots, 8$.

Filter	z-transform transfer functions
$P_1(z)$	$H_3(z^4)H_1(z^2)H_1(z)$
$P_2(z)$	$H_3(z^2)H_1(z)$
$P_3(z)$	$H_2(z^4)H_1(z^2)H_1(z)$
$P_4(z)$	$H_1(z^4)H_1(z^2)H_1(z)$
$P_5(z)$	$H_2(z^2)H_1(z)$
$P_6(z)$	$H_1(z^2)H_1(z)$
$P_7(z)$	$H_2(z)$
$P_8(z)$	$H_1(z)$

The highpass filters $Q_i(z), i = 1, \cdots, 8$ are symmetric with the lowpass filters $P_i(z), i = 1, \cdots, 8$ at the mid-frequency point. Therefore, the prototype filters which used to form the passbands of the lowpass filters form the passbands of the highpass filters simultaneously when they are interpolated by even interpolation factors. The z-transform transfer functions of the highpass filters $Q_i(z), i = 1, \cdots, 8$ are shown in Table 4.3.

Table 4.3: z-transform transfer functions of highpass filters $Q_i(z), i = 1, \cdots, 8$.

filter	z-transform transfer functions
$Q_1(z)$	$H_3(z^4)H_1(z^2)H_{1c}(z)$
$Q_2(z)$	$H_3(z^2)H_{1c}(z)$
$Q_3(z)$	$H_2(z^4)H_1(z^2)H_{1c}(z)$
$Q_4(z)$	$H_1(z^4)H_1(z^2)H_{1c}(z)$
$Q_5(z)$	$H_2(z^2)H_{1c}(z)$
$Q_6(z)$	$H_1(z^2)H_{1c}(z)$
$Q_7(z)$	$H_{2h}(z)$
$Q_8(z)$	$H_{1c}(z)$

Note $H_{1c}(z)$ and $H_{2h}(z)$ in Table 4.3 are the complement of $H_1(z)$ and the mirror image of $H_2(z)$ according to the center frequency respectively. They are defined as

$$H_{1c}(z) = z^{-(N_1-1)/2} - H_1(z), \tag{4.3}$$

$$H_{2h}(z) = \sum_{n=1}^{N_2} h_2(n)(-z^{-1})^n = (\sum_{n=2k} h_2(n)z^{-n} - \sum_{n=2k+1} h(n)z^{-n}), \tag{4.4}$$

where N_1 and N_2 are the filter lengths of $H_1(z)$ and $H_2(z)$.

According to Equations (4.1),(4.2) and Tables 4.2 and 4.3, the proposed architecture is presented in Fig. 4.2. There are two outputs for $H_1(z)$ and $H_2(z)$. The lower branch represents the complementary output for $H_1(z)$ and the mirror image output for $H_2(z)$ according to the center frequency. Therefore, the upper branches of the rightmost filters generate eight lowpass filters and the lower branches generate eight highpass filters. The working principle of the proposed filter bank can be best illustrated by an example. Suppose we want to form the third subband $B_3(z)$. According to Fig. 4.1, $B_3(z) = P_3(z) - P_2(z)$ has lower and upper cut-off frequency at 500Hz and 750Hz respectively. Four filters, $H_2(z^4)$, $H_3(z^2)$, $H_1(z^2)$ and $H_1(z)$, are involved according to Table 4.2. The magnitude responses of various sub-filters are shown in Fig. 4.3. The cutoff frequencies of $H_1(z)$, $H_2(z)$ and $H_3(z)$ are listed in Table 4.4 under a sampling frequency of 16kHz.

The interpolated filter $H_2(z^4)$ produces an output with three passbands, as shown by the solid line in Fig. 4.3(a). In order to form the required lowpass filter as shown

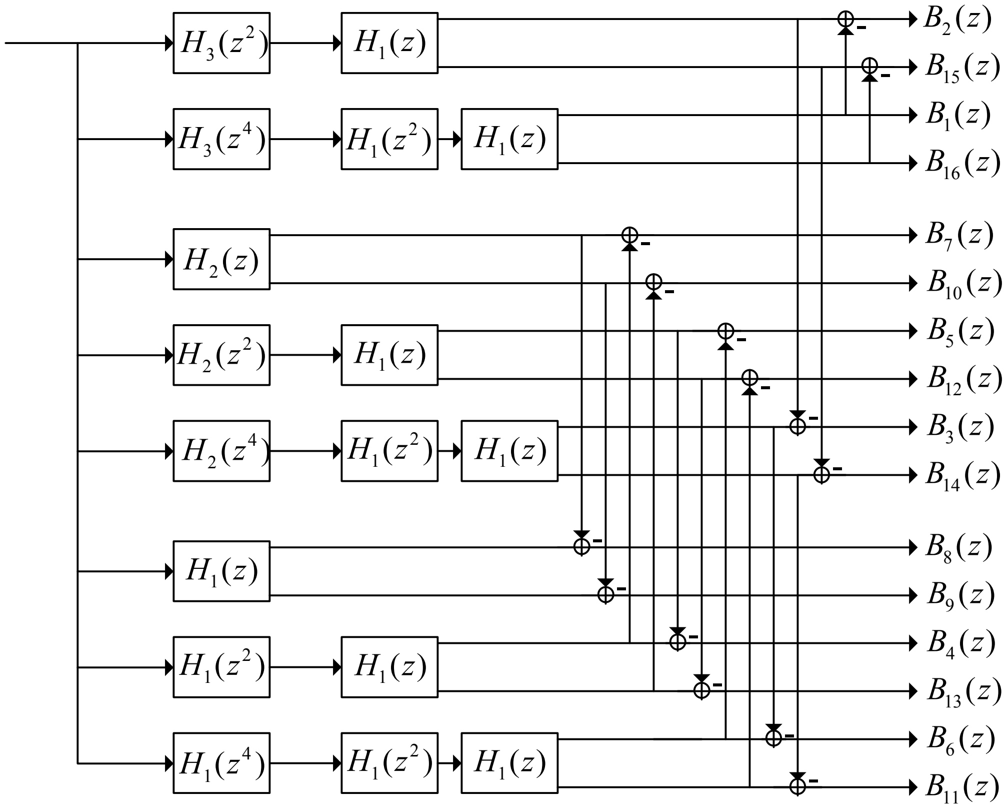


Figure 4.2: The block diagram of the 16-band non-uniform filter bank.

Table 4.4: Cutoff frequencies of $H_1(z)$, $H_2(z)$ and $H_3(z)$

Filter	Cut-off frequency	Normalized cut-off frequency
$H_1(z)$	4kHz	0.25
$H_2(z)$	3kHz	0.1875
$H_3(z)$	1kHz	0.0625

in Fig. 4.3(b), the unwanted passbands at mid and high frequencies are removed by masking filters $H_1(z^2)$ and $H_1(z)$. Similarly, $H_3(z^2)$ is filtered by $H_1(z)$ to produce a lowpass filter. The subband $B_3(z)$ is generated by subtracting the output in Fig. 4.3(d) from that in Fig. 4.3(b). Note that leading delays should be added to the sub-filters to make sure that they have the same group delay when connected in parallel. This is necessary to avoid frequency dependant delay [101].

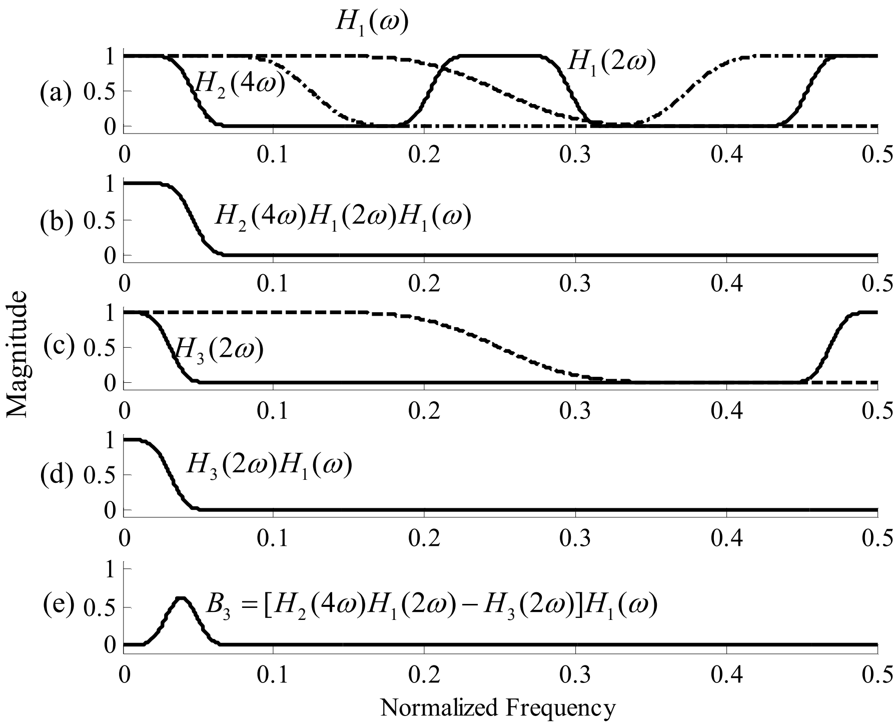


Figure 4.3: The formation of subbands $B_3(z)$.

4.3 Implementation of the Filter Bank

Power consumption is mainly related to the number of multipliers. Models sharing multipliers and flip-flops in the proposed structure would be very helpful.

In Fig. 4.2, there are four basic structures. One is the complementary structure. Based on the prototype filter $H_1(z)$, the hardware cost of producing the complement $H_{1c}(z)$ can be minimized because the delays are shared between the original and complementary filters, as shown in Fig. 4.4.

Secondly, frequency masking is achieved by a cascaded structure consists of $H_1(z)$ and its interpolated version. This cascaded structure can be effectively implemented through a multiplier-sharing scheme. An example is shown in Fig. 4.4, where T is a latch used to store previous signal for the next round and multipliers are shared between $H_1(z)$ and $H_1(z^2)$.

Additionally, we need to produce highpass filter $H_{2h}(z)$ which is symmetric with $H_2(z)$ around the mid-point frequency. Recall Equation (4.4), $H_{2h}(z)$ can be obtained by substituting substraction for addition alternately.

Now considering the parallel structure such as $H_2(z)$, $H_2(z^2)$ and $H_2(z^4)$. Since they have the same input and the filters are interpolated filters from the same pro-

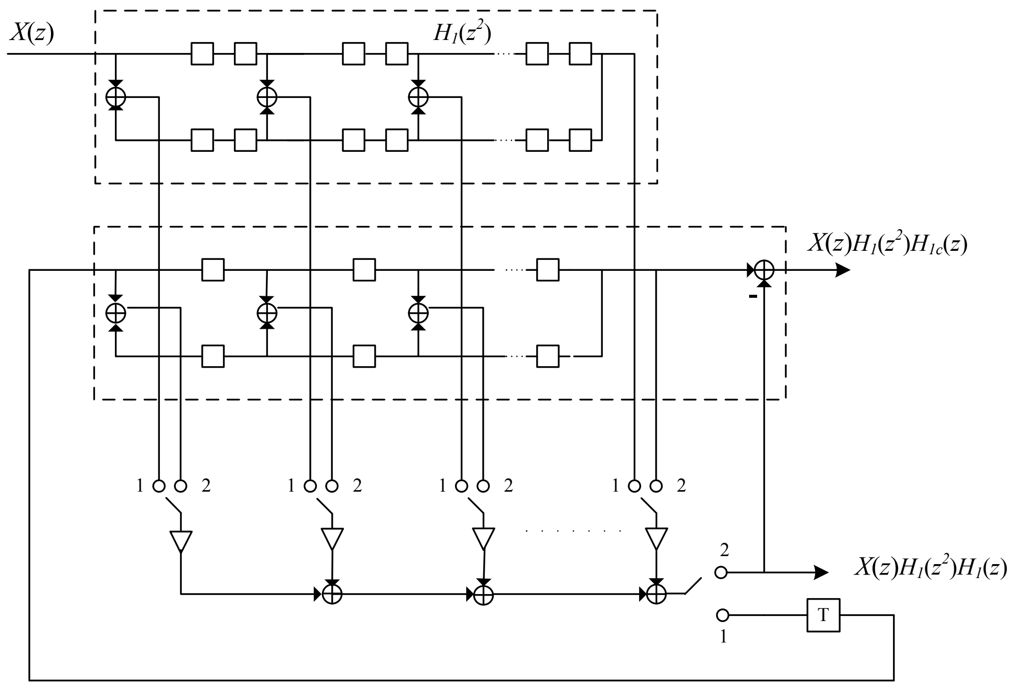


Figure 4.4: Implementation of cascaded structure.

totype filter, the multipliers, delays and adders can be shared among these filters. An implementation example with $H_{2h}(z)$ is shown Fig. 4.5. For this example, the length of $H_2(z)$ is 5.

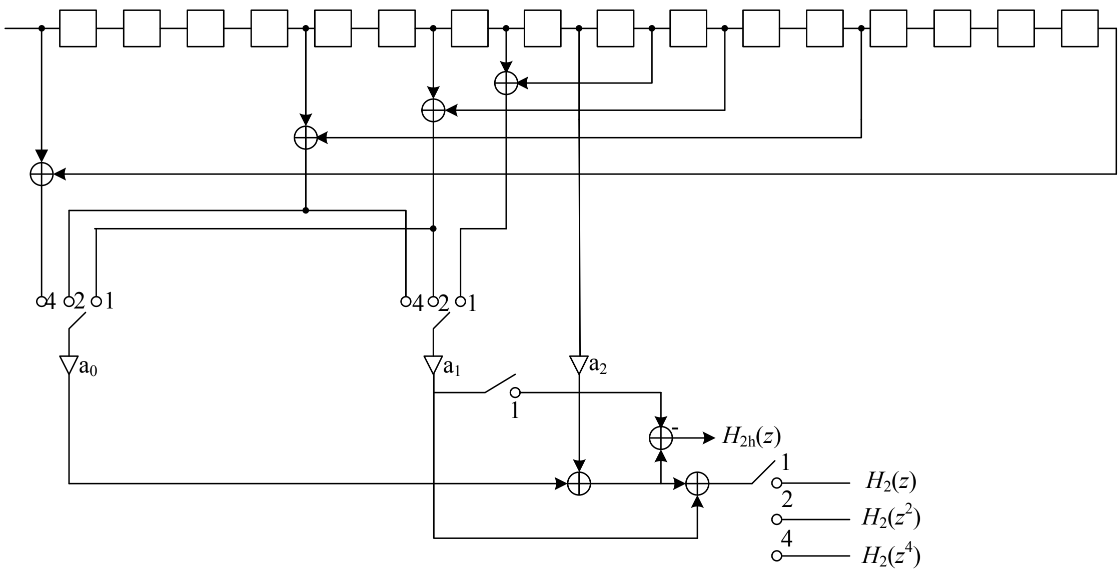


Figure 4.5: Implementation of parallel structure.

4.4 A Design Example

Suppose an audiogram is sampled and N data points are collected. a_i represents the intensity of the i^{th} sample among N data points. Similarly, define the sampled magnitude response of the filter bank as $B_k = [b_{k,1}, b_{k,2}, \cdots, b_{k,N}]$, where $b_{k,i}$ is the intensity of the i^{th} sample within the k^{th} subband. The gain for each subband is

x_k, $k = 1, \cdots, 8$. let

$$A = [a_1, a_2, \cdots, a_N]^T, \tag{4.5}$$

$$B = [B_1, B_2, \cdots, B_{16}], \tag{4.6}$$

and

$$X = [x_1, x_2, \cdots, x_{16}]^T. \tag{4.7}$$

In order to make sure that input signal in each subband is amplified properly, the gains are chosen in such a way that they minimize the maximum matching error between the objective curve and the output of the filter bank, as shown in (4.8). The optimized gain for each subband can be easily found using a Matlab program named fminimax.

$$\underset{\mathrm{x}}{\mathrm{minimax}} \left\{ ||A - B \cdot X|| \right\}. \tag{4.8}$$

Suppose a filter bank with normalized transition bandwidth 0.15 is the desired filter bank. According to the cut-off frequencies of the three prototype filters shown in Table 4.4, the passband and stopband edges of $H_1(z)$, $H_2(z)$ and $H_3(z)$ are at (0.1750, 0.3250), (0.1125, 0.2625) and (0.025, 0.10) accordingly. For a minimum stopband attenuation of 60dB, the lengths of $H_1(z)$, $H_2(z)$ and $H_3(z)$ are 27, 25 and 49, respectively. Since $H_3(z)$, $H_2(z)$ and $H_1(z)$ are linear phase filters and $H_1(z)$ is also a half band filter, 45 coefficients are needed to implement the 16-band filter bank. The frequency response of the filter bank is shown in Fig. 4.6.

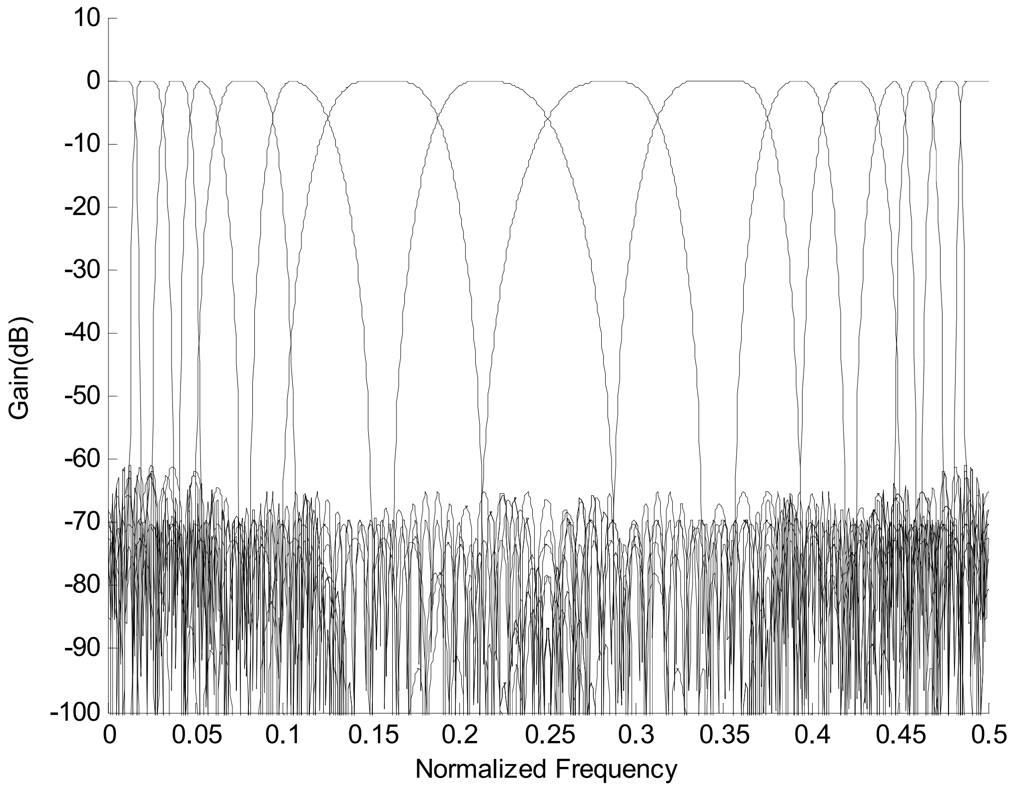


Figure 4.6: Frequency response of the proposed filter bank.

4.5 Influence of the Transition Bandwidth

The outputs of the rightmost filters of Fig. 4.2 are eight lowpass and eight highpass frequency responses. They are named from 1 to 16 from low frequency to high frequency. The design starts with the selection of bandedges of the three prototype filters. Suppose the transition bandwidth of $H_1(z)$ is $\triangle B$, the transition bandwidths of $H_2(z)$ and $H_3(z)$ are set to be $\triangle B$ and $\triangle B/2$ to ensure that the passband edge of the i^{th} lowpass filter $\omega_{p,i}$, is larger than the stopband edge of the $(i-1)^{th}$ lowpass filter $\omega_{s,i-1}$.

The transition bandwidths of the prototype filters determine the complexity of overall filter bank and affect the matching errors. To find out their influence, several non-uniform filter banks with normalized transition bandwidths ranging from 0.05 to 0.22 were designed. The hearing loss pattern due to aging shown in Fig. 2.3 is selected as the objective curve because it is the most common hearing loss. Matching errors and filter complexities are listed in Table 4.5.

It can be seen that as the transition bandwidth increases, the matching error becomes smaller. This is because the matching curves get smoother. Also the complexity decreases as the transition bandwidth increases. This can be expected because the complexity of the prototype filters is inversely proportional to the transition bandwidth. Considering both the complexity and the maximum matching error, 0.2 is a good choice as the value of $\triangle B$. It requires 34 multiplications in the implementation of a filter bank with stopband attenuation of 60dB.

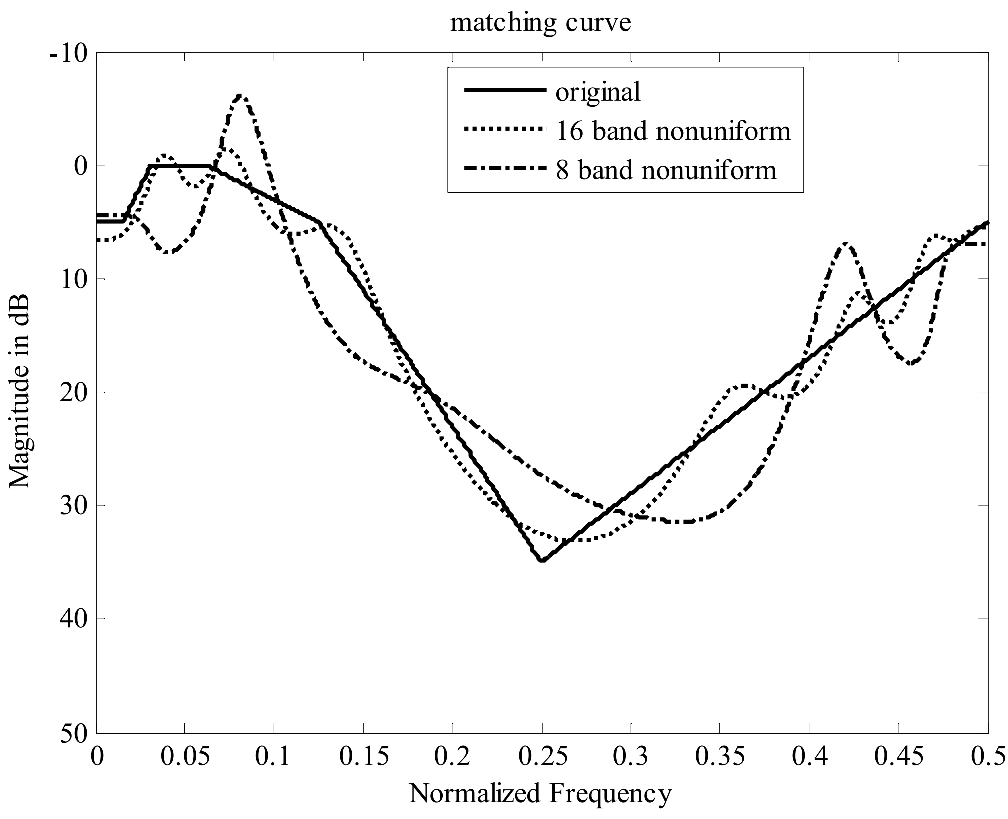
Table 4.5: Influence of the transition bandwidth.

Transition bandwidth (normalized)	Multiplications	Maximum matching error(dB)(Minimax)
0.05	75+38+19=132	2.1044
0.10	38+19+10=67	0.5587
0.15	25+13+7=45	0.3961
0.18	21+10+6=39	0.3974
0.20	19+10+5=34	0.3837
0.22	19+9+5=33	0.4113

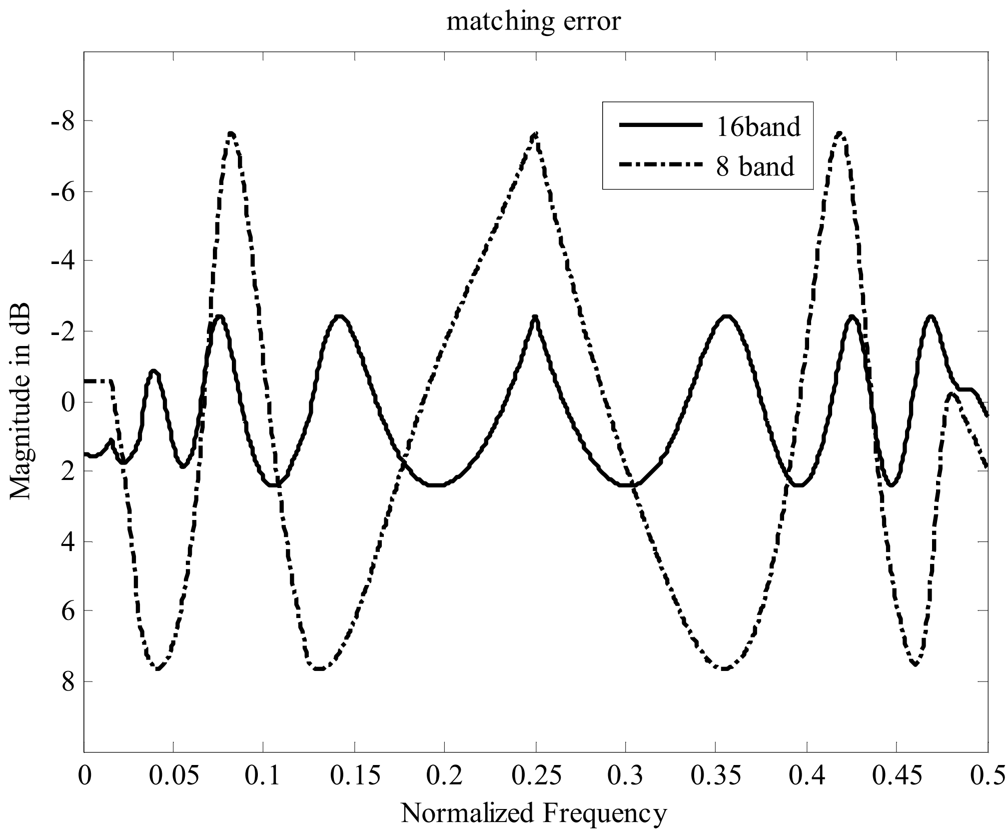
4.6 Performance Evaluation

The performance of the proposed filter bank is compared with the 8-bank non-uniform filter bank in Chapter 3 and a uniform counterpart. The effectiveness is examined using two types of audiograms: (1) the mild hearing loss in the high frequencies and (2) the mild to moderate hearing loss in low frequencies. In the first case, the communication distance is reduced to 1 or 2 meters, whereas the distance is up to 12 meters for normal hearing. In the second case, the primary effect will be a loss of overall loudness.

The matching results of the proposed filter bank and the 8-band non-uniform filter bank is shown in Fig. 4.7 for mild hearing loss in high frequency. The maximum matching error for the proposed filter bank is 2.42 dB against 7.65 dB of the 8-band non-uniform filter bank. The matching curve and matching error for mild to moderate hearing loss in low frequencies is shown in Fig. 4.8. The maximum matching error for the proposed filter bank is about 0.26 dB, while the maximum matching error of the 8-band non-uniform filter bank is 2.11 dB. The above examples clearly indicate the performance enhancement of proposed filter bank when compared with the 8-band non-uniform filter bank. The above comparisons do not take into consideration the recruitment-phenomenon [102].

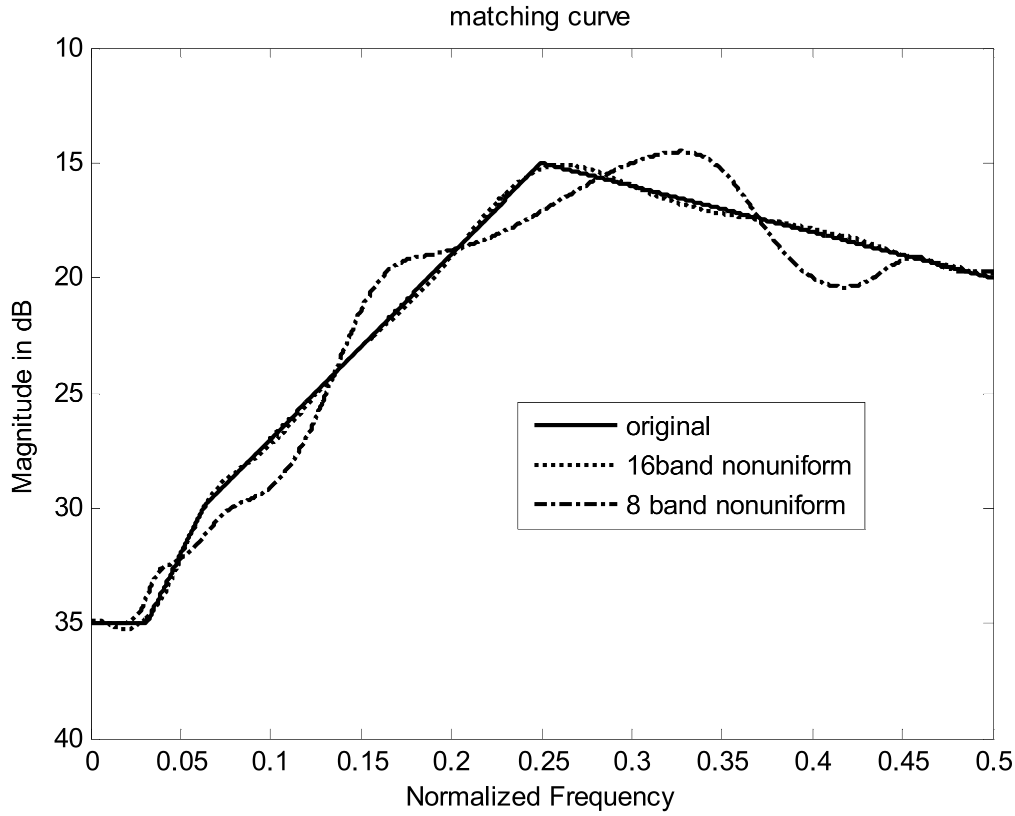


(a)



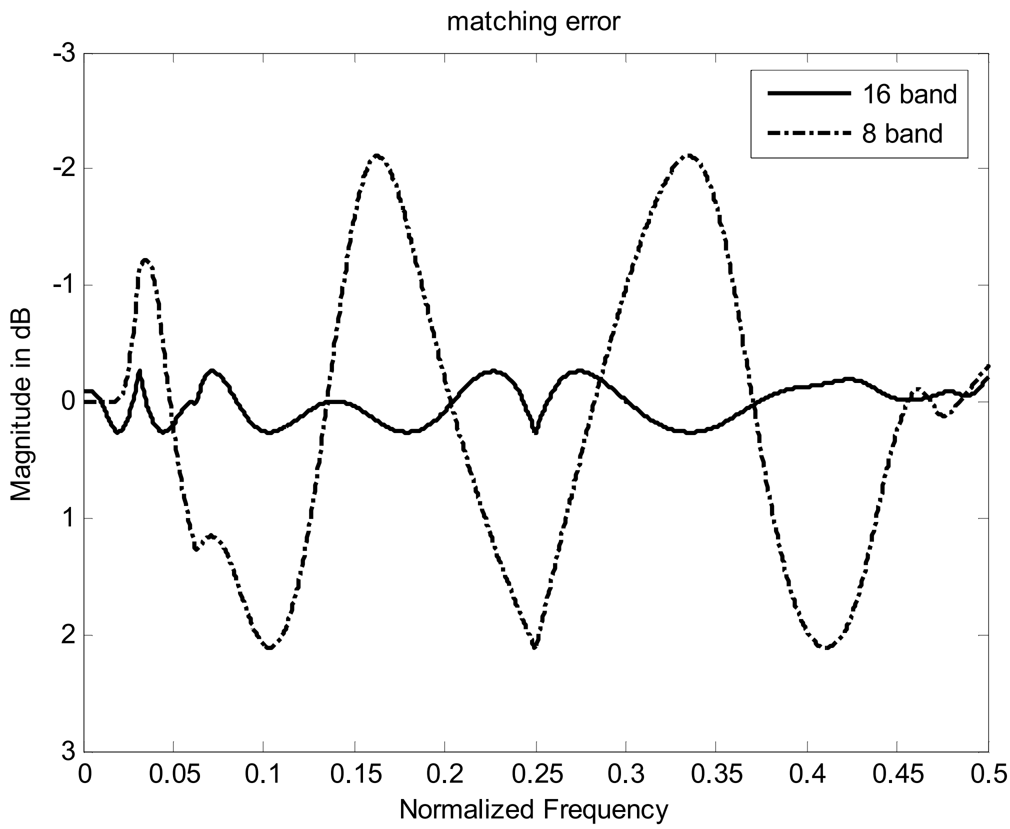
(b)

Figure 4.7: Matching results for mild hearing losses in high frequencies (a) matching curve; (b) matching error.



(a)

It is interesting to compare the performance of the proposed filter bank with that of a 16-band uniform filter bank. The assumption is that the two types of filter banks have the same number of subbands, use the same method to determine the gain and match the same audiogram. We used the efficient approach proposed in [97] to design the uniform filter bank. Four linear phase half band FIR filters with lengths 131, 67, 35 and 19, respectively, are used to produce the 16 subbands with attenuation of 60dB. 64 multiplications are needed to obtain all the subbands. Comparing to the uniform scheme, the propose structure achieves more than 46%



(b)

Figure 4.8: Matching results for mild to moderate hearing loss in low frequencies(a) matching curve; (b) matching error.

savings in terms of the number of multiplications.

Audiograms with hearing loss in low frequency are most suitable to the comparison. The matching errors of uniform and non-uniform filter banks for mild to moderate hearing loss in low frequencies is shown in Fig. 4.9. The proposed filter bank performs better than the uniform one especially in low frequency range.

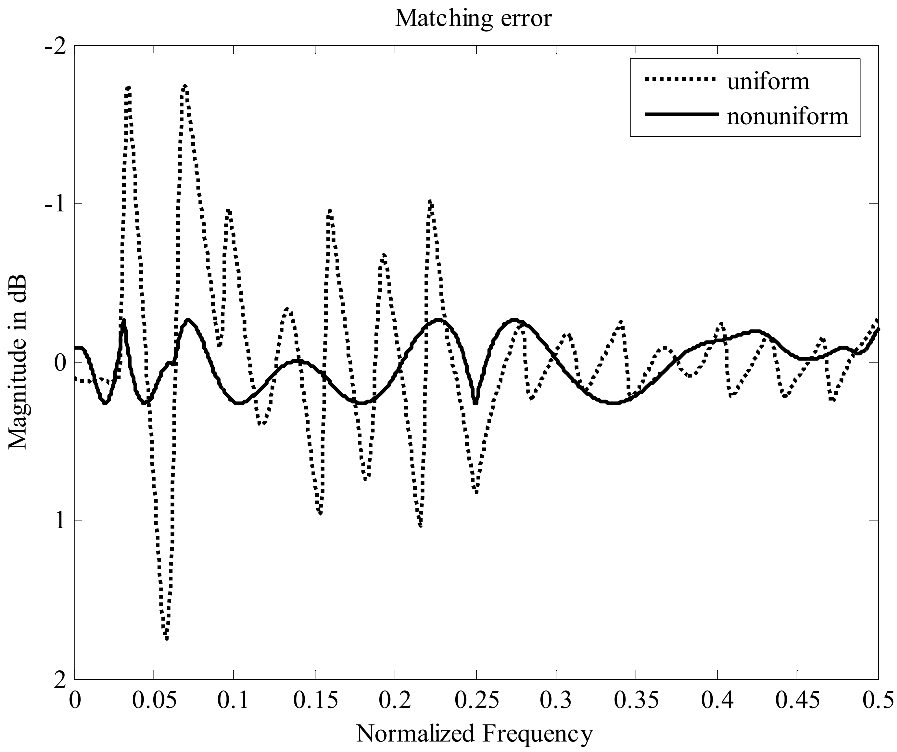


Figure 4.9: Matching errors of the proposed filter bank and uniform filter bank for most common hearing loss.

Another advantage of the proposed structure is that it reduces delay greatly while achieving the low complexity. This is very important because the quality of hearing can be enhanced by combining speech with lip-reading. To avoid mismatch in speech and visual observation, the processing delay should be small. The group delay of the proposed filter bank is 6.4ms under the sampling frequency of 16 kHz, which is about 48% of the delay in Chapter 3 which is about 13 ms and 37.8% of the delay of the uniform filter bank which is about 17ms. The delay is reduced effectively because more prototype filters and smaller interpolation factors are employed compared with the structure in Chapter 3 and the uniform structure.

Comparison of the number of multipliers, the number of adders and group delay among the non-uniform 8-band, non-uniform 16-band and uniform 16-band filter banks is shown in Table 4.6.

Table 4.6: Comparison of the three different filter banks

Filter bank	Multipliers	Adders	Delays(ms)
nonuniform 8-band	15	30	13
nonuniform 16-band	34	64	6.4
uniform 16-band	64	128	17

4.7 Summary

In this chapter, a 16-band non-uniformly spaced digital FIR filter bank has been proposed. The use of frequency-response masking technique leads to significant savings in terms of number of multiplications. The overall filter delay is significantly reduced as the result of novel filter structure which reduces the interpolation factor for the prototype filters. Comparing to the filter bank in Chapter 3, the proposed filter bank doubles the number of bands yet reduces delay by 48%.

Chapter 5

Low Complexity Design of Sharp FIR Filters Based on Frequency-Response Masking Approach

5.1 Introduction

The frequency response masking technique [16] is an efficient way to design sharp filters with arbitrary bandwidth. The zero-phase frequency response of a FRM filter can be written as

$$H(\omega) = G(\omega)H_{ma}(\omega) + [1 - G(\omega)]H_{mc}(\omega), \tag{5.1}$$

where $H_{ma}(\omega)$ and $H_{mc}(\omega)$ are the frequency response of the masking filters whose lengths must be both odd or even. $G(\omega)$ is the frequency response of the band-edge shaping filter. For conventional FRM, $G(\omega) = H_a(M\omega)$. That means instead of designing a sharp filter $H(z)$ with transition bandwidth $\triangle B$ directly, a filter $H_a(z)$ with transition bandwidth $M \cdot \triangle B$ is designed and then interpolated by a factor of M. The unwanted passbands are removed by two masking filters $H_{ma}(z)$ and $H_{mc}(z)$. The overall filter $H(z)$ is obtained by combining the outputs of the two masking filters.

The role of $G(z)$ is to form the desired transition band. An interpolated filter $H_a(z^M)$ is a good candidate of $G(z)$. However it is not necessary for $G(z)$ to have periodical passbands and stopbands as long as it serves the role of band-edge shaping.

Much effort has been made to reduce the complexity of FRM scheme [21–23]. In this chapter, a novel structure which can further reduce the complexity of FRM filters is proposed. The basic idea is to employ single filter frequency masking approach [6] [7] to design $G(z)$.

This chapter is organized as follows. In Section 5.2, the idea of the proposed scheme is presented. Then in Section 5.3, the details of the design procedure are described. Ripple analysis is presented in Section 5.4. Implementation of the proposed structure is discussed in Section 5.5. In Section 5.6, a design example is given. In Section 5.7, the proposed structure is extended. Conclusion is drawn in Section 5.8.

5.2 Proposed Scheme

The proposed structure is shown in Fig. 5.1. At first glance, the structure looks like a 2-stage FRM filter. However, this is not the case. For a 2-stage FRM filter, P should equal to Q, while in this structure, P and Q may not be the same.

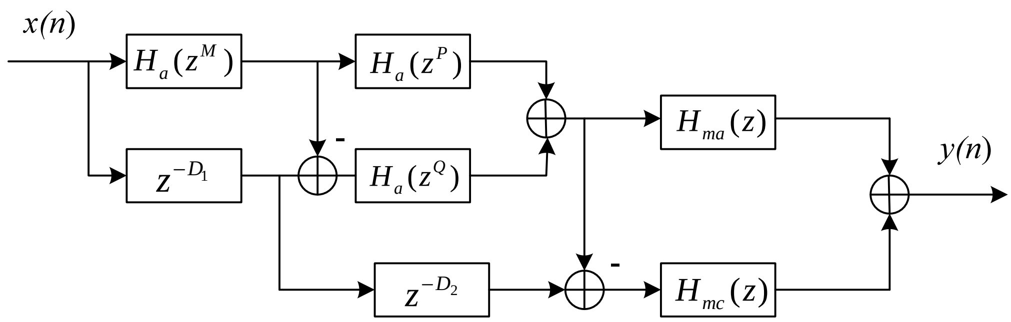


Figure 5.1: The proposed synthesis structure.

According to Fig 5.1, the zero-phase frequency response of the band-edge shaping filter $G(z)$ can be given by (5.2)

$$G(\omega) = H_a(M\omega)H_a(P\omega) + [1 - H_a(M\omega)]H_a(Q\omega). \tag{5.2}$$

The parameters required to be determined are as follows.

1)The passband and stopband edges, θ_a and ϕ_a, of the prototype filter $H_a(z)$.

2)The interpolation factors M, P and Q.

3)The passband and stopband edges, ω_{pma} and ω_{sma}, of masking filter $H_{ma}(z)$.

4)The passband and stopband edges , ω_{pmc} and ω_{smc}, of masking filter $H_{mc}(z)$.

Similar to original FRM design, there are two different design cases, i.e. Case A and Case B. In a Case A design, one of the transition bands of $H_a(z^M)$ forms the transition band of the overall filter. In a Case B design,the complement of $H_a(z^M)$ forms the transition band of the overall filter.

It should be noted that sometimes instead of $H_a(z^P)$ and $H_a(z^Q)$, their complements are used. Define two variables $Case_P$ and $Case_Q$ as

$Case_P$ (or $Case_Q$)= A, if the original filter is used,

$Case_P$ (or $Case_Q$)= B, if the complementary filter is used.

A simple example is shown in Fig. 5.2. It can be seen from Fig. 5.2(e) that the frequency response of $G(z)$ is not periodical. Additionally, it can be estimated that P and Q cannot be very large numbers since large P and Q leads to small

d_1 and d_2, shown in Fig. 5.2(e), which results in long lengths of $H_{ma}(z)$ and $H_{mc}(z)$.

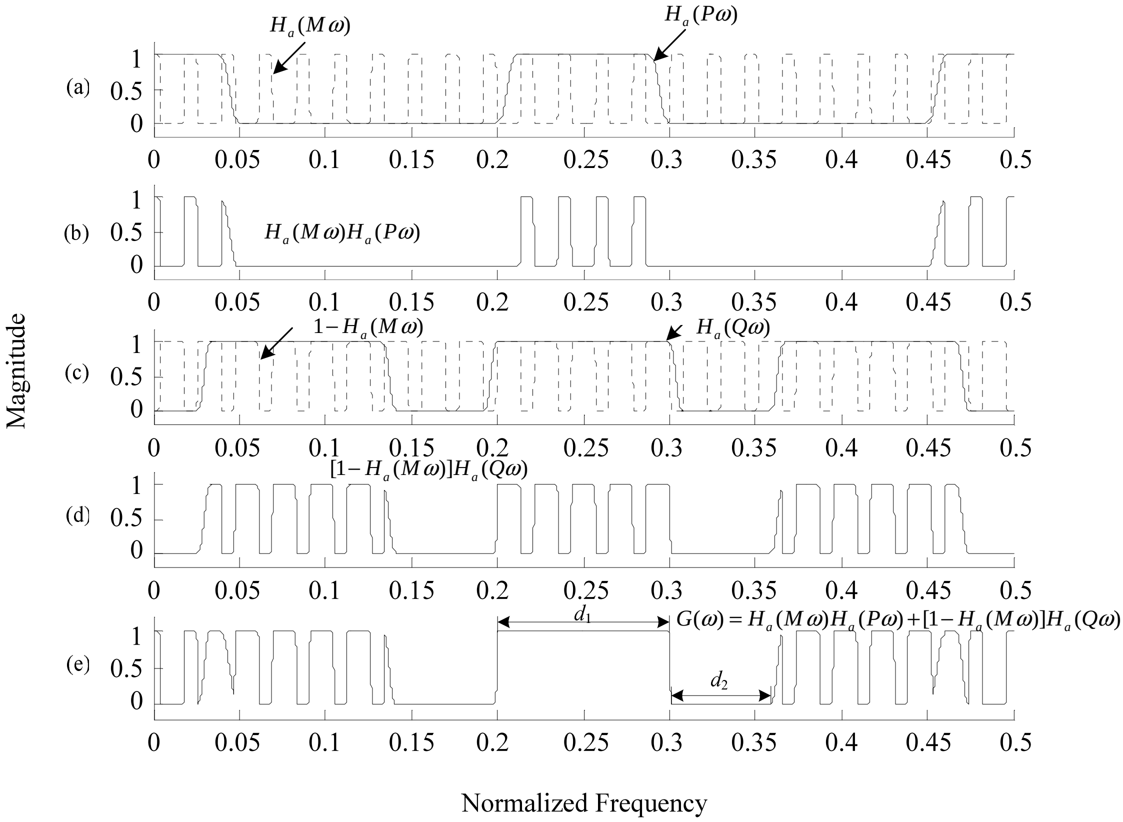


Figure 5.2: The process of synthesizing the band-edge shaping filter $G(z)$.

5.3 Design Procedure

Suppose the passband and stopband edges of the overall filter $H(z)$ are ω_p and ω_s respectively. The passband and stopband edges of the prototype filter $H_a(z)$, θ_a and ϕ_a can be found by the following equations [16].

For Case A,

$$m = \lfloor \omega_p M / 2\pi \rfloor, \tag{5.3}$$

$$\theta_a = \omega_p M - 2m\pi, \tag{5.4}$$

$$\phi_a = \omega_s M - 2m\pi. \tag{5.5}$$

For Case B,

$$m = \lceil \omega_s M / 2\pi \rceil, \tag{5.6}$$

$$\theta_a = 2m\pi - \omega_s M, \tag{5.7}$$

$$\phi_a = 2m\pi - \omega_p M, \tag{5.8}$$

where M is the interpolation factor. $\lfloor \omega_p M/2\pi \rfloor$ is the largest integer smaller than $\omega_p M/2\pi$, and $\lceil \omega_s M/2\pi \rceil$ is the smallest integer larger than $\omega_s M/2\pi$.

Only one of the two sets satisfies the condition $0 < \theta_a < \phi_a < \pi$.

5.3.1 Determination of Interpolation Factors M, P and Q

The goal is to find a particular set of parameters [M, P, Q] which leads to a design with the lowest complexity. The parameters are obtained by an exhaustive search. For a given M, P and Q should satisfy several conditions in order to minimize the complexity. These conditions are discussed for Case A and Case B respectively.

A. Case A design

Let us discuss the determination of P first. The frequency responses of the sub-filters in the upper branch of the synthesis structure for $G(z)$ is shown in Fig. 5.3. The shadowed passband is used to form the transition band of the overall filter . Six points ω_1, ω_2, ω_3, x_1, x_2, and x_3 are important to derive the conditions that determine the value of P. To keep the complete passband which forms the transition band of overall filter, inequalities (5.9) and (5.10) should be satisfied. To remove at least part of the next unwanted passband, inequality (5.11) should be satisfied.

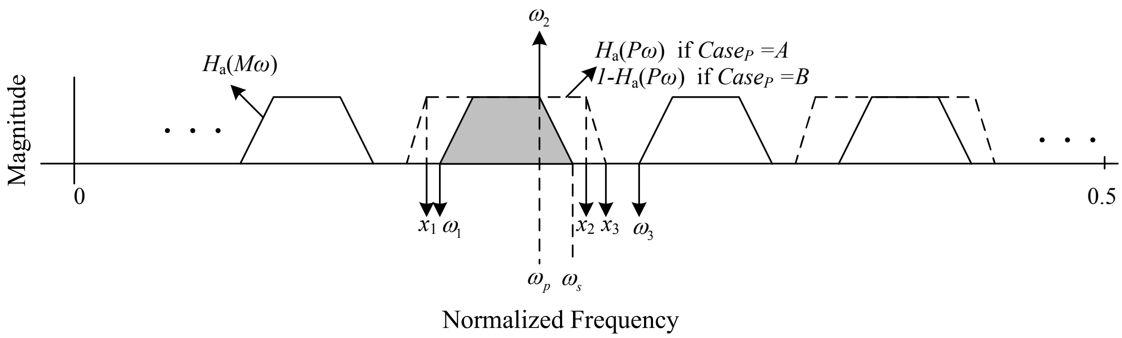


Figure 5.3: Frequency responses of the sub-filters in the upper branch of the synthesis structure for $G(z)$, $Case = A$.

The desired P is the smallest integer which satisfies the the following inequations.

$$x_1 \le \omega_1, \tag{5.9}$$

$$x_2 \geq \omega_2, \tag{5.10}$$

$$x_3 \leq \omega_3, \tag{5.11}$$

where the values of ω_1, ω_2, and ω_3 are given by (5.12) to (5.14),

$$\omega_1 = \frac{2\pi m - \phi_a}{M}, \tag{5.12}$$

$$\omega_2 = \omega_s, \tag{5.13}$$

$$\omega_3 = \frac{2\pi(m+1) - \phi_a}{M}, \tag{5.14}$$

and the values of x_1, x_2, and x_3 are different for $Case_P = A$ and $Case_P = B$, as shown in Equations (5.15) to (5.17).

$$x_1 = \begin{cases} \dfrac{2\pi p - \theta_a}{P} & , Case_P = A, \\ \dfrac{2\pi(p-1) + \phi_a}{P} & , Case_P = B, \end{cases} \tag{5.15}$$

$$x_2 = \begin{cases} \dfrac{2\pi p + \theta_a}{P} & , Case_P = A, \\ \dfrac{2\pi p - \phi_a}{P} & , Case_P = B, \end{cases} \tag{5.16}$$

$$x_3 = \begin{cases} \dfrac{2\pi p + \phi_a}{P} & , Case_P = A, \\ \dfrac{2\pi p - \theta_a}{P} & , Case_P = B. \end{cases} \tag{5.17}$$

The way to determine the interpolation factor Q is similar with the way to determine the interpolation factor P. The frequency responses of the sub-filters in the lower branch of the synthesis structure for $G(z)$ is shown in Fig. 5.4. The important points are ω_4, ω_5, ω_6, ω_7, y_1, y_2, y_3 and y_4. To keep the first passband left to ω_p, inequalities (5.18) and (5.19) should be satisfied. To remove the first unwanted passband right to ω_p completely, inequalities (5.20) and (5.21) should be satisfied.

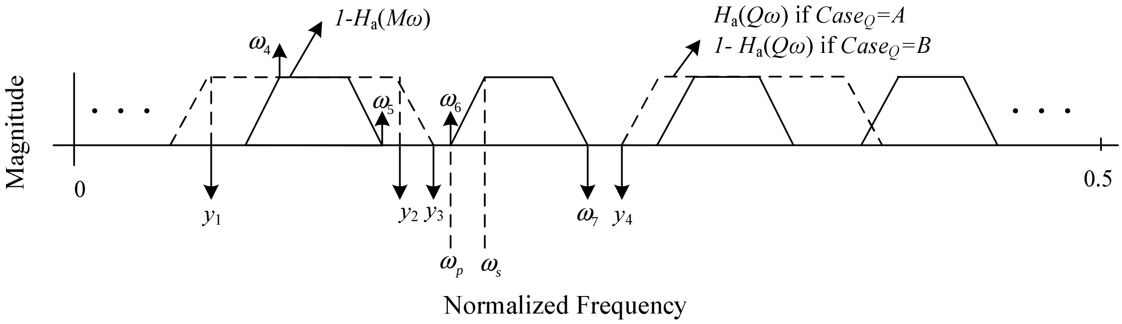


Figure 5.4: Frequency responses of the sub-filters in the lower branch of the synthesis structure for $G(z)$, $Case = A$.

The desired Q is the smallest integer which satisfies the following inequalities.

$$y_1 \leq \omega_4, \tag{5.18}$$

$$y_2 \geq \omega_5, \tag{5.19}$$

$$y_3 \leq \omega_6, \tag{5.20}$$

$$y_4 \geq \omega_7, \tag{5.21}$$

where the values of $\omega_i, i = 4, \cdots, 7$ can be obtained according to (5.22) to (5.25),

$$\omega_4 = \frac{2\pi(m-1)+\phi_a}{M}, \tag{5.22}$$

$$\omega_5 = \frac{2\pi m-\theta_a}{M}, \tag{5.23}$$

$$\omega_6 = \omega_p, \tag{5.24}$$

$$\omega_7 = \frac{2\pi(m+1)-\theta_a}{M}, \tag{5.25}$$

and the values of $y_i, i = 1, \cdots, 4$ are shown in (5.26) to (5.29),

$$y_1 = \begin{cases} \dfrac{2\pi q-\theta_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi(q-1)+\phi_a}{Q} & , Case_Q = B, \end{cases} \tag{5.26}$$

$$y_2 = \begin{cases} \dfrac{2\pi q+\theta_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi q-\phi_a}{Q} & , Case_Q = B, \end{cases} \tag{5.27}$$

$$y_3 = \begin{cases} \dfrac{2\pi q+\phi_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi q-\theta_a}{Q} & , Case_Q = B, \end{cases} \tag{5.28}$$

$$y_4 = \begin{cases} \dfrac{2\pi(q+1)-\phi_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi q+\theta_a}{Q} & , Case_Q = B. \end{cases} \tag{5.29}$$

B. Case B design

For a Case B design, the complement of $H_a(z^M)$ forms the transition band of the overall filter. Therefore, the roles of $H_a(z^P)$ and $H_a(z^Q)$ are exchanged. P is the smallest integer which can satisfy inequalities (5.18), (5.19),(5.20) and (5.21),

$$\omega_4 = \frac{2\pi(m-1)-\theta_a}{M}, \tag{5.30}$$

$$\omega_5 = \frac{2\pi(m-1)+\phi_a}{M}, \tag{5.31}$$

$$\omega_6 = \omega_p, \tag{5.32}$$

$$\omega_7 = \frac{2\pi m+\phi_a}{M}, \tag{5.33}$$

and

$$y_1 = \begin{cases} \dfrac{2\pi p-\theta_a}{P} & , Case_P = A, \\ \dfrac{2\pi(p-1)+\phi_a}{P} & , Case_P = B, \end{cases} \tag{5.34}$$

$$y_2 = \begin{cases} \dfrac{2\pi p+\theta_a}{P} & , Case_P = A, \\ \dfrac{2\pi p-\phi_a}{P} & , Case_P = B, \end{cases} \tag{5.35}$$

$$y_3 = \begin{cases} \dfrac{2\pi p+\phi_a}{P} & , Case_P = A, \\ \dfrac{2\pi p-\theta_a}{P} & , Case_P = B, \end{cases} \tag{5.36}$$

$$y_4 = \begin{cases} \dfrac{2\pi(p+1) - \phi_a}{P} & , Case_P = A, \\ \dfrac{2\pi p + \theta_a}{P} & , Case_P = B. \end{cases} \tag{5.37}$$

Q is the smallest integer which can satisfy inequalities (5.9), (5.10) and (5.11),

$$\omega_1 = \frac{2\pi(m-1) + \theta_a}{M}, \tag{5.38}$$

$$\omega_2 = \omega_s, \tag{5.39}$$

$$\omega_3 = \frac{2\pi m + \theta_a}{M}, \tag{5.40}$$

and

$$x_1 = \begin{cases} \dfrac{2\pi q - \theta_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi(q-1) + \phi_a}{Q} & , Case_Q = B, \end{cases} \tag{5.41}$$

$$x_2 = \begin{cases} \dfrac{2\pi q + \theta_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi q - \phi_a}{Q} & , Case_Q = B, \end{cases} \tag{5.42}$$

$$x_3 = \begin{cases} \dfrac{2\pi q + \phi_a}{Q} & , Case_Q = A, \\ \dfrac{2\pi q - \theta_a}{Q} & , Case_Q = B. \end{cases} \tag{5.43}$$

5.3.2 Determination of the Band-edges of $H_{ma}(z)$

θ_a and ϕ_a are already obtained. Also M, P and Q are determined. We can begin to determine the band-edges of $H_{ma}(z)$ and $H_{mc}(z)$. Let us first consider $H_{ma}(z)$. It is clear that ω_{pma} equals to ω_p and only ω_{sma} needs to be determined.

A. Case A design

The parameters used to calculate ω_{sma} for Case A design is shown in Fig. 5.5.

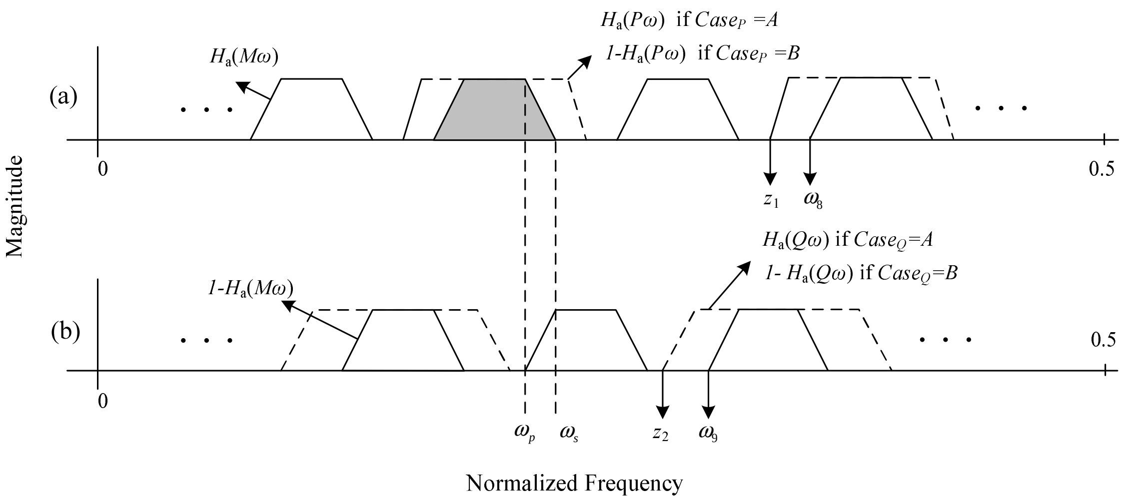


Figure 5.5: Illustration of the process to determine the band-edges of $H_{ma}(z)$ for Case A design.

The values of z_1 and z_2 are

$$z_1 = \begin{cases} \dfrac{2\pi(p+1) - \phi_a}{P} & , Case_P = A \\ \dfrac{2\pi p + \theta_a}{P} & , Case_P = B \end{cases} \tag{5.44}$$

$$z_2 = \begin{cases} \dfrac{2\pi(q+1) - \phi_a}{Q} & , Case_Q = A \\ \dfrac{2\pi q + \theta_a}{Q} & , Case_Q = B \end{cases} \tag{5.45}$$

Step 1

For the upper branch, find if z_1 is located in any of the stopband region R_k. R_k is the stopbands given by (5.46)

$$R_k = \left[\frac{2\pi(k-1) + \phi_a}{M}, \frac{2\pi k - \phi_a}{M}\right], \tag{5.46}$$

where $k = m+1, \cdots, \lfloor M/2 \rfloor$.

Define ω_{sma_up} as

$$\omega_{sma\_up} = \begin{cases} \omega_8 & , z_1 \in R_{k_0}, \\ z_1 & , z_1 \notin R, \end{cases} \tag{5.47}$$

where $R = \bigcup_k R_k$, and

$$\omega_8 = \frac{2\pi k_0 - \phi_a}{M}. \tag{5.48}$$

Step 2

For the lower branch, find if z_2 is located in any of the stopband region R_l. R_l is given by (5.49)

$$R_l = \left[\frac{2\pi l - \theta_a}{M}, \frac{2\pi l + \theta_a}{M}\right], \tag{5.49}$$

where $l = m + 1, \cdots, \lfloor M/2 \rfloor$.

Define ω_{sma_dn} as

$$\omega_{sma\_dn} = \begin{cases} \omega_9 & , z_2 \in R_{l_0}, \\ z_2 & , z_2 \notin R, \end{cases} \tag{5.50}$$

where$R = \bigcup_l R_l$, and

$$\omega_9 = \frac{2\pi l_0 + \theta_a}{M}. \tag{5.51}$$

Step 3

Compare ω_{sma_up} and ω_{sma_dn}. The stopband edge of $H_{ma}(z)$ is the smaller one of the two variables.

$$\omega_{sma} = \min(\omega_{sma\_up}, \omega_{sma\_dn}) \tag{5.52}$$

B. Case B design

The process for Case B design is the same with that for Case A design. The only difference is the values of the variables, which are listed below.

$$z_1 = \begin{cases} \dfrac{2\pi(q+1) - \phi_a}{Q} & , Case_Q = A \\ \dfrac{2\pi q + \theta_a}{Q} & , Case_Q = B \end{cases} \tag{5.53}$$

$$z_2 = \begin{cases} \dfrac{2\pi(p+1)-\phi_a}{P} & , Case_P = A \\ \dfrac{2\pi p+\theta_a}{P} & , Case_P = B \end{cases} \tag{5.54}$$

$$R_k = \left[\frac{2\pi k-\theta_a}{M}, \frac{2\pi k+\theta_a}{M}\right], k = m+1, \cdots, \lfloor M/2 \rfloor \tag{5.55}$$

$$R_l = \left[\frac{2\pi(l-1)+\phi_a}{M}, \frac{2\pi l-\phi_a}{M}\right], l = m+1, \cdots, \lfloor M/2 \rfloor \tag{5.56}$$

$$\omega_8 = \frac{2\pi k_0+\theta_a}{M} \tag{5.57}$$

$$\omega_9 = \frac{2\pi l_0-\phi_a}{M} \tag{5.58}$$

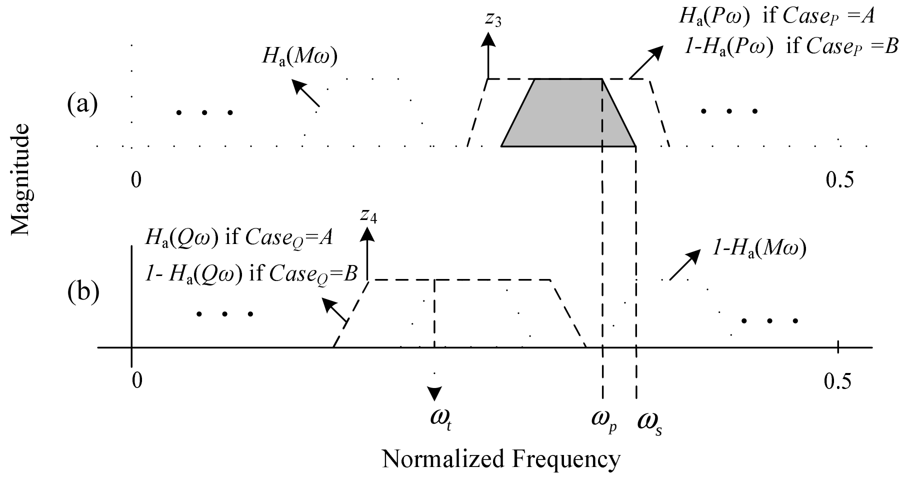
5.3.3 Determination of the Bandedges of $H_{mc}(z)$

For masking filter $H_{mc}(z)$, ω_{smc} equals to ω_s and the passband edge, ω_{pmc}, needs to be determined.

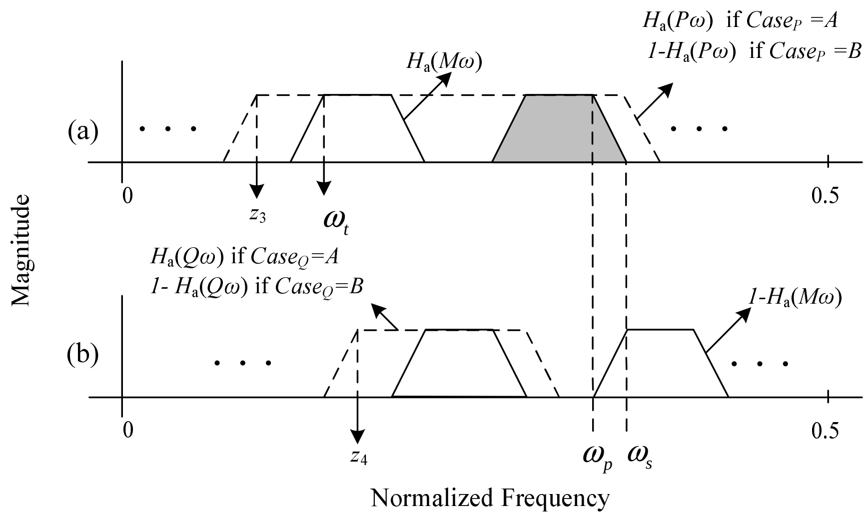
A. Case A design

The parameters for Case A design is shown in Fig.5.6. There are two situations to be considered according to the position of z_3 and z_4.

The values of z_3 and z_4 are given by



(a) $z_3 \geq z_4$



(b) $z_3 < z_4$

Figure 5.6: Illustration of the process to determine the band-edges of $H_{mc}(z)$ for Case A design.

$$z_3 = \begin{cases} \dfrac{2\pi p - \theta_a}{P} & , Case_P = A \\ \dfrac{2\pi(p-1) + \phi_a}{P} & , Case_P = B, \end{cases} \tag{5.59}$$

$$z_4 = \begin{cases} \dfrac{2\pi q - \theta_a}{Q} & , Case_Q = A \\ \dfrac{2\pi(q-1) + \phi_a}{Q} & , Case_Q = B. \end{cases} \tag{5.60}$$

When $z_3 \geq z_4$, find if z_3 is located in any of the stopbands $R_k = \left[\frac{2\pi(k-1)+\phi_a}{M}, \frac{2\pi k-\phi_a}{M}\right], k = 1, \cdots, m$. If $z_3 \in R_{k_0}$, z_4 is compared with ω_t. ω_{pmc} equals to the larger one of z_4 and ω_t. If z_3 is not located in any of the stopbands, ω_{pmc} equals to z_3.

Similarly, when $z_3 < z_4$, find if z_4 is located in any of the stopbands $R_l = \left[\frac{2\pi l-\theta_a}{M}, \frac{2\pi l+\theta_a}{M}\right], l = 0, \cdots, m-1$. If $z_4 \in R_{l_0}$, ω_t becomes important. ω_{pmc} equals to the larger one of z_3 and ω_t. If z_4 is not located in any of the stopbands, ω_{pmc} equals to z_4.

In conclusion, when $z_3 \geq z_4$,

$$\omega_{pmc} = \begin{cases} \max(z_4, \omega_t) & , z_3 \in R_{k_0} \\ z_3 & , z_3 \notin R, \end{cases} \tag{5.61}$$

where$R = \bigcup_k R_k$, and

$$\omega_t = \frac{2\pi(k_0 - 1) + \phi_a}{M}. \tag{5.62}$$

and when $z_3 < z_4$

$$\omega_{pmc} = \begin{cases} \max(z_3, \omega_t) & , z_4 \in R_{l_0} \\ z_4 & , z_4 \notin R, \end{cases} \tag{5.63}$$

where $R = \bigcup_l R_l$, and

$$\omega_t = \frac{2\pi l_0 - \theta_a}{M}. \tag{5.64}$$

B. Case B design

The way to find ω_{pmc} is the same with that of Case A design. The values of z_3 and z_4 are given by

$$z_3 = \begin{cases} \dfrac{2\pi q - \theta_a}{Q} & , Case_Q = A \\ \dfrac{2\pi(q-1) + \phi_a}{Q} & , Case_Q = B, \end{cases} \tag{5.65}$$

$$z_4 = \begin{cases} \dfrac{2\pi p - \theta_a}{P} & , Case_P = A \\ \dfrac{2\pi(p-1) + \phi_a}{P} & , Case_P = B. \end{cases} \tag{5.66}$$

when $z_3 \geq z_4$,

$$R_k = \left[\frac{2\pi k - \theta_a}{M}, \frac{2\pi k + \theta_a}{M}\right], k = 0, \cdots, m-1, \tag{5.67}$$

$$\omega_t = \frac{2\pi k_0 - \theta_a}{M}. \tag{5.68}$$

and when $z_3 < z_4$

$$R_l = \left[\frac{2\pi(l-1) + \phi_a}{M}, \frac{2\pi l - \phi_a}{M}\right], l = 1, \cdots, m, \tag{5.69}$$

$$\omega_t = \frac{2\pi(l_0 - 1) + \phi_a}{M}. \tag{5.70}$$

5.4 Ripple Analysis

To set the ripples of different sub-filters, it is helpful to have some understanding on the ripple of the overall filter. The ideal frequency responses of the overall filter and the two masking filters are shown in Fig. 5.7. The ripple analysis is carried on under the assumption that $Case = A$, $Case_P = A$ and $Case_Q = A$.

The frequency response of the overall filter can be written in the form of (5.71),

$$\begin{aligned} H(\omega) = \{H_a(M\omega)[H_a(P\omega) - H_a(Q\omega)] + H_a(Q\omega)\}[H_{ma}(\omega) - H_{mc}(\omega)] \\ + H_{mc}(\omega). \end{aligned} \tag{5.71}$$

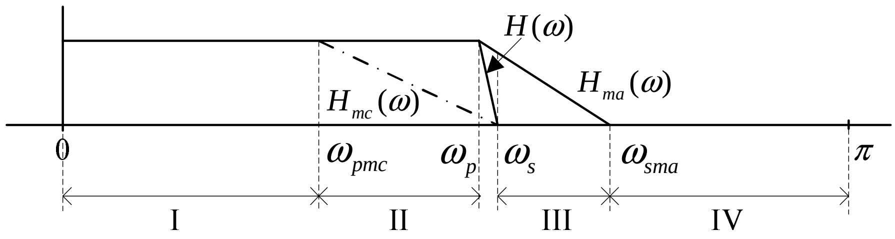


Figure 5.7: The ideal frequency response of the overall filter and the two masking filters.

$H(\omega)$ can be represented using (5.72),

$$H(\omega) = G(\omega) + d(\omega), \tag{5.72}$$

where $G(\omega)$ is the ideal frequency response and $d(\omega)$ is the ripple. The ripple in the transition band are assumed to be 0.

Similarly, the frequency responses of $H_a(z^M)$, $H_a(z^P)$, $H_a(z^Q)$, $H_{ma}(z)$ and $H_{mc}(z)$ can be written as (5.73)-(5.77).

$$H_a(M\omega) = G_M(\omega) + d_M(\omega), \tag{5.73}$$

$$H_a(P\omega) = G_P(\omega) + d_P(\omega), \tag{5.74}$$

$$H_a(Q\omega) = G_Q(\omega) + d_Q(\omega), \tag{5.75}$$

$$H_{ma}(\omega) = G_{ma}(\omega) + d_{ma}(\omega), \tag{5.76}$$

$$H_{mc}(\omega) = G_{mc}(\omega) + d_{mc}(\omega). \tag{5.77}$$

Applying Equations from (5.73) to (5.77) to Equation (5.72), it becomes

$$\begin{aligned} G(\omega)+d(\omega) = \{&(G_M(\omega)+d_M(\omega))[(G_P(\omega)+d_P(\omega))-(G_Q(\omega)+d_Q(\omega))] \\ &+(G_Q(\omega)+d_Q(\omega))\}\cdot[(G_{ma}(\omega)+d_{ma}(\omega))-(G_{mc}(\omega)+d_{mc}(\omega))] \\ &+(G_{mc}(\omega)+d_{mc}(\omega)). \end{aligned} \tag{5.78}$$

After some manipulations and by omitting the second order items, (5.78) results in (5.79),

$$\begin{aligned} d(\omega) =\ & d_M(\omega)\cdot[G_P(\omega)-G_Q(\omega)][G_{ma}(\omega)-G_{mc}(\omega)] \\ &+d_P(\omega)\cdot G_M(\omega)[G_{ma}(\omega)-G_{mc}(\omega)] \\ &+d_Q(\omega)\cdot[1-G_M(\omega)][G_{ma}(\omega)-G_{mc}(\omega)] \\ &+d_{ma}(\omega)\cdot[G_M(\omega)G_P(\omega)+(1-G_M(\omega))G_Q(\omega)] \\ &+d_{mc}(\omega)\cdot[1-G_M(\omega)G_P(\omega)-(1-G_M(\omega))G_Q(\omega)] \\ &+[G_M(\omega)G_P(\omega)+(1-G_M(\omega))G_Q(\omega)][G_{ma}(\omega)-G_{mc}(\omega)] \\ &+G_{mc}(\omega)-G(\omega). \end{aligned} \tag{5.79}$$

Now ripple analysis can be carried on according to different regions from I to IV based on (5.79).

A. Region I: $\omega \leqslant \omega_{pmc}$

In this region, $G(\omega)=G_{ma}(\omega)=G_{mc}(\omega)=1$, that is $G_{ma}(\omega)-G_{mc}(\omega)=0$ and

$G_{mc}(\omega) - G(\omega) = 0$. (5.79) becomes

$$d(\omega) = Ad_{ma}(\omega) + (1 - A)d_{mc}(\omega), \tag{5.80}$$

where

$$A = G_M(\omega)G_P(\omega) + (1 - G_M(\omega))G_Q(\omega). \tag{5.81}$$

Because $G_M(\omega)$, $G_P(\omega)$, and $G_Q(\omega)$ are all between 0 and 1, there are

$$0 \leqslant A \leqslant \max(G_P(\omega), G_Q(\omega)) = 1. \tag{5.82}$$

Therefore, (5.80) becomes

$$|d(\omega)| \leqslant \max(|d_{ma}(\omega)|, |d_{mc}(\omega)|). \tag{5.83}$$

It can be seen from (5.83) that in this region, the ripple of the overall filter is up-limited by the maximum value of the ripples of the masking filters.

B. Region II : $\quad \omega_{pmc} < \omega \leqslant \omega_p$

In this region, there are several conditions :

1) $G(\omega) = G_{ma}(\omega) = 1$

2) $d_{mc}(\omega) = 0$

3) When $G_M(\omega) = 1$, $G_P(\omega) = 1$. When $G_M(\omega) = 0$, $G_Q(\omega) = 1$. When $0 <$

$G_M(\omega) < 1$, $G_P(\omega) = G_Q(\omega) = 1$.

Condition 3) implies that $G_M(\omega)G_P(\omega) + (1 - G_M(\omega))G_Q(\omega) = 1$. Considering the three conditions, (5.79) becomes (5.84)

$$\begin{aligned} d(\omega) = d_M(\omega) \cdot [G_P(\omega) - G_Q(\omega)][1 - G_{mc}(\omega)] \\ + [d_P(\omega) \cdot G_M(\omega) + d_Q(\omega)(1 - G_M(\omega))][1 - G_{mc}(\omega)] \\ + d_{ma}. \end{aligned} \tag{5.84}$$

It is worth noting that

$$0 < |G_P(\omega) - G_Q(\omega)| < 1,$$

$$0 < (1 - G_{mc}(\omega)) < 1,$$

$$0 < G_M(\omega) < 1,$$

$$0 < (1 - G_M(\omega)) < 1.$$

Therefore, (5.84) results in

$$|d(\omega)| \leqslant [1 - G_{mc}(\omega)] \{|d_M(\omega)| + \max(|d_P(\omega)|, |d_Q(\omega)|)\} + |d_{ma}(\omega)|. \tag{5.85}$$

It can be seen from (5.85) that in this region, the ripple is larger than that of region I. Since $G_{mc}(\omega)$ keeps decreasing from 1 to 0 as the frequency increases in this region, there are large ripples near the passband edge ω_p.

C. Region III: $\omega_s \leqslant \omega < \omega_{sma}$

In this region, the conditions are as follows.

1) $G(\omega) = G_{mc}(\omega) = 0$.

2) $d_{ma}(\omega) = 0$.

3) When $G_M(\omega) = 1$, $G_P(\omega) = 0$. When $G_M(\omega) = 0$, $G_Q(\omega) = 0$. When $0 < G_M(\omega) < 1$, $G_P(\omega) = G_Q(\omega) = 0$.

Condition 3) implies that $G_M(\omega)G_P(\omega) + (1 - G_M(\omega))G_Q(\omega) = 0$. Considering the above conditions, (5.79) becomes (5.86)

$$\begin{aligned} d(\omega) = & d_M(\omega) \cdot [G_P(\omega) - G_Q(\omega)]G_{ma}(\omega) \\ & + [d_P(\omega) \cdot G_M(\omega) + d_Q(\omega) \cdot (1 - G_M(\omega))]G_{ma}(\omega) \\ & + d_{mc}(\omega). \end{aligned} \tag{5.86}$$

Therefore

$$|d(\omega)| \leqslant G_{ma}(\omega)\{|d_M(\omega)| + \max(|d_P(\omega)|, |d_Q(\omega)|)\} + |d_{mc}(\omega)|. \tag{5.87}$$

In this region, $G_{ma}(\omega)$ keeps decreasing from 1 to 0 as frequency increases. There are large ripples near the stopband edge ω_s.

D. Region IV: $\omega_{sma} \leqslant \omega < \pi$

In this region, $G(\omega) = G_{ma}(\omega) = G_{mc}(\omega) = 0$. The expressions are the same as Equation (5.80) to (5.83). In this region, the ripple of the overall filter is up-limited by the maximum value of the ripples of the masking filters.

In conclusion, in region I and IV the ripple of the overall filter is determined by the maximum value of the ripples of the masking filters. In region II and III, The ripple of the overall filter is influenced by several sub-filters and becomes larger than that of region I and IV. And the largest ripple occurs at the region near the band edges of the overall filter. According to (5.85) and (5.87), it is safe to set the ripple of $H_a(z)$ be

$$\delta_a = \frac{\delta_d - \max(\delta_{ma}, \delta_{mc})}{2}, \tag{5.88}$$

where δ_d is the desired ripple of the overall filter and δ_{ma} and δ_{mc} are the ripples of the two masking filters.

5.5 Implementation of the Scheme

The proposed structure makes use of $H_a(z)$ several times to form the band-edge shaping filter. It is necessary to find an implementation structure where multipliers can be reused. If we rewrite (5.2) to (5.89), it is possible to share the multipliers.

$$G(z) = H_a(z^M)[H_a(z^P) - H_a(z^Q)] + H_a(z^Q) \tag{5.89}$$

A simple FIR filter in Fig. 5.8(a) consists of two parts, multipliers, represented by block "M" in Fig. 5.8(b) and delay elements and adders, represented by block "$D\&A$" in Fig. 5.8(b).

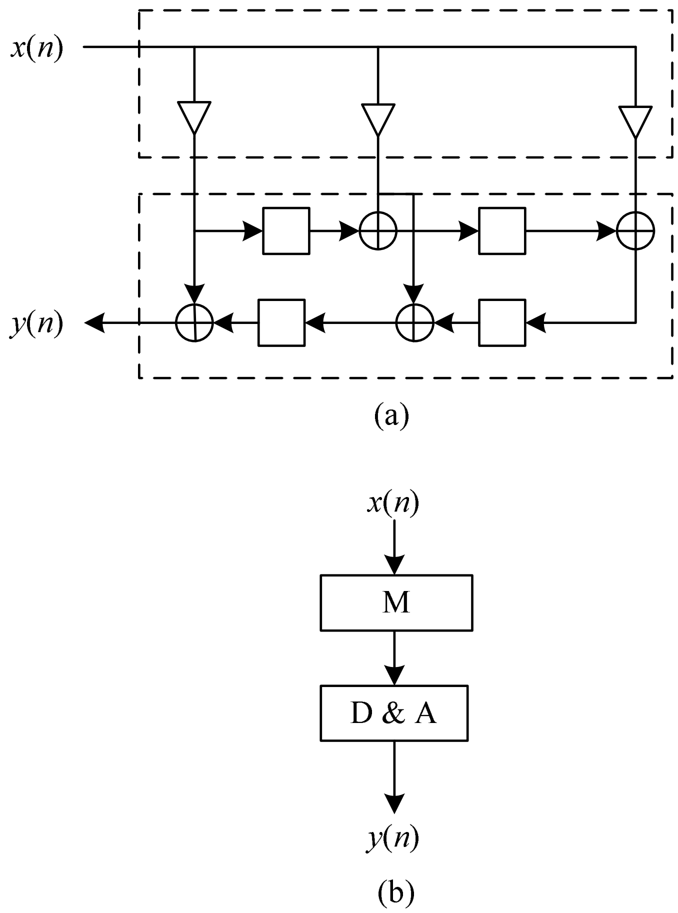


Figure 5.8: Two-part structure of a FIR filter.

Equation (5.89) can be realized using the structure in Fig. 5.9. Block "D" is a delay block. The numbers in the brackets in the "$D\&A$" blocks represent the interpolation factors. Firstly switch 1 is closed. The input signal passes through two systems with transfer functions $H_a(z^P)$ and $H_a(z^Q)$ respectively. The difference of the two outputs is kept in a latch T. The output of $H_a(z^Q)$ is delayed to wait for addition. Then switch 1 is open and switch 2 is closed. The signal in latch

T passes through a system with transfer function $H_a(z^M)$. This output and the delayed output from $H_a(z^Q)$ in the previous round are added together to get the final output signal. In this way, the multipliers can be shared.

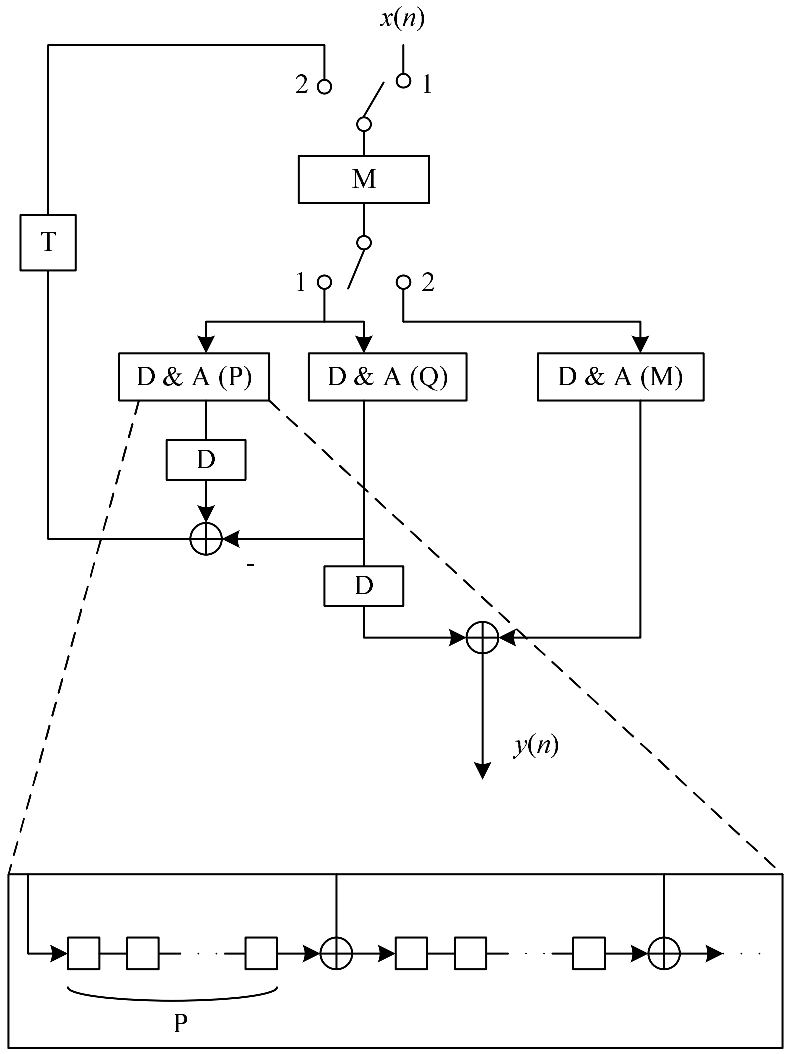


Figure 5.9: Implementation of the proposed synthesis structure (suppose $P < Q$).

5.6 A Design Example

The whole process to synthesize a FIR filter with the proposed structure can be demonstrated by the following example. Suppose the desired sharp filter has passband and stopband edge at 0.6π and 0.602π, respectively. The passband ripple is 0.01 and the stopband attenuation is 40dB. Using the proposed method, the three interpolation factors are 46, 4 and 6 respectively. The lengths of $H_a(z)$, $H_{ma}(z)$ and $H_{mc}(z)$ are 47, 37 and 23 respectively. It is a Case B design. The implementation requires 55 multipliers and 107 adders. The magnitude response of the overall filter is shown in Fig. 5.10.

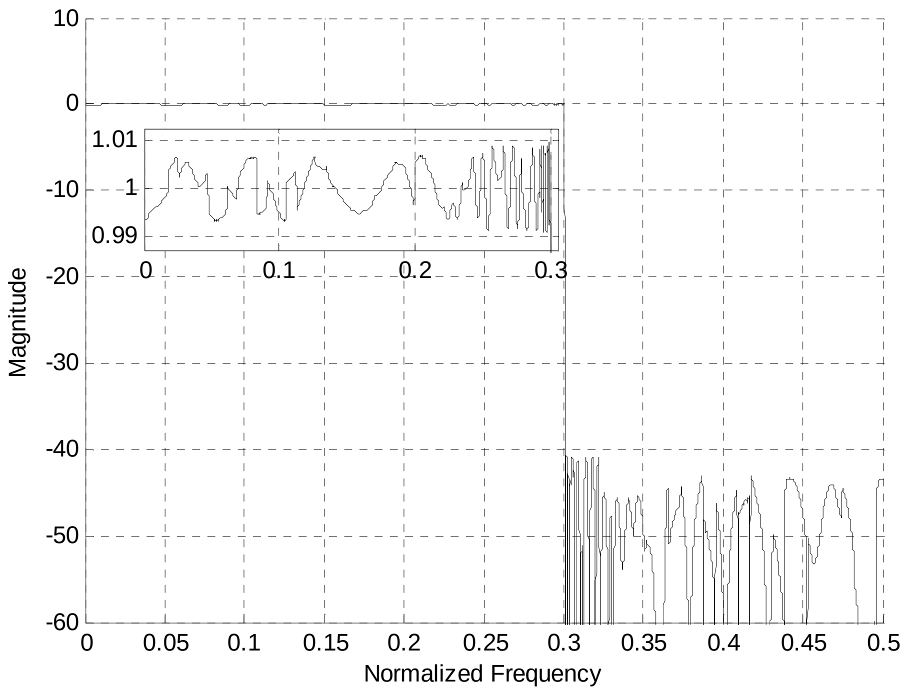


Figure 5.10: The magnitude response of the overall filter.

For comparison, different FRM designs are summarized in Table 5.1. It can be seen from Table 5.1 that the proposed method achieves more than 57% and 40% savings in terms of the number of multipliers compared to the traditional 1-stage and 2-stage FRM method. The proposed method also achieves more than 34% savings in terms of the number of multipliers and more than 26% savings in group delay compared to SFFM method.

Table 5.1: Comparison of different design methods.

Designs	Multipliers	Adders	Group Delays
conventional	925	1850	925
1-stage FRM	130	256	1000
2-stage FRM	92	177	1105
IFIR-FRM	104	202	1084
SFFM case A	86	170	1638.5
SFFM case B	84	165	2150.5
Proposed	55	107	1214

5.7 Extended Structure

In case of designing filters with very narrow transition bands, the proposed structure can be extended to the structure shown in Fig. 5.11. In the figure, $P_i > P_{i+1}$ and $Q_j > Q_{j+1}$. The reduction of complexity is at the expense of delay.

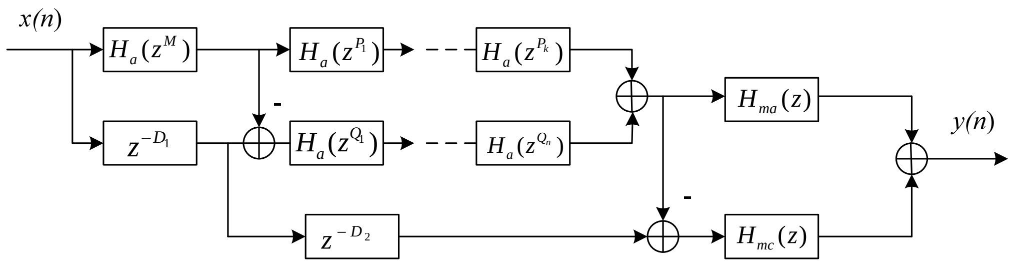


Figure 5.11: The extended structure of the proposed structure.

5.8 Summary

In this chapter, a low complexity design for sharp FIR filter based on FRM technique was proposed. Instead of using the interpolated filter as band-edge shaping filter, a band-edge shaping filter with non-periodical frequency response was synthesized. The multipliers of the sub-filters were shared efficiently. Simulations showed that significant savings in terms of the number of multipliers are achieved compared with the traditional 1-stage and 2-stage FRM methods. Both the complexity and the group delay are reduced compared with the SFFM method.

Chapter 6

Low Complexity Serial Masking Scheme Based on Frequency-Response Masking Approach

6.1 Introduction

One of the most successful techniques to reduce the design complexity of sharp FIR filers is frequency response masking technique, which was introduced in [16] and developed in [17][18][27]. The structure has been given in Fig. 1.3. Three sub-filters and two delay blocks are needed to synthesize a FRM filter. The band-edge shaping filter $H_a(z^M)$ and its complement are masked by two parallel masking filters respectively. Then the results of the parallel processing are combined to form the desired output.

A lot of study has been carried on to further reduce the complexity of FRM design. In this chapter, instead of using parallel masking, a synthesis structure with serial masking is proposed. The proposed structure has low complexity compared to the conventional FRM structure.

This chapter is organized as follows. In Section 6.2, the proposed structure is discussed. In Section 6.3, the design procedure is addressed. In Section 6.4, simulation results are provided to testify the effectiveness of the proposed structure. A conclusion is drawn in Section 6.5.

6.2 Proposed Synthesis Structure

Unlike the conventional FRM scheme which uses parallel masking, the proposed structure achieves its desired output by two-step masking. Since the second step masking removes all the unwanted passbands at frequencies larger than the stopband edge, the first step masking filter can be any kind of filters as long as the frequency response of the desired filter at frequencies below the passband edge is produced before the second step masking. One good candidate of the first step masking filter is an interpolated filter.

The structure is different with different stopband edges and different implementation cases. These structures are discussed separately.

A. $\omega_s < \pi/2$ & *Case A*

When the stopband edge of the desired filter is smaller than $\pi/2$ and the synthesis is a Case A design, the structure shown in Fig. 6.1 is proposed. $H_a(z)$ is the prototype filter. The first step and second step masking filters are named $H_{m1}(z)$ and $H_{m2}(z)$ respectively. D_a and D_{m1} are group delays of $H_a(z^M)$ and $H_{m1}(z)$ respectively. L_a and L_{m1} are the lengths of $H_a(z)$ and $H_{m1}(z)$, which should be odd numbers.

$$D_a = \frac{M(L_a - 1)}{2} \tag{6.1}$$

$$D_{m1} = \frac{N(L_{m1} - 1)}{2} \tag{6.2}$$

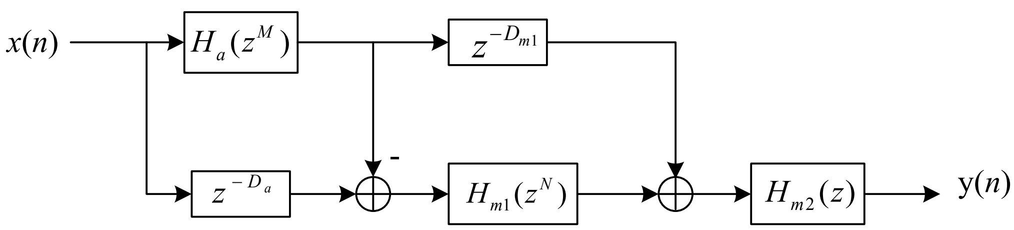


Figure 6.1: The proposed structure for $\omega_s < \pi/2$ & Case A.

The process is illustrated clearly in Fig. 6.2. First, the complement of $H_a(z^M)$ is masked by the interpolated formation of $H_{m1}(z)$ with interpolation factor N as shown in Fig. 6.2(c). Then combination of the frequency responses in Fig. 6.2(a) and (c) is further masked by $H_{m2}(z)$ as shown in Fig. 6.2(d) to get the desired frequency response. The z-transform transfer function of the overall filter can be written as

$$H(z) = \left[H_a(z^M)z^{-D_{m1}} + (z^{-D_a} - H_a(z^M))H_{m1}(z^N)\right] H_{m2}(z). \tag{6.3}$$

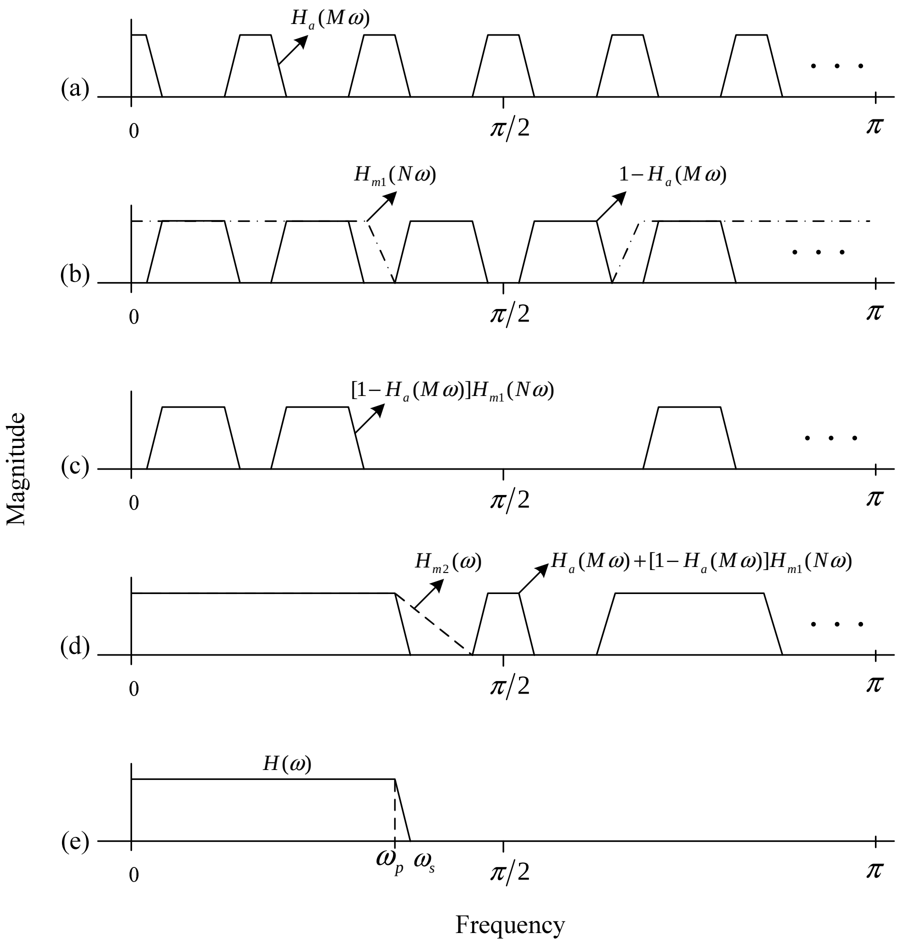


Figure 6.2: The process of obtaining the overall filter with $\omega_s < \pi/2$ & Case A.

B. $\omega_s < \pi/2$ & *Case B*

When the stopband edge is smaller than $\pi/2$ and it is a Case B design, the proposed structure is shown in Fig. 6.3. $H_a(z^M)$ is first masked by the interpolated filter of $H_{m1}(z)$ with interpolation factor N. The complement of $H_a(z^M)$ is delayed and then added to the result of the first step masking. Finally the result of the summation is masked by the second step masking filter $H_{m2}(z)$ to get the desired frequency response. The process is shown in Fig. 6.4. The z-transform transfer function of the overall filter can be written as

$$H(z) = \left[H_a(z^M)H_{m1}(z^N) + (z^{-D_a} - H_a(z^M))z^{-D_{m1}}\right] H_{m2}(z). \qquad (6.4)$$

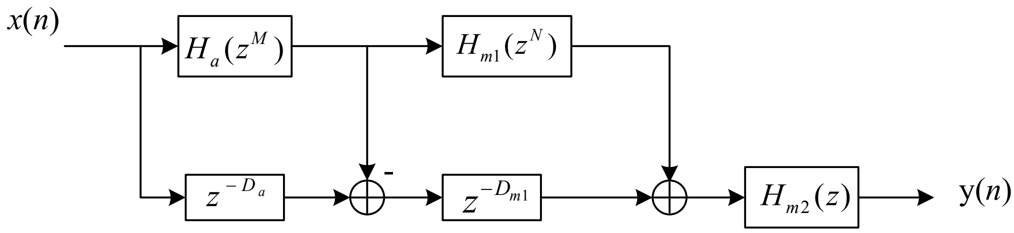


Figure 6.3: The proposed structure for $\omega_s < \pi/2$ & Case B.

C. $\omega_s > \pi/2$ & *Case A*

When the stopband edge is larger than $\pi/2$, the complement of the desired filter instead of the desired filter is synthesized. In a Case A design, the proposed structure

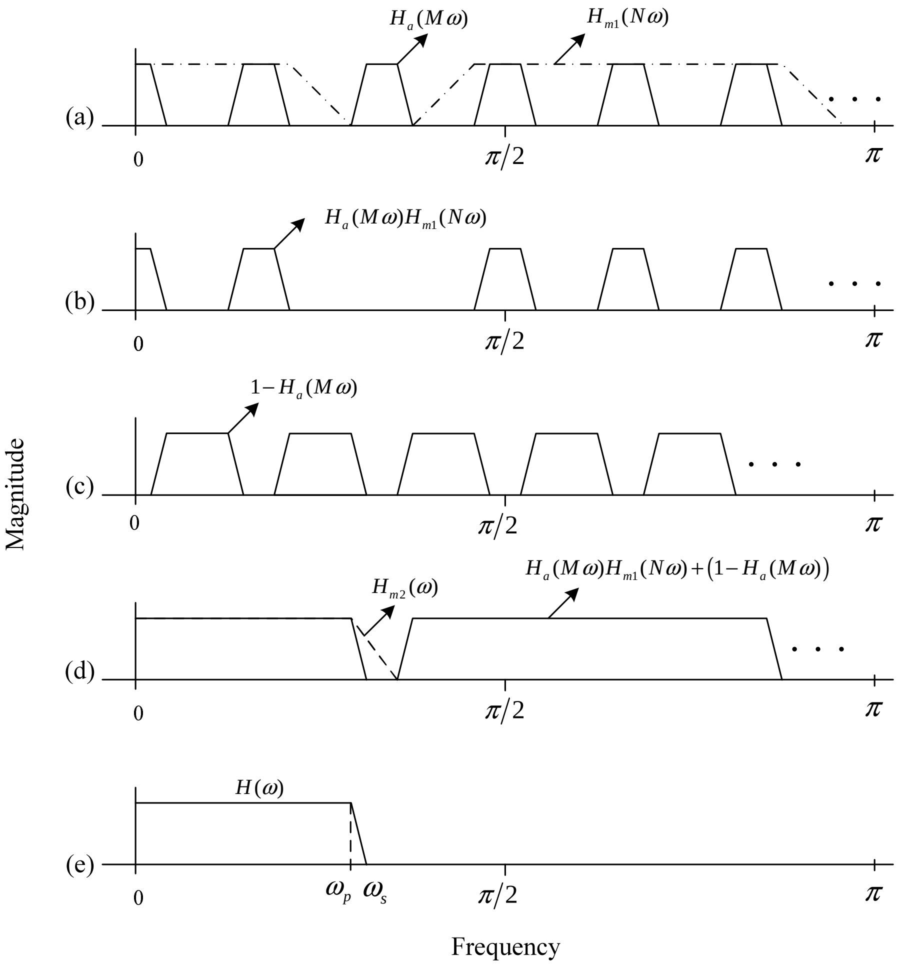


Figure 6.4: The process of obtaining the overall filter with $\omega_s < \pi/2$ & Case B.

is shown in Fig. 6.5. D_m is the group delay of the cascade of $H_{m1}(z)$ and $H_{m2}(z)$, i.e. $D = D_{m1} + D_{m2}$. D_{m1} is given by (6.2), and D_{m2} equals to $(L_{m2} - 1)/2$, where L_{m2} is the length of $H_{m2}(z)$, which should be odd in this case.

The structure is similar to the structure when $\omega_s < \pi/2$ and $Case = B$. The complement of $H_a(z^M)$ is delayed to add the cascade of $H_a(z^M)$ and $H_{m1}(z^N)$. The result is combined and then masked by $H_{m2}(z)$ to get the complement of the desired frequency response. The whole process is shown in Fig. 6.6. The z-transform transfer function of the overall filter can be written as (6.5).

$$H(z) = z^{-(D_a+D_m)} - \left[H_a(z^M)H_{m1}(z^N) + (z^{-D_a} - H_a(z^M)z^{-D_{m1}}\right] H_{m2}(z). \quad (6.5)$$

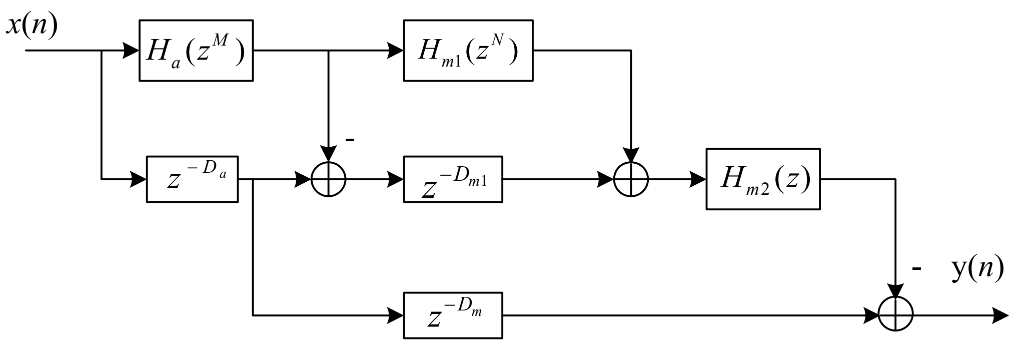


Figure 6.5: The proposed structure for $\omega_s > \pi/2$ & Case A.

D. $\omega_s > \pi/2$ & *Case B*

When $\omega_s > \pi/2$ and it is a Case B design, the proposed structure is shown in

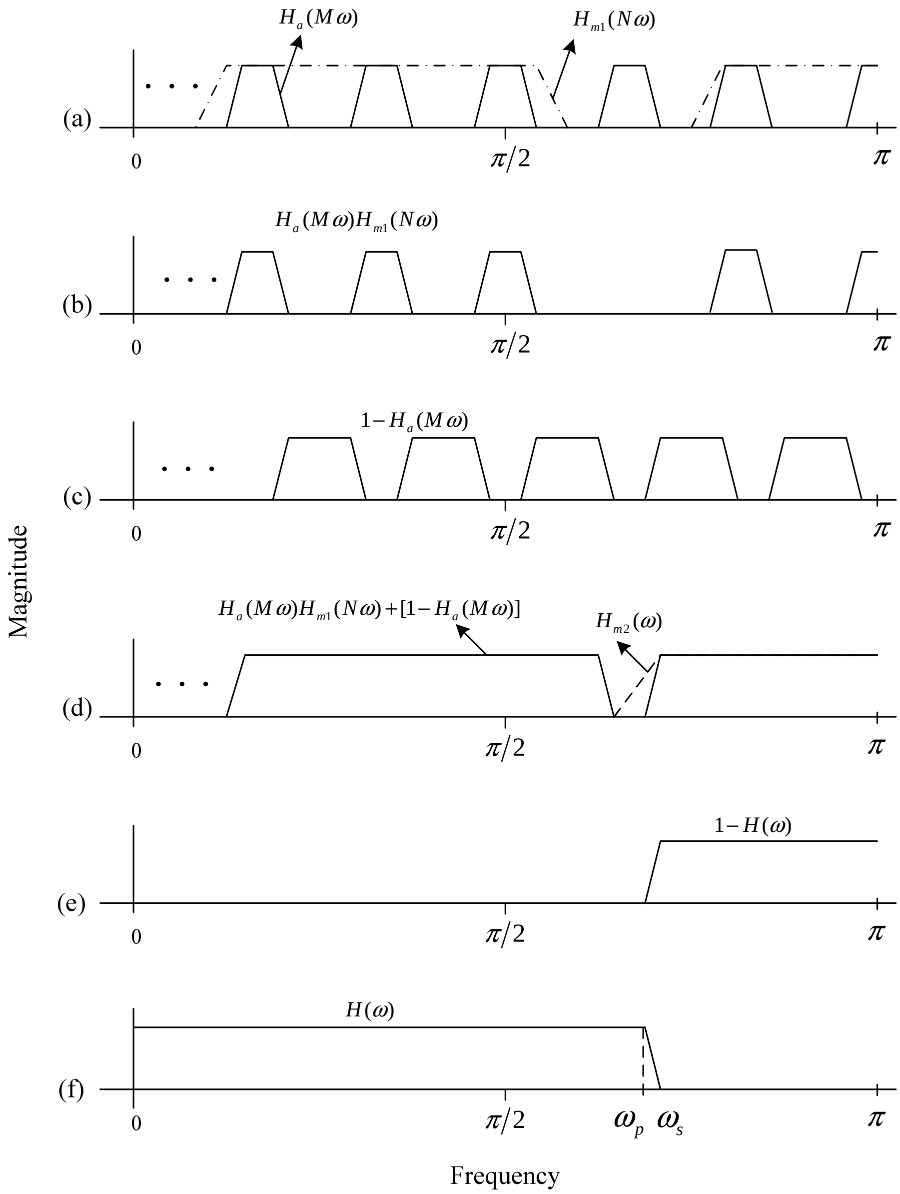


Figure 6.6: The process of obtaining the overall filter with $\omega_s > \pi/2$ & Case A.

Fig. 6.7. First, the complement of $H_a(z^M)$ is masked by the interpolated filter of $H_{m1}(z)$ with interpolation factor N. Then the sum of the fist-step masking and the delayed $H_a(z^M)$ is masked by $H_{m2}(z)$ to get the complement of the desired frequency response. The z-transform transfer function of the overall filter can be written as in (6.6). The whole process is shown in Fig. 6.8.

$$H(z) = z^{-(D_a+D_m)} - \left[H_a(z^M)z^{-D_{m1}} + (z^{-D_a} - H_a(z^M))H_{m1}(z^N)\right] H_{m2}(z). \quad (6.6)$$

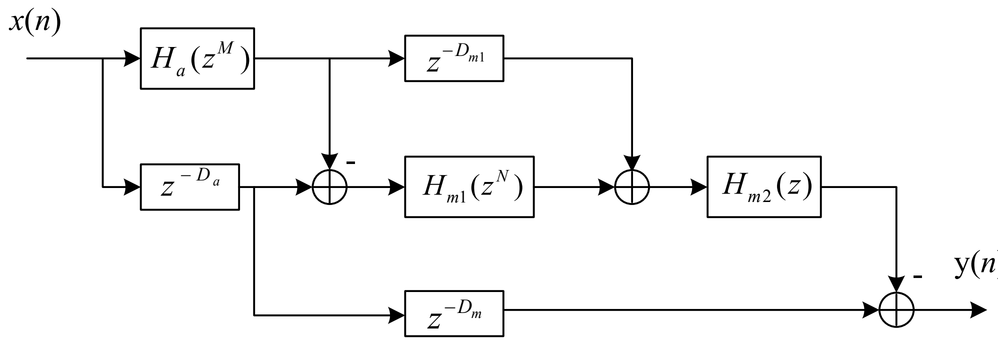


Figure 6.7: The proposed structure for $\omega_s > \pi/2$ & Case B.

There are several facts which are to be noted.

1) For all the structures, the lengths of $H_a(z)$ and $H_{m1}(z)$ should be odd.

2) The length of $H_{m2}(z)$ can be odd or even when $\omega_s < \pi/2$ and should be odd when $\omega_s > \pi/2$.

3) When $\omega_s < \pi/2$, $H_{m2}(z)$ is a lowpass filter and when $\omega_s > \pi/2$, $H_{m2}(z)$ is a highpass filter.

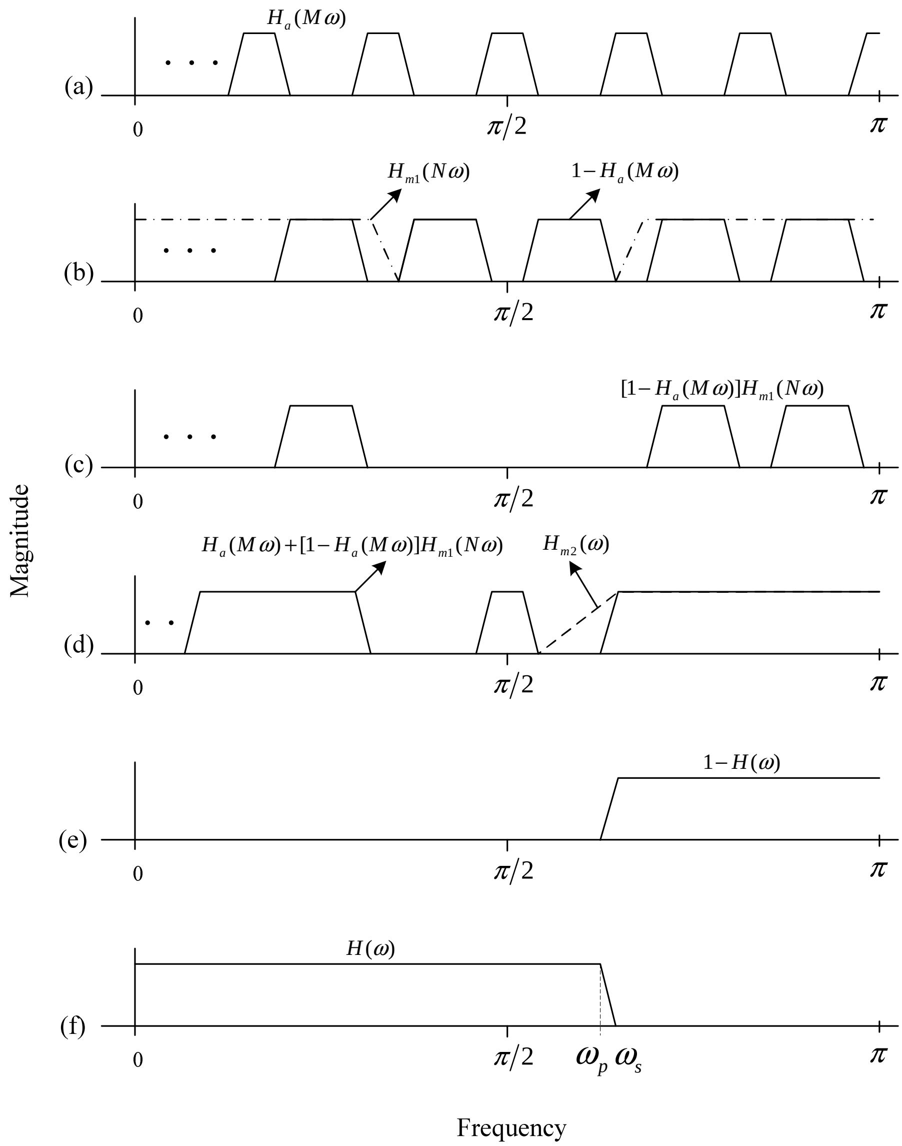


Figure 6.8: The process of obtaining the overall filter with $\omega_s > \pi/2$ & Case B.

6.3 Design Procedure

The design procedure is illustrated under the assumption of $\omega_s < \pi/2$ and $CaseA$. M is carefully chosen to make sure that the band-edge shaping filter forms the transition band of the overall filter. For a given M,

Step 1 : Determine the band-edges, θ_a and ϕ_a of $H_a(z)$. The equations used are given by (5.3) to (5.8).

step 2: Determine the band-edges, ω_{pm1} and ω_{sm1} of $H_{m1}(z)$ and the value of N. For $\omega_s < \pi/2$ & Case A, N should satisfy $N \cdot \omega_p < \pi$. For a given N, ω_{pm1} and ω_{sm1} can be obtained using (5.3) to (5.8) with $\omega'_p = (2m\pi - \theta_a)/M$ and $\omega'_s = \omega_p$ as the band-edges of the desired filter.

step 3: Determine the band-edges, ω_{pm2} and ω_{sm2} of $H_{m2}(z)$. It is clear that $\omega_{pm2} = \omega_p$ and $\omega_{sm2} = \dfrac{2(m+1)\pi - \phi_a}{M}$.

By exhaustive search, M and N are chosen to be the ones which can make the complexity of the overall filter lowest. The design procedures of others three cases are the same.

6.4 Design Examples

The design method is illustrated using two lowpass examples. The first example is a filter with stopband edge larger than $\pi/2$ and the second example is a filter with stopband edge smaller than $\pi/2$.

A. Example 1

Suppose the desired filter is a lowpass FIR filter with passband and stopband edges at 0.6π and 0.602π respectively. The passband ripple is 0.001 and the stopband attenuation is 60dB. Using the conventional FRM scheme, interpolation factor M equals to 14 and the lengths of prototype filter and two masking filters are 241, 82 and 114 respectively. 219 multipliers and 437 adders are needed. The group delay is 1736.5. Using the proposed method, the lengths of the $H_a(z)$, $H_{m1}(z)$ and $H_{m2}(z)$ are 205, 83, 93, respectively, with $M = 16$ and $N = 2$. Altogether 192 multipliers and 381 adders are needed. The group delay is 1760. The number of multipliers is reduced by 12.3% and the number of adders is reduced by 12.8% while the group delay increases by 1.3%. The frequency response of overall filter is shown in Fig. 6.9.

B. Example 2

Another Example is a lowpass FIR filter with passband and stopband edges at 0.2π and 0.21π respectively. The passband ripple is 0.001 and the stopband attenuation is 60dB. According to [17], FRM method is more efficient than IFIR method for

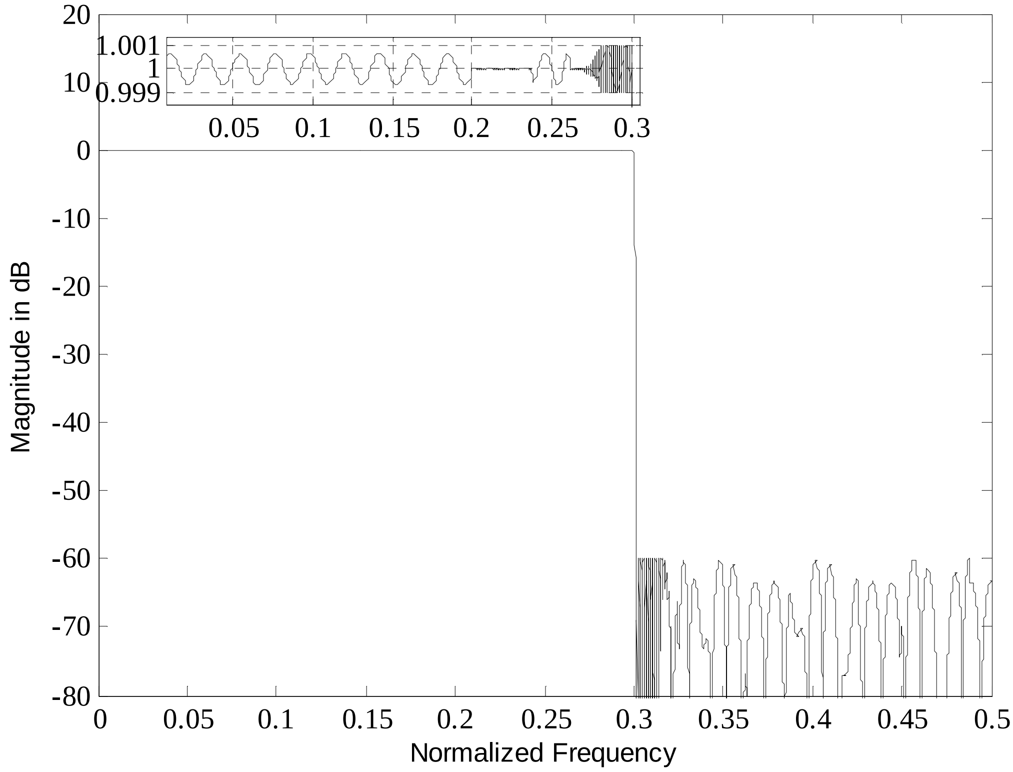


Figure 6.9: Frequency response of the overall filter in example 1.

the design of such a filter. Using the conventional FRM scheme, it is a Case B design with $M = 7$ and the lengths of the prototype filter and the masking filters are 97, 56 and 42 respectively. 98 multipliers and 195 adders are needed to obtain the overall filter. The group delay is 363.5. Using the proposed method, it is a Case B design with $M = 7$ and $N = 3$. The lengths of the $H_a(z)$, $H_{m1}(z)$ and $H_{m2}(z)$ are 95, 21 and 45 respectively. Altogether 82 multipliers and 161 adders are needed. The group delay is 381. The number of multipliers is reduced by 15.5% and the number of adders is reduced by 17.4% while the group delay increases by 4.8%. The frequency response of overall filter is shown in Fig. 6.10.

It can be seen from the two examples that using the proposed scheme, the complexity of filters are reduced effectively especially if the passband of the desired filter is narrow. The group delay is slightly increased as a cost.

6.5 Summary

In this chapter, a low complexity design based on FRM technique was proposed. Instead of using parallel masking to mask the bandedge shaping filter and its complement, the proposed structure uses two-step serial masking. Since the first step masking filter can be an interpolated filter, the complexity of the overall filter is reduced. Examples showed that the proposed structure works effectively especially for sharp filters with narrow passbands.

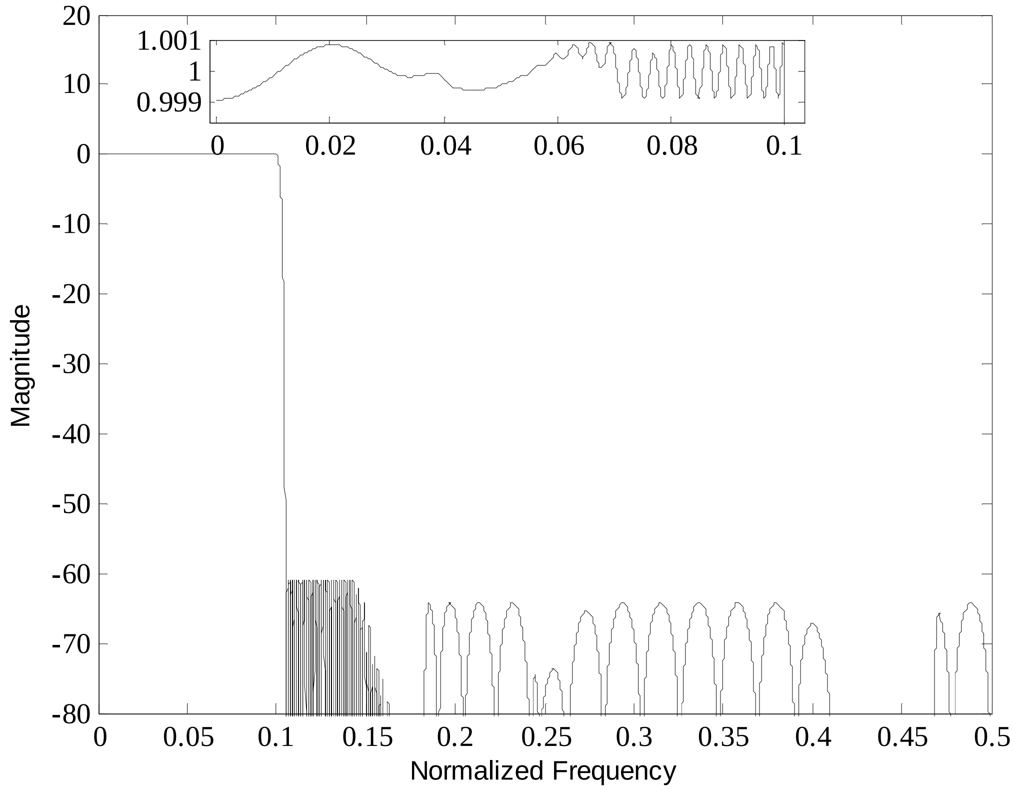


Figure 6.10: Frequency response of the overall filter in example 2.

Chapter 7

A 1 GHz Decimation Filter for Sigma-Delta ADC

7.1 Introduction

The requirement for high-speed analog to digital (A/D) conversion is growing since the advent of software defined radio. A soft-radio makes use of software on a micro-processor or mostly digital signal processor to process the radio signals of different standards. The conversion of radio frequency signals to digital signals requires a very high-speed A/D converter operating at Giga Hertz range. An efficient way to achieve A/D conversion at such high sampling rates is to use a bandpass delta-sigma converter [117–122], which consists of a delta-sigma modulator and a decimation filter. The modulator samples the input analog signal at a much higher rate than the Nyquist frequency by trading off resolution in magnitude for resolution in time, so that instead of having a highly accurate multi-bit output stream at Nyquist rate, they can have the same accuracy with a single bit output stream at several times as high as Nyquist rate.

The main function of the decimation filters is anti-aliasing while accomplishing the reduction of sampling rate. It receives a 1-bit data stream at a very high sampling rate from the output of Sigma-Delta modulator and outputs a multi-bits data stream at the Nyquist rate. The frequency response of a decimation filter should satisfy the desired requirements on both the passband drop and the out of band attenuation. When designing decimation filters, considerations include passband drop ($f = f_c$), aliasing error and worst case imaging error ($f = 1/M - f_c$), where f_c is the edge of the useful signal bandwidth, and M is the decimation factor. These parameters are illustrated Fig. 7.1 clearly.

The conversion rate between input and output signals can be as high as 1000 times, e.g. a GHz input signal can be down-sampled to a MHz output signal. Under such a situation, very long FIR filters are needed to prevent aliasing in the desired band. It is very difficult to implement a very long FIR filter in Giga Hertz range. Therefore, such a decimation filter is normally implemented in a multi-stage architecture, as shown in Fig. 7.2. The first stage of the decimation filter consists of several comb filters which are simple and do not require any multipliers. Hence these filters can

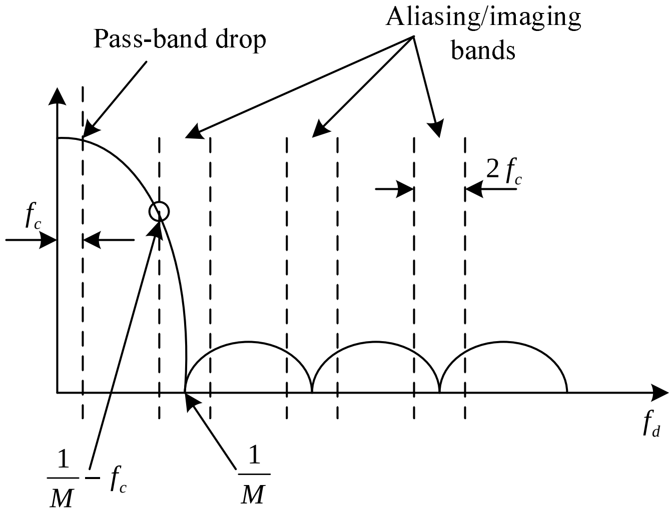


Figure 7.1: Important parameters according to the requirements of decimation filters.

operate at very high-speed. The second stage of the filter acts as an equalizer, as the comb filter has a Sinc-like frequency response and introduces a droop towards the edge of the passband. This droop is compensated in the second stage. The final stage is a generic FIR filter which provides adequate decimation factor.

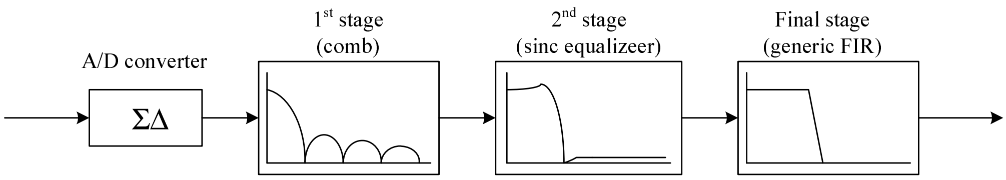


Figure 7.2: A/D Converter with 3-stage decimation filter.

The study focuses on the implementation of the first-stage comb filter. The comb filters operate at the highest sampling rate and usually have huge power consumption. Once the operating frequency is reduced, it is easy to design the rest stages.

This chapter consists of four sections. In Section 7.2, the concept and the implementation of comb filters are briefly reviewed. In Section 7.3, a filter suitable for high speed implementation is proposed. The power consumption is calculated using the power figure of merit (PFOM). A conclusion is given in Section 7.4.

7.2 Overview of Comb Filters

In a multi-stage decimation filter design, the initial stages are normally implemented using comb filters. Comb filters are a class of linear-phase FIR filters and characterized by the following z-transform transfer function:

$$H(z) = \left(\frac{1 - z^{-M}}{1 - z^{-1}} \right)^{k}, \tag{7.1}$$

where k is the order of the comb filter and M is the decimation factor.

There are many ways to implement high-speed comb filters. The cascaded integrator and differentiator approach [124][125] separates the filter into two sections as shown in (7.2) and illustrated by Fig. 7.3. Each of the sections is a cascade of k identical filters. The down-sampler is between the integrator and differentiator stages. The advantage of the scheme is that the differentiator operates at the rate M times lower than the integrator. The drawback is that in order to ensure the stability of the integrator, the word-length of the integrator section has to be sufficient long which contributes to a high power consumption and a large chip area. The word-length is indicated by (7.3), where W_o and W_i are the output and input word-lengths respectively.

$$H(z) = \left(\frac{1}{1 - z^{-1}}\right)^k \cdot \left(1 - z^{-M}\right)^k. \tag{7.2}$$

$$W_o = W_i + k \cdot \log_2 M. \tag{7.3}$$

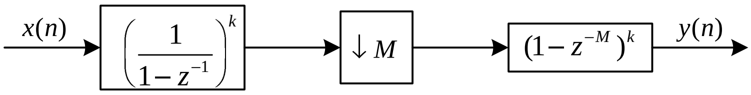


Figure 7.3: The cascaded integrator and differentiator approach.

The non-recursive implementation realizes the comb filter as a FIR filter [123], which is computationally efficient for a small decimation factor. If the decimation factor M is powers of two, (7.1) becomes (7.4). The commutative rule can be used in decimation process. The non-recursive architecture can be implemented by cascading identical filters followed by decimation with a factor of 2, as shown in Fig 7.4. It is clear that only the first $(1 + z^{-1})^k$ operates at high input data rate and the word-length of stage j is limited to (7.5). The filter speed is only limited by the first stage which has the smallest word-length . Because as the word-length increases, the sampling rate decreases, the power consumption is reduced compared to the recursive architecture. Additionally, this structure does not have stability problem.

$$H(z) = \prod_{i=0}^{\log_2 M-1} (1 + z^{-2^i})^k. \tag{7.4}$$

$$W_j = W_i + kj. \tag{7.5}$$

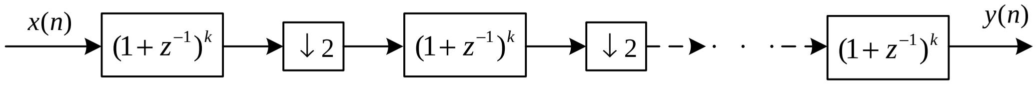


Figure 7.4: Nonrecursive implementation of cascaded CIC with powers of two decimation factor.

Besides the above methods, the table look-up method speeds up the filtering process by storing pre-computed partial sums in memory [126]. Comb filters can also be decomposed using polyphase structure, thus the multiplications and additions are performed at a frequency much lower than the signal rate, which will be discussed in detail in Section 7.3. Furthermore, all additions from the branches of polyphase structure can be gathered into one Wallace tree, which will significantly reduce the power consumption according to [132].

7.3 Design of the Decimation Filter

The output of a bandpass Sigma-Delta modulator is a bandpass signal. The signal is first down-converted to DC by I/Q mixer oscillators ($e^{-j\omega_c n}$), then passes through a complex lowpass filter $H(z)$, as shown in Fig. 7.5. Suppose the sampling frequency is f_s and the carrier frequency is f_c, ω_c can be presented as

$$\omega_c = \frac{f_c}{f_s} \cdot 2\pi. \tag{7.6}$$

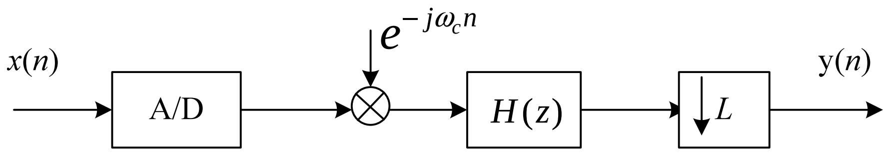


Figure 7.5: Structure of bandpass down-sampling.

As mentioned above, decimation is usually realized by multistage architecture. Since the input data rate is very high, the low power implementation of the first stage decimation filter is the most difficult task in the whole design. An efficient design is to use polyphase structure to divide the filter into M parallel branches which operate at a sampling rate M times lower than the input rate. The decomposition is presented by (7.7). The coefficients of the branch filters $E_i(z)$ are listed in (7.8). After the down-samplers are brought forward to the front of the branch filters, the down-sampler and the delay in each branch can be replaced with a commutative switch.

$$H(z) = \sum_{i=0}^{M-1} z^{-i} E_i(z^M). \tag{7.7}$$

$$
\begin{aligned}
e_0(n) &= [h(0), h(M), \cdots] \\
e_1(n) &= [h(1), h(M+1), \cdots] \\
&\vdots \\
e_{M-1}(n) &= [h(M-1), h(2M+1), \cdots].
\end{aligned}
\tag{7.8}
$$

The number of branches of the polyphase structure can be reduced if the ratio of sampling frequency to carrier frequency is special. In this work, the sampling frequency f_s is 4GHz and the carrier frequency f_c is 1GHz, thus ω_c equals to $\pi/2$. The I/Q mixer oscillators are $[1, 0, -1, 0, \cdots]$ and $[0, -1, 0, 1, \cdots]$ respectively. In each serial, zero and nonzero appear alternatively.

Block diagrams of decimation by 2 and by 4 are shown in Fig. 7.6 and Fig. 7.7, respectively. Half of the branches of the polyphase structure could be eliminated due to the special ratio.

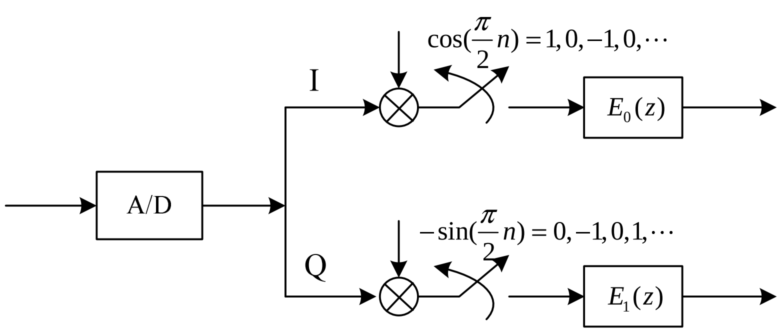


Figure 7.6: An efficient structure for $f_s = 4f_c$, decimate by 2.

Power consumption is proportional to the word-length and the operating frequency.

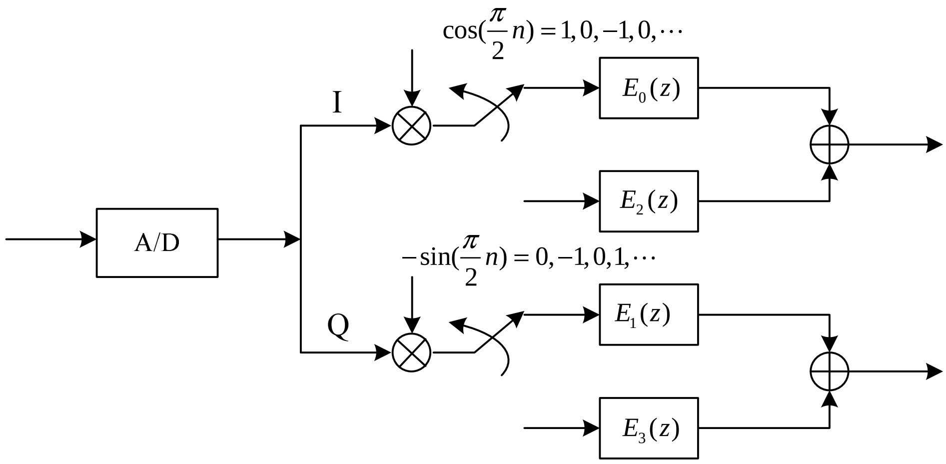


Figure 7.7: An efficient structure for $f_s = 4f_c$, decimate by 4.

In Fig.s 7.6 and 7.7 no flip-flops and adders operate at input data rate, which reduces the power consumption effectively. Since the input rate is as high as Giga Hz, the first stage of decimation should be as simple as possible. To determine the decimation factor for the first stage, it is necessary to compare the power consumption of different decimation schemes, for a decimation by 4, it can be done by either two cascaded decimation by 2 or one decimation by 4. Since the modulator is a 4^{th} order bandpass, the order of comb filter is at least 3 [128][129]. The frequency response of the first stage decimation filter is shown in Fig. 7.8.

For decimation by 4, the two branches left is $E_0(z) = 1 + 12z^{-1} + 3z^{-2}$ and $E_2(z) = 6 + 10z^{-1}$ for the I branch. The architecture is shown in Fig. 7.9. $E_0(z)$ is implemented by transposed form to limit the critical path to 1 multiply and 1 add. The multiplication of 12 is achieved in two stages. If the input signal $x(n)$

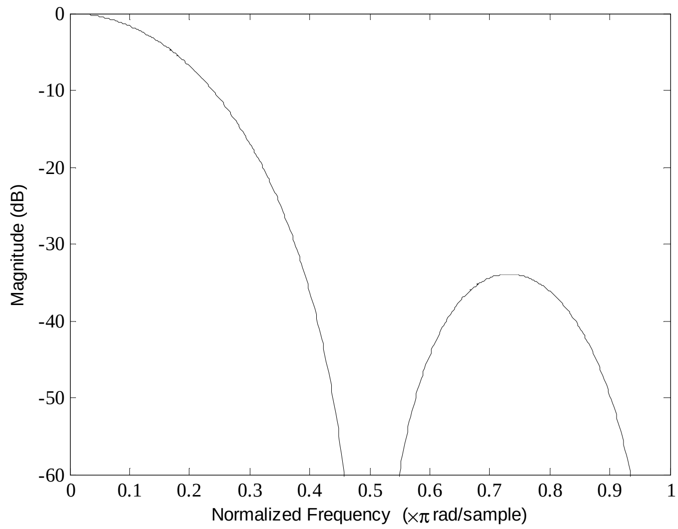


Figure 7.8: Frequency response of the first stage decimation filter

is first multiplied by 3, the word-length will be increased. Fortunately, due to the data-broadcast structure, the multiplication can be done after the summation of $4x(n)$ and $x(n-1)$. $E_2(z)$ is implemented by direct form. To reduce the internal word-length, the sum of $3x(n)$ and $5x(n-1)$ are shifted instead of calculating the sum of $6x(n)$ and $10x(n-1)$ directly. The multiplications in the architecture are replaced by several shifts and adders. The two branches left for the Q branch is $E_1(z) = 3 + 12z^{-1} + z^{-2}$ and $E_3(z) = 10 + 6z^{-1}$. The architecture is shown in Fig. 7.10.

For decimation by 2, the branch left is $1 + 3z^{-1}$ for I branch and $3 + z^{-1}$ for Q branch, which can be implemented by direct form I easily. The discussion of such a simple structure is omitted. Decimation by 4 can be obtained using two stages

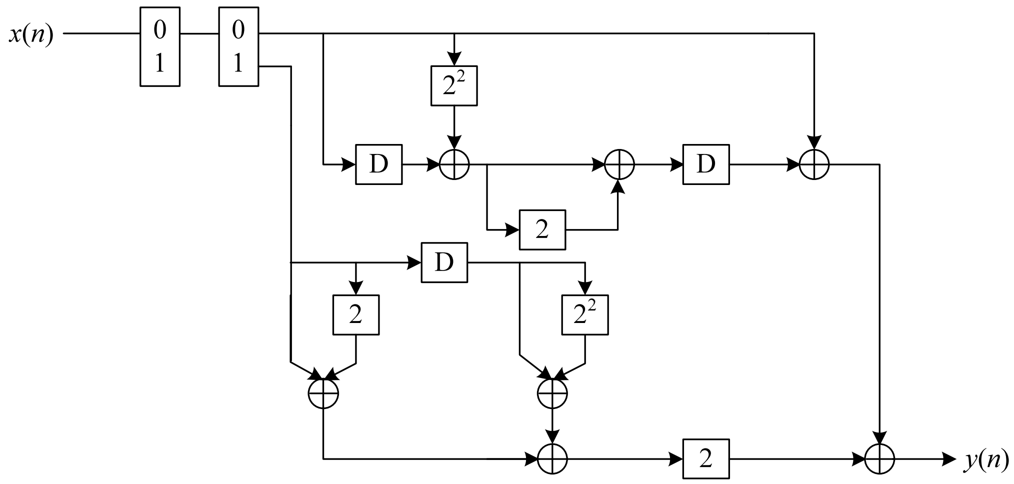


Figure 7.9: Implementation of decimation by 4 (I branch)

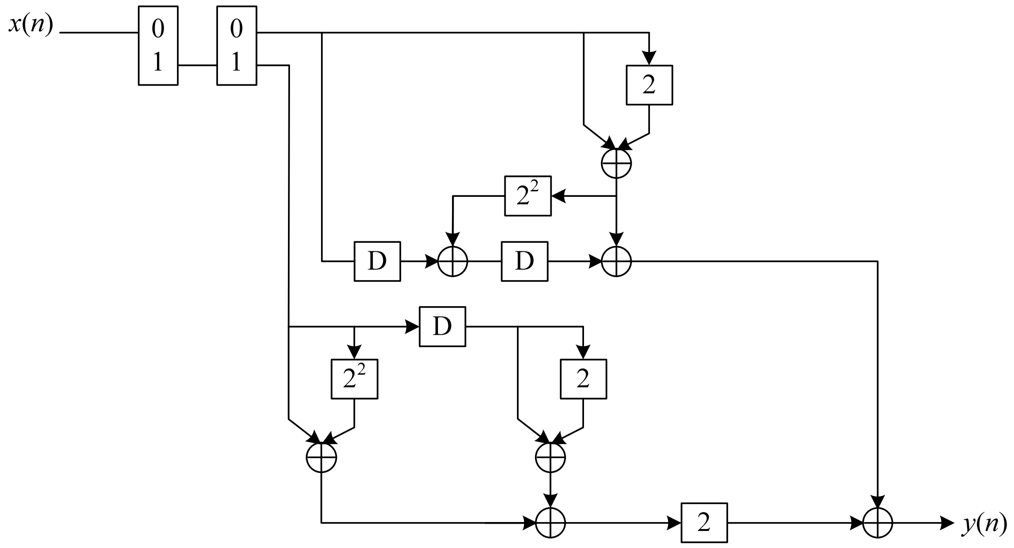


Figure 7.10: Implementation of decimation by 4 (Q branch)

of decimation by 2.

The power efficiency is evaluated between the two approaches, two stages of decimation by 2 and decimation by 4. The power consumptions of 2 stages of decimation by 2 and decimation by 4 filters are shown in Table 7.1 and Table 7.2, respectively. In this calculation, all the architectures work at the same supply voltage, which is normalized to 1. The power consumption of a D flip-flop is defined as 1. The Power Figure of Merit (PFOM) is introduced as data rate multiplies power consumption of 1-bit logic circuit. It enables us to compare the power efficiency among different filter structures. Note that the power consumption of a one-bit full adder is about 1/25 of the power consumption of the D flip-flop according to the circuits used in the simulation [130][131]. For example, in Table 7.1, the value of adder PFOM in I branch of 1st stage was calculated as $(1 + 2) \times 2/25 = 0.24$. The power consumption of the whole structure is proportional to the total value of its PFOM. Results shows that the power consumption of decimation by 4 is about 77.6% of that of decimation by two stages of decimation by 2.

The power consumption of the circuit with a 1.8 V supply was found to be $6mW$ for the in-phase part and $4mW$ for the quadrature part. The power consumption of the proposed structure is $10\mu W/MHz$ since the operating frequency is 1GHz, which is 62.5% of about $16\mu W/MHz$ in [132].

Table 7.1: Power consumption for 2 Stages of decimation by 2

$M_1 = 2$(1st stage)		Number	Word-length	Data Rate(GHz)	PFOM
I (or Q)	Flip-flop	1	1	2	2
	Adder	2	1 and 2	2	0.24
$M_2 = 2$(2nd stage)		Number	Word-length	Data Rate	PFOM
I(or Q)	Flip-flop	2	3	1	6
	Adder	2	3 and 4	1	0.28
		2	3 and 4	1	0.28
		1	5	1	0.2
Total PFOM=PFOM $(M_1) \cdot 2+$ PFOM $(M_2) \cdot 2 = 2.24 \cdot 2 \cdot +6.76 \cdot 2 = 18$					

Table 7.2: Power consumption for decimation by 4

$M = 4$		Number	Word-length	Data Rate(GHz)	PFOM
I	Flip-flop	2	1	1	2
		1	4	1	4
	Adder	3	3	1	0.36
		2	2 and 5	1	0.28
		2	4	1	0.32
Q	Flip-flop	2	1	1	2
		1	4	1	4
	Adder	3	4	1	0.48
		2	3	1	0.24
		2	2 and 5	1	0.28
Total PFOM=6.96 +7 =13.96					

7.4 Summary

In this chapter, a 1 GHz decimation filter has been introduced as the first-step decimation filter. It is carefully designed using polyphase structure in order to minimize the power consumption. The power consumption of decimation by 4 and two stages of decimation by 2 are compared in the form of Power Figure of Merit. Decimation by 4 is a better choice.

Chapter 8

Conclusions

The cost of implementing digital FIR filters and filter banks increases as the transition bandwidth decreases. To reduce the design complexity, efficient design techniques including IFIR and FRM are developed. The basic ideas of these techniques are interpolation and masking. The role of interpolation is to form the desired transition band. The resulted interpolated filter has very sparse coefficients which contributes to the low complexity of the overall filer. The role of masking is to remove the unwanted subbands produced by interpolation. The IFIR and FRM methods are closely related to each other. The IFIR method can be considered as the special case of FRM method. In this thesis, new design methods were developed to further reduce the complexity of the design of sharp FIR filters and filter banks.

The thesis consists of two parts. In the first part, computationally efficient filter banks were designed to meet the requirements of hearing amplification. Considering the fact that human ear has higher resolution at low frequencies than at high frequencies, non-uniform filter banks were designed to compensate the hearing loss. Due to the non-uniform allocation of frequency bands, the ability of compensation is enhanced. In Chapter 3, an 8-band non-uniformly spaced digital FIR filter bank was proposed. The use of two half-band FIR filters as prototype filters and the combination of FRM technique lead to significant savings in terms of number of multipliers, e.g. a minimum 80dB stopband attenuation is achieved using only 15 coefficients. Additionally, the performance of the proposed filter bank is enhanced by an optimal gain allocation process that helps to minimize the least square error between the objective audiogram and the magnitude response of the filter bank.

Though the complexity of the 8-band non-uniform filter bank is low, its delay is not small enough. In Chapter 4, a 16-band non-uniform low delay FIR filter bank was proposed. The delay of the overall filter is significantly reduced as the result of novel filter structure which reduces the interpolation factors of the prototype filters. Compared to the 8-band non-uniform filter bank in Chapter 3 and the 16-band uniform filter bank, group delay is reduced by 48% and 37.8%, respectively.

In the second part, new synthesis methods were proposed to simplify the design of FIR filters. A conventional FRM filter uses an interpolated filter as the band-edge shaping filter. In Chapter 5, a band-edge shaping filter with non-periodical frequency response was proposed. The multipliers of the sub-filters were shared efficiently. Simulation shows that great savings in terms of the number of multipliers are achieved compared to the traditional 1-stage FRM, 2-stage FRM, IFIR-FRM, and SFFM methods and the delay is reduced compared to SFFM method.

Another efficient design method based on FRM was proposed in Chapter 6. Instead of using parallel masking to mask the bandedge shaping filter and its complement, the proposed structure uses two-step serial masking. The first step masking filter can be an interpolated filters, hence the complexity of the overall filter is reduced. Examples show that design complexity is reduced effectively especially for sharp filters with narrow passbands.

In Chapter 7, a 1 GHz decimation filter was introduced as the first-step decimation filter. It was designed using polyphase structure to minimize the power consumption. The power consumption of decimation by 4 and that of two stages of decimation by 2 are compared in the form of Power Figure of Merit. Filter of decimation by 4 is proved to be a better first-step decimation filter.

Recommendation for Future Research

Among the filters and filter banks proposed in the thesis, the optimization methods adopted is linear programming and Remez algorithm. It is interesting to explore if the design complexity can be further reduced using non-linear programming algorithms. Additionally, in the design methods proposed, the sub-filters are not jointly designed. It is worth exploring the effectiveness of joint design in further work.

Bibliography

[1] J. F. Kaiser, "Nonrecursive digital filter design using i_0-sinh window function," *Proc.of IEEE Int. Symp. Circuits Syst.*, pp. 20-30, Apr. 1974.

[2] Y. Neuvo, C. Y. Dong and S.K. Mitra, "Interpolated finite impulse response filters," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 32, pp. 563-570, June 1984.

[3] T. Saramaki, T. Neuvo and S. K. Mitra, "Design of computationally efficient interpolated FIR filters," *IEEE Transactions on Circuits and Systems*, vol. 35, Issue 1, pp.70-88 , Jan. 1988,

[4] H. Kikuchi, H. Watanabe and T. Yanagisawa, "Interpolated FIR filters using cyclotomic polynomials," *Proc.of IEEE Int. Symp. Circuits Syst.*, vol.3, pp. 2009-2012, Jun. 1988.

[5] D. Pang, L. A. Ferrari and P. V. Sankar, "A unified approach to general IFIR filter design using B-spline functions," *IEEE Transactions on Signal Processing*, vol 39, Issue 9, pp. 2115-2118, Sept. 1991.

[6] O. Gustafsson, H. Johansson and L. Wanhammar, "Design and efficient implementation of high-speed narrow-band recursive digital filters using single filter

frequency masking techniques," *Proc.of IEEE Int. Symp. Circuits Syst.*, vol.3, pp. 359-362, may 2000.

[7] O. Gustafsson, H. Johansson, and L. Wanhammar, "Narrow-band and wide-band single filter frequency masking FIR filters," *Proc.of IEEE Int. Symp. Circuits Syst.*, vol. 2, pp. 181-184, may 2001.

[8] A. Abousaada, T. Aboulnasr and W. Steenaart, "An echo tail canceller based on adaptive interpolated FIR filtering," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 39, Issue 7, pp. 409-416, July 1992.

[9] Yuan-Pei Lin and P. P. Vaidyanathan, "An iterative approach to the design of IFIR matched filters," *Proc.of IEEE Int. Symp. Circuits Syst.*, vol. 4, pp. 2268-2271, June 1997.

[10] G. Jovanovic-Dolecek and J. Javier Diaz Carmona, "Lowpass minimum phase filter design using IFIR filters," *Electronics Letters*, vol. 33, Issue 23, pp. 1933-1935, Nov. 1997.

[11] Jae-Wan Kim, Byung-Moo Min and Jang-Sik Yoo, "An area-efficient sigma-delta DAC with a current-mode semi-digital IFIR reconstruction filter," *Proc.of IEEE Int. Symp. Circuits Syst.*, vol. 1, pp. 344-347, June 1998.

[12] Cheng-Shing Wu and An-Yeu Wu, "A novel cost-effective multi-path adaptive interpolated FIR (IFIR)-based echo canceller," *Proc.of IEEE Int. Symp. Circuits Syst.*, vol. 5, pp. 453-456, May 2002.

[13] L. S. Nielsen and J. Sparso, "Designing asynchronous circuits for low power:

an IFIR filter bank for a digital hearing aid," *Proc.of IEEE*, vol. 87, Issue 2, pp. 268-281, Feb. 1999.

[14] K. Rajgopal and S. Venkataraman, "A Delayless Adaptive IFIR Filterbank Structure for Active Noise Control," *TENCON 2005 IEEE Region 10*, pp. 1-6, Nov. 2005.

[15] Zhongqi Jing and Adly Fam, "A new structure for narrow transition band, lowpass digital filter design," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, Issue 2, pp. 362-370, Apr. 1984.

[16] Y. C. Lim, "Frequency-response masking approach for the synthesis of sharp linear phase digital filters," *IEEE Trans. Circuits Syst.*, CAS-33, pp. 357-364, Apr. 1986.

[17] Y. C. Lim and Y. Lian, "The optimal design of one- and two- dimensional FIR filters using the frequency response masking technique," *IEEE Trans. Circuits Syst.*, Part 2, vol. 40, pp. 88-95, Feb. 1993.

[18] Y. C. Lim and Y. Lian, "Frequency-response masking approach for digital filter design: complexity reduction via masking filterfactorization," *IEEE Trans. Circuits Syst.II*, vol 41, Issue 8, pp. 518 - 525, Aug. 1994.

[19] Y. Lian, "A new frequency-response masking structure with reduced complexity for FIR filter design," *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 609-612, May 2001.

[20] L. Zhang, Y. Lian and C. C. Ko, "A new approach for design sharp FIR filters using frequency-response masking technique," *Proc.* 9^{th} *IEEE DSP Workshop.*

[21] Y. Lian, L. Zhang and C. C. Ko, "An improved frequency response masking approach for designing sharp FIR filters," *Signal Processing*, vol. 81, No. 12, pp. 2573-2581(9), Dec. 2001.

[22] Chunzhu Yang and Y. Lian, "A modified structure for the design of sharp FIR filters using frequency response masking technique," *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 237 - 240, May 2002.

[23] Chunzhu Yang and Y. Lian, "New Structures for Single Filter Based Frequency-Response Masking Approach," *IEEE Asia Pacific Conference on Circuits and Systems*, pp. 69-72, Dec. 2006.

[24] K. Supramaniam and Yong Lian, "Complexity reduction for frequency-response masking filters using cyclotomic polynomial prefilters," *IEEE International Symposium on Circuits and Systems*, May 2006.

[25] Jianghong Yu and Yong Lian, "Frequency-response masking based filters with the even-length bandedge shaping filter," *IEEE International Symposium on Circuits and Systems*, vol. 5, pp. 536 - 539, May 2004.

[26] T. Saramaki and Yong Ching Lim, "Use of the Remez algorithm for designing FIR filters utilizing the frequency-response masking approach," *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 449-455, 30 May - 2 June 1999.

[27] T. Saramaki and H. Johansson, "Optimization of FIR filters using the frequency-response masking approach," *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 177-180, May 2001.

[28] Ya Jun Yu and Yong Ching Lim, "Genetic algorithm approach for the optimization of multiplierless sub-filters generated by the frequency-response masking technique," 9^{th} *International Conference on Electronics, Circuits and Systems*, vol. 3, pp. 1163-1166, Sept. 2002.

[29] Wu-Sheng Lu, and T. Hinamoto, "Optimal design of frequency-response-masking filters using semidefinite programming" *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, Issue 4, pp. 557-568, April 2003.

[30] Wu-Sheng Lu and T. Hinamoto, "Optimal design of FIR frequency-response-masking filters using second-order cone programming," *IEEE International Symposium on Circuits and Systems*, vol.3, pp. III-878 - III-881, May 2003.

[31] W. R. Lee, V. Rehbock, K. L. Teo and L. Caccetta, "A weighted least-square-based approach to FIR filter design using the frequency-response masking technique," *Signal Processing Letters*, vol. 11, Issue 7, pp. 593-596, July 2004.

[32] W. R. Lee, L. Caccetta, K. L. Teo and V. Rehbock, "A weighted least squares approach to the design of FIR filters synthesized using the modified frequency response masking structure," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 53, Issue 5, pp. 379-383, May 2006.

[33] Wu-Sheng Lu and T. Hinamoto, "Improved design of frequency-response-masking filters using enhanced sequential quadratic programming," *International Symposium on Circuits and Systems*, vol. 5, pp. 528 - 531, May 2004.

[34] YaJun Yu, Yong Ching Lim, Kok Lay Teo and Guohui Zhao, "Frequency-response-masking technique incorporating extrapolated impulse response band-

edge shaping filter," *Proceedings of the 2004 International Symposium on Circuits*, vol. 5, pp. 532 - 535, May 2004.

[35] Y. C. Lim and S. H. Low, "The synthesis of sharp diamond-shaped filters using the frequency response masking approach," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 2181-2184, Apr. 1997.

[36] S. H. Low and Y. C. Lim, "A new approach to design sharp diamond-shaped filters," *Signal Processing*, vol. 67, Number 1, pp. 35-48(14), May 1998.

[37] Y. C. Lim and S. H. Low "Frequency-response masking approach for the synthesis of sharp two-dimensional diamond-shaped filters," *Signal Processing*, vol. 45, Issue 12, pp. 1573 - 1584, Dec. 1998.

[38] T, Saramaki, Y. C. Lim and R. Yang, "The synthesis of half-band filter using frequency-response masking technique," *IEEE Transactions on Digital Signal Processing*, vol. 42, Issue 1, pp. 58 - 60, Jan. 1995.

[39] S. H. Low and Y. C. Lim, "Synthesis of 2-D half-band filters using the frequency response masking technique," *Proceedings of the 1998 IEEE International Symposium on Circuit and Systems*, vol. 5, pp. 57-60, May 1998.

[40] S. H. Low and Y. C. Lim, "Multi-stage approach for the design of 2-D half-band filters using the frequency response masking technique," *2001 IEEE International Symposium on Circuit and Systems*, vol. 2, pp. 557-560, May 2001.

[41] M. D. Lutovac and L. D. Milic, "IIR filters based on frequency-response masking approach," 5^{th} *International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Service*, vol. 1, pp. 163-170, Sept. 2001.

[42] H. H. Chen, S. C. Chan and K. L. Ho, "A semi-definite programming (SDP) method for designing IIR sharp cut-off digital filters using frequency-response masking," *Proceedings of the 2004 International Symposium on Circuits and Systems*, vol. 3, pp. 149-152, May 2004.

[43] Yongzhi Liu and Zhiping Lin, "Design of complex FIR filters using the frequency-response masking approach," *Proceedings of the 2005 International Symposium on Circuits and Systems*, vol. 3, pp. 2024-2027, May 2005.

[44] Yongzhi Liu and Zhiping Lin, "Synthesis of very sharp Hilbert transformer using the frequency-response masking technique," *IEEE Transactions on Signal Processing*, vol. 53, Issue 7, pp. 2595-2597, July 2005.

[45] Yong Ching Lim, Yajun Yu and T. Saramaki, "Optimum masking levels and coefficient sparseness for Hilbert transformers and half-band filters designed using the frequency-response masking technique," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, Issue 11, pp. 2444-2453, Nov. 2005.

[46] Yong Lian and Jianghong Yu, "The reduction of noises in ECG signal using a frequency response masking based FIR filter," *2004 IEEE International Workshop on Circuits and Systems*, vol. 52, pp. S2/4 - 17-20, Dec. 2004.

[47] Lihong Zhou, Wenjiang Pei, Pengcheng Xi and Zhenya He, "Frequency-Response Masking Approach for Design of Intermediate Frequency Filters in CDMA and Wideband GSM Modules," *IEEE Asia Pacific Conference on Circuits and Systems*, pp. 736-739, Dec. 2006.

[48] S. L. Netto, P. S. R. Diniz and L. C. R. Barcellos, "Efficient implementation for cosine-modulated filter banks using the frequency response masking approach,"

IEEE International Symposium on Circuits and Systems, vol. 3, pp. 229-232, May 2002.

[49] L. Rosenbaum, P. Lowenborg and M. Johansson, "Cosine and sine modulated FIR filter banks utilizing the frequency-response masking approach," *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 882 - 885, May 2003.

[50] M. B. Furtado Jr., P. S. R. Diniz and S. L. Netto, "Optimization techniques for cosine-modulated filter banks based on the frequency-response masking approach," *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 890 - 893, May 2003.

[51] M. B. Furtado Jr., P. S. R. Diniz and S. L. Netto, "On the design of high-complexity cosine-modulated transmultiplexers based on the frequency-response masking approach," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 52, Issue 11, pp. 2413 - 2426, Nov. 2005.

[52] Ching-Shun Lin and C. Kyriakakis, "Frequency response masking approach for designing filter banks with rational sampling factors," *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 99 - 102, Oct. 2003.

[53] Ching-Shun Lin, Yong Ching Lim, R. Bregovic and T. Saramaki, "Frequency response masking based design of two-channel FIR filterbanks with rational sampling factors and reduced implementation complexity," *Proceedings of the 4^{th} International Symposium on Image and Signal Processing and Analysis*, pp. 121 - 126, Sept. 2005.

[54] R. E. Crochiere and L. R. Rabiner, Multirate Digital Signal Processing, Prentice Hall, Inc., 1983.

[55] A. Biem, S. Katagiri and E. McDermott, "An application of discriminative feature extraction to filter-bank-based speech recognition," *IEEE Transactions on Speech and Audio Processing*, vol. 9, Issue 2, pp. 96-110, Feb. 2001.

[56] Y. Shao and C. H. Chang, "A Generalized Time-Frequency Subtraction Method for Robust Speech Enhancement Based on Wavelet Filter Banks Modeling of Human Auditory System," *IEEE Transactions on Systems and Man and Cybernetics-Part B: Cybernetics*, accepted for future publication, vol. PP, Issue 99, pp. 877-889, 2007.

[57] J. Kovacevic and M. Vetterli, "FCO sampling of digital video using perfect reconstruction filter banks," *IEEE Transactions on Image Processing*, vol. 2, Issue 1, pp. 118-122, Jan. 1993.

[58] C. T. Chiu, "Parallel implementation of transform-based DCT filter-bank for video communications," *IEEE Transactions on Consumer Electronics*, vol. 40, Issue 3, pp. 473-475, Aug. 1994.

[59] T. Uto, T. Oka and M. Ikehara, "M-Channel Nonlinear Phase Filter Banks in Image Compression: Structure, Design, and Signal Extension," *IEEE Transactions on Signal Processing*, vol. 55, Issue 4, pp. 1339-1351, Apr. 2007.

[60] K. A. Kotteri, A. E. Bell and J. E. Carletta, "Multiplierless filter Bank design: structures that improve both hardware and image compression performance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, Issue 6, pp. 776-780, June 2006.

[61] T. Tanaka,Y. Hirasawa and Y. Yamashita, "Variable-length lapped transforms with a combination of multiple synthesis filter banks for image coding," *IEEE Transactions on Image Processing*, vol. 15, Issue 1, pp. 81-88, Jan. 2006.

[62] B. R. Horng, H. Samueli and A. N. Jr. Wilson, "The design of low-complexity in linear-phase FIR filter banks using powers-of-two coefficients with an application to subband image coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, Issue 4, pp. 318-324, Dec. 1991.

[63] F. J. Harris, C. Dick and M. Rice, "Digital receivers and transmitters using polyphase filter banks for wireless communications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, Issue 4, Part 2, pp. 1395-1412, Apr. 2003.

[64] X. G. Xia, "New Precoding for Intersymbol Interference Cancellation Using Nonmaximally Decimated Multirate Filter Banks with Ideal FIR Equalizers," *IEEE Transactions on Signal Processing*, vol. 45, no. 10, pp. 2431-2441, Oct. 1997.

[65] A. Scaglione, G. B. Giannakis and S. Barbarossa, "Redundant Filter Bank Precoders and Equalizers Part I: Unification and Optimal Designs," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 1988-2006, July 1999.

[66] A. V. Krishna and K. V. S. Hari, "Filter Bank precoding for FIR equalization in high-rate MIMO communications," *IEEE Transactions on Signal Processing*, vol. 54, Issue 5, pp. 1645-1652, May 2006.

[67] P. P. Vaidyanathan, Y. P. Lin, S. Akkarakaran and S. M. Phoong, "Optimalilty

of Principal Component Filter Banks for Discrete Multitone Communication Systems," *Proceedings of IEEE ISCAS*, Geneva, May 2000.

[68] G. B. Giannakis, "Filter Banks for Blind Channel Identification and Equalization," *IEEE Signal Processing Letters*, vol. 4, no. 6, pp. 184-187, June 1997.

[69] A. M. Jalil, H. Amindavar and F. Almasganj, "Subband blind equalization using wavelet filter banks," *IEEE International Symposium on Circuits and Systems*, vol. 6, pp. 5730-5733, May 2005.

[70] P. P. Vaidyanathan, Multirate Systems and Filter Banks, PTR Prentice Hall, 1992.

[71] J. Wintermantel and N. J. Fliege, "A class of DFT polyphase filter banks for orthogonal multiple carrier data transmission," *IEEE International Symposium on Circuits and Systems*, vol. 3, pp. 1332-1335, May 1992.

[72] C. D. Creusere and S. K. Mitra, "Efficient image scrambling using polyphase filter banks," *IEEE International Conference on Image Processing*, vol. 2, pp. 81-85, Nov. 1994.

[73] M. Sablatash, J. Lodge, "Theory and design of spectrum-efficient bandwidth-on-demand multiplexer-demultiplexer pairs based on wavelet packet tree and polyphase filter banks," *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1797-1800, May 1998.

[74] Y. C. Lim and B. Farhang-Boroujeny, "Fast filter bank (FFB)," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 39, Issue 5, pp. 316-318, May 1992.

[75] Lee Jun Wei and Y. C. Lim, "Designing the fast filter bank with a minimum complexity criterion," *IEEE Seventh International Symposium on Signal Processing and Its Applications*, vol. 2, pp. 279-282, July 2003.

[76] Y. C. Lim and Jun Wei Lee, "Matrix formulation: fast filter bank," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 133-136, May 2004.

[77] T. Blu, "An iterated rational filter bank for audio coding," *IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, pp. 81-84, June 1996.

[78] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding shemes," *IEEE Conference on Acoustics, Speech, and Signal Processing*, pp. 191-195, May 1977.

[79] Kuansan Wang and S.A. Shamma, "Spectral shape analysis in the central auditory system," *IEEE Transactions on Speech and Audio Processing*, vol. 3, Issue 5, pp. 382-395, Sept. 1995.

[80] T. Koizumi, M. Mori and S. Taniguchi, "Speech recognition based on a model of human auditory system Spoken Language," *Fourth International Conference on Spoken Language Processing*, vol. 2, pp. 937- 940, 1996.

[81] S. Seneff, "A computational model for the peripheral auditory system: Application of speech recognition research," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 11, pp. 1983-1986, Apr 1986.

[82] Marion A. Hersh, Michael A. Johnson and etc., Assistive technology for the hearing-impaired, deaf and deafblind, Springer-Verlag London Limited, 2003.

[83] A. M. Engebretson, "Benefits of digital hearing aids," *IEEE Engineering in Medicine and Biology Magazine*, vol. 13, Issue 2, pp. 238-248, April-May 1994.

[84] T. J. Gennosa, R. Batra and S. Kuwada, "Estimating human hearing thresholds using potentials that follow sinusoidally amplitude modulated (SAM) tones," *Proceedings of the IEEE Seventeenth Annual Northeast Bioengineering Conference*, pp. 125-126. April 1991.

[85] J. P. Mobley, S. Gilman, "A microprocessor-controlled analog hearing aid," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1077-1078, Nov. 1989.

[86] R. Singh Rana, "Computer aided system simulation of micropower CMOS analog hearing aid," 10^{th} *Annual IEEE International ASIC Conference*, pp. 343-347, Sept. 1997.

[87] A. B. Bhattacharyya, R. S. Rana, S.K. Guha, and etc., "A micropower analog hearing aid on low voltage CMOS digital process," 9^{th} *International Conference on VLSI Design*, pp. 85-85, Jan. 1996.

[88] J. C. Ventura, "Digital audio gain control for hearing aids," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 2049-2052, May 1989.

[89] J. C. Tejero, S. Bernal, J. A. Hidaldo, J. Fernandez, R. Urquiza and A. Gago, "A digital hearing aid that compensates loudness for sensorineural hearing impairments," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 2991-2994, May 1995.

[90] Young-Cheol Park, Dong-Wook Kim and In-Young Kim, "Design of high-performance digital hearing aid processor," *Electronics Letters*, vol. 34, Issue 17, pp. 1631-1632, Aug. 1998.

[91] F. Nielsen and J. Sparso, "A 85 W asynchronous filter-bank for a digital hearing aid," 45^{th} *IEEE International Solid-State Circuits Conference*, pp. 108-109, Feb. 1998.

[92] E. Onat, M. Ahmadi, G. A. Jullien and W. C. Miller, "Optimized delay characteristics for a hearing instrument filter bank," *Proceedings of the* 43^{rd} *IEEE Midwest Symposium on Circuits and Systems*, vol. 3, pp. 1074-1077, Aug. 2000.

[93] H. G. McAllister, N. D. Black, N. Waterman and M. Li, "Audiogram matching using frequency sampling filters," *Proceedings of the* 16^{th} *IEEE Annual International Conference on Engineering Advances: New Opportunities for Biomedical Engineers*, vol. 1, pp. 249 - 250, Nov. 1994.

[94] A. B. Hamida, "Implication of new technologies in deafness healthcare: deafness rehabilitation using prospective design of hearing aid systems," *IEEE International Symposium on Technology and Society*, pp. 85-90, Sept. 2000.

[95] H. Li, G. A. Jullien, V. S. Dimitrov, M. Ahmadi and W. Miller, "A 2-digit multidimensional logarithmic number system filter bank for a digital hearing aid architecture," *IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 760-763, May 2002.

[96] Y. Wei and Y. Lian, "A Computationally Efficient Non-Uniform Digital Filter Bank for Hearing Aid," *IEEE International Workshop on Biomedical Circuits and Systems*, pp. S1.3.INV-17-20, Dec. 2004.

[97] Y. C. Lim, "A digital filter bank for digital audio systems," *IEEE Transactions on Circuits and Systems*, vol. 3, pp. 848-849, Aug 1986.

[98] T. Lunner and J. Hellgren, "A digital filterbank hearing aid design, implementation and evaluation," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. 3661-3664, April 1991.

[99] A. B. Hamida, "An adjustable filter-bank based algorithm for hearing aid systems," *International Conference on Industrial Electronics, Control and Instrumentation*, vol. 3, pp. 1187-1192, 1999.

[100] R. Brennan and T. Schneider, "An ultra low-power DSP system with a flexible filterbank," *The* 35^{th} *Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 809 - 813, Nov. 2001.

[101] M. A. Stone and B. C. J. Moore, "Tolerable hearing-aid delays: III. effects on speech production and perception of across-frequency variation in delay," *Ear and Hearing*, vol. 24 (2), pp. 175-183, April 2003.

[102] B. C. J. Moore, "Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids," *Ear and Hearing*, vol. 17(2), pp. 133-160, April 1996.

[103] Robert E. Sandlin, Text book of hearing aid amplification, Singular Publishing Group, 2000.

[104] R. Brennan and T. Schneider, "A flexible filterbank structure for extensive signal manipulations in digital hearing aids," *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, vol. 6, pp. 569-572, 31 May-3 June 1998.

[105] E. Zwicker and H. Fastl, Psychoacoustics, Facts and Models, Berlin: Springer Verlag, 1990.

[106] B. A. Daurich, L. R. Rabiner and T. B. Martin, "On the effects of varying filterbank parameters on isolated word recognition," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-31, no. 4, pp. 793-807, 1983.

[107] Guojun Zhou, J. H. L. Hansen and J. F, Kaiser, "Nonlinear feature based classification of speech under stress," *IEEE Trans.on Speech and Audio Processing*, vol. 9, pp. 201-215, 2001.

[108] P. Noll, "Digital audio coding for visual communications," *Proc. of IEEE*, vol. 83, pp. 925-943, June, 1995.

[109] N. Westerlund, J. M. de Haan, M. Dahl and I. Claesson, "Low distortion SNR-based speech enhancement employing critical band filter banks," *Proceedings of the 2003 Joint Conference on Information, Communications and Signal Processing*, vol. 1, pp. 129-133, Dec. 2003.

[110] Peterson, T. and S. Boll, "Critical band analysis-synthesis," *IEEE International Conference on ICASSP'81*, vol. 6, pp. 773-775, 1981.

[111] B. Carnero and Andrzej Drygajlo, "Perceptual speech coding and enhancement using framed-synchronized fast wavelet packet transform algorithms," *IEEE Trans. on Signal Processing*, vol. 47, no. 6, pp. 1622-1635, 1999.

[112] Ching-Ta Lu and Hsiao-Chuan Wang, "Speech enhancement using robust weighting factors for critical-band-wavelet-packet transform," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, pp. I - 721-4, May 2004.

[113] Chao Wang and Yit-Chow Tong, "An improved critical-band transform processor for speech applications," *Proceedings of the 2004 International Symposium on Circuits and Systems*, vol. 3, pp. III - 461-4 , May 2004.

[114] J. M. Kates, "An auditory spectral analysis model using the chirp z-transform," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-31, no. 1, pp. 148-156, 1983.

[115] Y. Lian and Y. Wei, "A Computationally Efficient Non-Uniform FIR Digital Filter Bank for Hearing Aid," *IEEE Trans. on Circuits and Systems I:Regular Papers*, vol. 52, pp. 2754-2762, Dec. 2005.

[116] Michael Valente, Hearing Aids: STANDARDS, OPTIONS, AND LIMITATIONS, Thieme, 2002.

[117] W. Gao and W. M. Snelgrove, "A 950-MHz IF Second-Order Integrated LC Bandpass Delta Sigma Modulator," *IEEE J. Solid-State Circuits*, vol. 33, No. 5, pp. 723-732, May 1998.

[118] A. Jayaraman, P. Asbeck, K. Nary, S. Beccue and Keh-Chung Wang, "Bandpass delta-sigma modulator with 800 MHz center frequency," *IEEE* 19^{th} *Annual Technical Digest on Gallium Arsenide Integrated Circuit (GaAs IC) Symposium*, pp. 95-98, Oct. 12-15, 1997.

[119] Y. F. Liang and Y. Lian, "A 250MHz CMOS Bandpass Delta Sigma Modulator using Continuous-Time Resonator Structure," *IEEE CAS Workshop on Wireless Communications and Networking*, California, USA, Sept. 2002.

[120] J. A. Cherry, W. M. Snelgrove and W. Gao, "On the design of a fourth-order

continuous-time LC delta-sigma modulator for UHF A/D conversion," *IEEE Trans.Circuits Syst. II*, vol. 47, pp. 518-530, June 2000.

[121] T. Kaplan, J. Cruz-Albrecht, M. Mokhtari, D. Matthews, J. Jensen and M.F. Chang, "1.3-GHz IF digitizer using a 4th-order continuous-time bandpass modulator," *IEEE Proc. Custom Int.Circ. Conference*, pp. 127-130, Sept. 2003.

[122] K. P. J. Thomas, R. S. Rana, and Y. Lian, "A 1 GHz CMOS Fourth-Order Continuous-Time Bandpass Sigma Delta Modulator for RF Receiver Front End A/D Conversion," *ASP-DAC* 2005, pp. 665-670, Shanghai, China, Jan. 2005.

[123] Y. Gao, L. Jia, and H. Tenhunen, "A fifth-order comb decimation filter for multi-standard transceiver applications," *The 2000 IEEE International Symposium on Circuits and Systems*, vol. 3, pp.89-92, Geneva, May 28-31, 2000.

[124] D. D. Kim and M. A. Brooke, "A 1.4G samples/sec comb filter design for decimation of sigma-delta modulator output," *Proceedings of the 2003 International Symposium on Circuits and Systems*, vol. 1, pp.I-1009 - I-1012, May 25-28, 2003.

[125] E. Hogenauer, "An economical class of digital filters for decimation and interpolation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, Issue 2, pp.155 - 162, Apr 1981.

[126] P. C. Maulik, M. S. Chadha, W. L. Lee and P. J. Crawley, "A 16-bit 250-khz delta-sigma modulator and decimation filter," *IEEE J. Solid-State Circuits*, vol. 35, pp. 458-467, Apr. 2000.

[127] Y. P. Xie, S. R. Whiteley and T. van Duzer, "High-speed decimation fil-

ter for delta-sigma analog-to-digital converte," *IEEE Transactions on Applied Superconductivity*, vol. 9, Issue 2, Part 3, pp. 3632 - 3635, June 1999.

[128] J. C. Candy, "Decimation for sigma delta modulation," *IEEE Transactions on Communications*, vol. 34, Issue 1, pp. 72 - 76, Jan 1986.

[129] E. N. Farag and R. Yan, and M. I. Elmasry, "A programmable power-efficient decmation filter for software radios," *International Symposium on Low Power Electronics and Design*, pp. 68 - 71, 18 - 20 Aug. 1997.

[130] R. Chandrasekaan, Y. Lian and R. S. Rana, "A High Speed Low Power CMOS Full Adder," US Patent Application Number 60/780, 193, filed March 08, 2006.

[131] R. Chandrasekaan, Y. Lian and R. S. Rana, "A High-Speed Low-Power D Flip-Flop," *The* 6^{th} *International Conference on ASIC(2005)*, vol. 1, pp. 82-85, Oct. 2005.

[132] H. Aboushady, Y. Dumonteix, M.-M. Louirat, and H. Mehrez, "Efficient polyphase decomposition of comb decimation filters in sigma-delta analog-to-digital converters," *IEEE Trans. Circuits Syst.-II*, vol. 48, no. 10, pp. 898-903, Oct. 2001.