Low Power Digital Filter Implementation for Hearing Aid Applications

[1]D. Joseph Jeyakumar, [2]J. Rajapaul Perinbam and [3]R. Rachel Selina
[1]Department of Media Science, Anna University, Chennai, India
[2]Department of Electronics and Communication Engineering,
Faculty of Information and Communication Engineering, Anna University, Chennai, India
[3]Anna University, Chennai, India

Abstract: The design of any battery operated device is limited by power consumption. The hearing aid prescription formula NAN-NL1 prescribes gain at ANSI 1/3 octave frequencies. Most digital hearing aid designs avoid ANSI S1.11 compliant filter banks due to the high computational complexity involved. IIR filter implementation of ANSI S1.11 design can be utilized only where linear phase is not important. An FIR-based ANSI S1.11 filter bank is presented for digital hearing aids which incorporates multi-rate algorithm to reduce data rates on bandwidth limited bands. FIR filter order is minimized in a systematic manner. Battery life longevity is preferred in electronic devices such as hearing aids. This has been achieved by incorporating various low power techniques in the design. Power reduction of about 30% is obtained when compared to the design of hearing aids available in the literature.

Key words: Hearing aids, filter bank, FIR digital filters, non-uniform filter bank, ANSI S1.11 1/3 octave filter bank, multi-rate systems

INTRODUCTION

Hearing aids are portable battery operated devices to compensate the hearing loss (auditory compensation) and to improve the speech intelligibility for the hearing impaired users by employing algorithms to reduce noise, cancel echo and enhance speech. The small dimension of hearing aids imposes a design constraint on the electronic devices and battery sizes. A generic digital hearing aid with blocks for auditory compensation, echo (feedback) cancellation, noise reduction/suppression and speech enhancement is shown in Fig. 1.

Hearing aid requires low amplification for low frequency signals and high amplification for high frequency signals. Auditory compensation performs frequency shaping to compensate the hearing loss. It consists of two modules: a filter bank and a compressor. The filter bank decomposes the input speech signal into different frequency bands and prescribed gains are applied on each band to compensate hearing loss. The signal's dynamic range is then reduced by the compressor to fit the processed speech into the limited hearing range of the user. The insertion gains, calculated using a fitting procedure are prescribed by NAL-NL1 (Ho and Young, 2004) on 1/3 octave frequencies defined by ANSI S1.11 standard. More, power is consumed by the filter bank of the auditory compensation system which is responsible for improving the quality of speech.
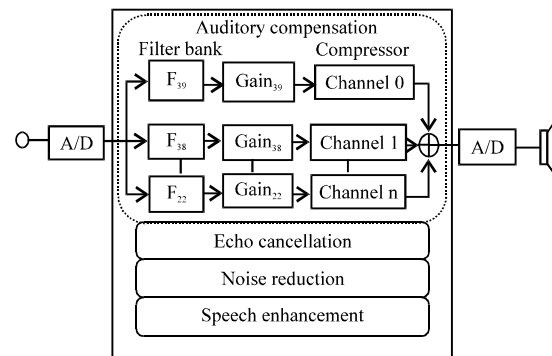

Fig. 1: Advanced digital hearing aid

Most existing ANSI S1.11 filter banks are implemented by Infinite Impulse Response filters (IIR) (Lozano and Carlosena, 2003) to reduce the computational complexity. But linear phase is required in feedback cancellation (Chong *et al.*, 2006) in digital hearing aids. Thus, a low complexity FIR filter bank that adheres to ANSI S1.11 standard is desired for digital hearing aids. This study presents a multi-rate FIR filter bank to reduce data rates of the bandwidth limited bands. Also, low power techniques have been incorporated to obtain a power efficient design. Various power estimation tools for FPGAs exist that help quantify power utilization for various implementations. Reducing the number of

Corresponding Author: D. Joseph Jeyakumar, Department of Media Science, Anna University, Chennai, India

multiplications and additions is possible by reducing the order of the filter. However, this is not easily achievable as the filter requirements depend on the order of the filter.

ANSI S1.11 STANDARD

ANSI S1.11 provides the performance specification for the octave and fractional-octave-band filter bank. It describes the definition of the 1/3-octave bands and the performance requirement of ANSI S1.11 filter banks (ANSI, 2004). It defines 43 1/3-octave bands covering the frequency range of 0-20 kHz. Each 1/3-octave band is specified by its mid-band frequency f_m and the bandwidth Δf. The nth band's mid-band frequency $f_m(n)$ is defined by:

$$f_m(n) = 2^{\left(\frac{n-30}{3}\right)} \times f_r \qquad (1)$$

where, f_r is the reference frequency set to 1 kHz. Two band edge frequencies lower $f_1(n)$ and upper $f_2(n)$ of the nth band are determined by $f_1(n) = f_m(n) \times 2^{-1/6}$ and $f_2(n) = f_m(n) \times 2^{1/6}$. The bandwidth of nth band can be calculated by:

$$\Delta f(n) = f_2(n) - f_1(n) \qquad (2)$$

ANSI S1.11 standard describes three classes of filters: class 0, 1 and 2. For each filter class, the performance requirements are specified in terms of parameters such as relative attenuation, linear operating range, environment sensitivities (e.g., humidity and temperature), maximum output signal, terminating impedance and so on. Figure 2 illustrates the ANSI S1.11 class-2 filter specification for the nth 1/3-octave band where $M_n(w)$ and $m_n(w)$ denote the limits on the minimum and maximum attenuations of the nth band filter, respectively. It shows that a class-2 filter allows a 1 dB ripple in its pass-band region (i.e., $f_1 \sim f_2$) and at least 60 dB attenuation at frequencies lower than f_1' (= $f_m \times 0.184$) and at frequencies $> f_2'$ (= $f_m \times 5.434$). The relative
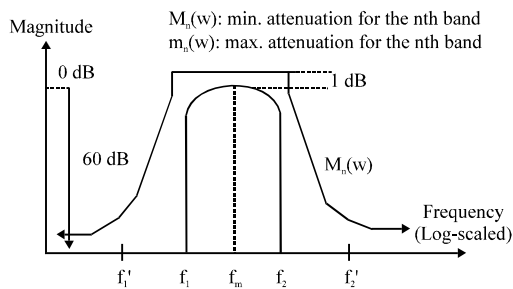
attenuation is an important design parameter from the hardware implementation point of view because it strongly relates to the filter order. The specifications of class-0 and class-1 filters are stricter than class-2 filters. Here, 18-band filter bank (Kuo *et al.*, 2007) for the 22nd~39th band (covering 140-8980 Hz) is designed based on the class-2 specifications with sampling rate of 24 kHz in order to provide good sound quality.

MULTI-RATE FILTER BANK DESIGN

The complexity effective design of an 18-band ANSI S1.11 filter bank (Kuo *et al.*, 2010) which implements the 22nd-39th 1/3-octave bands (denoted by $F_{22} \sim F_{39}$) for hearing aids is shown in Fig. 3. Symbols x and y stand for the input and output sequences in the auditory compensation system. The 18 bands $F_{22} \sim F_{39}$ cover six octaves which are indexed by k = 1, 2, ... , 6, starting from high frequencies towards low frequencies. With the filter bank, the input x(n) is first decomposed into 18 frequency-selected outputs, $y_{22} \sim y_{39}$. They are separately amplified by the prescribed insertion gains and then processed by the compressor. Finally, the outputs $y_{22}' \sim y_{39}'$ are combined to form the output sequence y(n). For simplicity, if the filter F_n, $22 \leq n \leq 39$, is ideal, then the following magnitude response is obtained:

$$\left|F_n(e^{jw})\right| = \begin{cases} 1 & f_1(n) \leq w \leq f_2(n) \\ 0 & \text{otherwise} \end{cases} \qquad (3)$$

where, $f_1(n)$ and $f_2(n)$ are the upper and lower band edge frequencies of the nth octave band. According to ANSI specification, the highest frequency $f_2(36)$ in the 2nd octave is below $\pi/2$ when the sampling rate is set to 24 kHz. Since, each 1/3 octave band in the 2nd octave is band limited. The straightforward FIR design of the 1/3-octave filter bank requires extremely high order filters
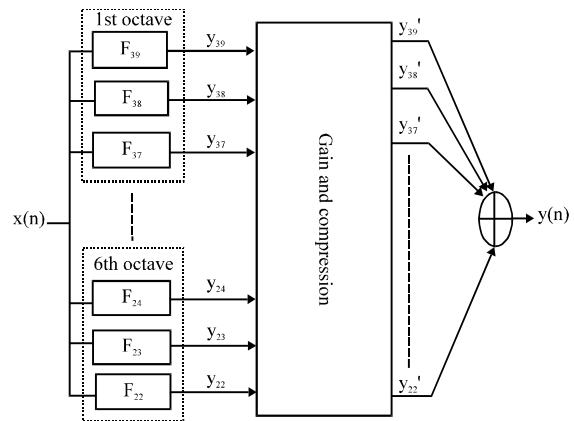

Fig. 2: ANSI S1.11 class-2 filter specification (Ho and Young, 2004)
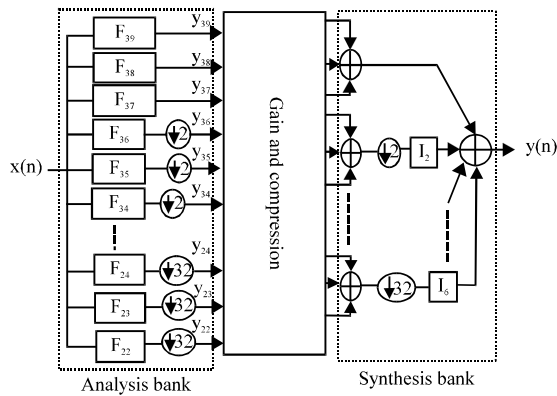

Fig. 3: ANSI S1.11 18-band parallel filters

Fig. 4: Multi-rate filter bank of eighteen filters

Fig. 5: Analysis bank

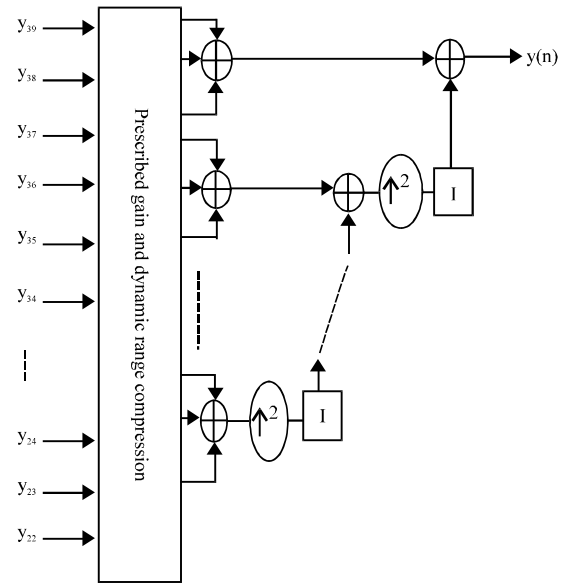Fig. 6: Dynamic range compression of synthesis bank

The proposed multi-rate filter bank minimizes the sampling rate of the band-limited channels to reduce the computation complexity of the filter bank and compressor.

DESIGN OF FILTER COEFFICIENTS

To realize the multi-rate filter bank in Fig. 4, the orders of filters F_{37}, F_{38}, F_{39} and D should be determined as small as possible for reducing the computational complexity. First, the optimal values of pass-band ripple, stop-band attenuation, pass-band and stop-band frequencies for F_{37}, F_{38}, F_{39}, D and I are calculated in such a way that all eighteen 1/3-octave bands adhere to ANSI S1.11 specifications and to minimize the overall complexity. According to the ANSI S1.11 class-2 filter specifications, the pass-band ripple and the stop-band attenuation are 1 and -60 dB, respectively. Since, the orders of the designed filters are dominated by the filters' transition bandwidths, the problem of finding fs1, fp1, fp2 and fs2 is then translated into finding the transition bandwidths (denoted as TBW1 and TBW2), fs1 and fs2 where TBW1 = fp1-fs1 and TBW2 = fs2-fp2. Besides, to simplify the design process, the band-pass filter specifications of F_{37}~F_{39} are decomposed into low-pass and high-pass specifications and the range of TBW1 and TBW2 are found separately. An exhaustive search of the pass-band and stop-band frequencies in their possible ranges is to be made to obtain optimal results. The filter sampling is also suppressed. Though, the coefficient set of F_{37}~F_{39} determined in the first step satisfy the specifications, the total ripple of D, I and F_{37}~F_{39} could make the bands

since the bandwidths of the bands in small octaves (e.g., F_{22}) are very narrow. To overcome these problems, the filter bank is designed with a Multi-Rate algorithm which is shown in the Fig. 4. This multi-rate filter bank only implements the highest octave's three filters (i.e., F_{37}, F_{38} and F_{39}) and a low-pass decimation filter (denoted by D). Consequently by the IFIR techniques and the noble identity, researchers can implement the filters F_{22}~F_{36} and I2~I6 with filters F_{39}, F_{38}, F_{37}, D and I.

According to the definition of 1/3-octave bands, the bandwidth of F_n is exactly half of that of F_{n+3}. Therefore, each octave has the identical frequency characteristics if the sampling rate is reduced by 2 for every octave. Hence, the multi-rate filter bank could iteratively calculate the lower octaves with F_{37}~F_{39}, when the input signal is band-limited by D and down-sampled by 2 for each octave as shown in Fig. 5.

The synthesis block would contain an interpolation filter (denoted as I) to suppress the imaging distortion caused by up sampling. The dynamic range compression of the synthesizer is depicted in Fig. 6.

22~36 violate the specifications. Thus, the 2nd step refines the coefficients so that ripple in each band response does not violate the specifications. F_{37}~F_{39} are redesigned with different ripples. For each ripple used in F_{37}~F_{39}, the maximum possible value of the ripple of D and I are determined to make the 22nd~36th bands meet the specifications. The final solution consists of the best result with minimal computation complexity where 41, 33, 27, 35 and 41 tap filters were obtained for F_{37}, F_{38}, F_{39}, D and I, respectively.

VLSI IMPLEMENTATION OF ANSI FILTER BANK

The proposed ANSI S1.11 1/3-octave filter bank contains six octaves and each octave is based on the identical set of filters F_{37}, F_{38}, F_{39} and D as illustrated in Fig. 5. For an area-efficient implementation, the octaves are folded as demonstrated in Fig. 7. However, this incurs the scheduling problem. A good scheduling algorithm should avoid computation conflicts and minimize the required storage elements. Here, the Recursive Pyramid Algorithm (RPA) is considered. According to RPA, the first octave is calculated every other time slot, the second octave is calculated every four time slots, the third every eight time slots and so on. Figure 8 shows the computation scheduling of the folded multi-rate filter bank in which the variables y_n and x_k are the output of the nth band and input of the kth octave, respectively as illustrated in Fig. 6.

Based on the RPA scheduling, researchers design the proposed filter bank with the hardware block diagram shown in Fig. 9 which consists of modules of the memory controller, the memory block, the Multiply and Accumulate (MAC) unit.

The memory block contains a read only memory that stores the coefficients of filters F_{37}, F_{38}, F_{39} and D. The datapath of the MAC unit is shown in Fig. 10 which consists of a multiplier adder and a tmp register together with four accumulators, namely acc_F_{37}, acc_F_{38}, acc_F_{39} and acc_D, respectively. The MAC unit performs the FIR filtering computations of the four filters: F_{37}, F_{38}, F_{39} and D in an interleaved style. For instance, the computations of the ith tap of these four filters are carried out as follows. Input is first read out and stored in the tmp register. Then, in order to exploit the coefficient symmetry property of linear phase FIR filters, again input is read out and accumulated into the one stored in the tmp register. During the 3rd-6th cycles, the ith coefficients of the four filters are consecutively read out from the coefficient ROM. They are then multiplied with the value stored in
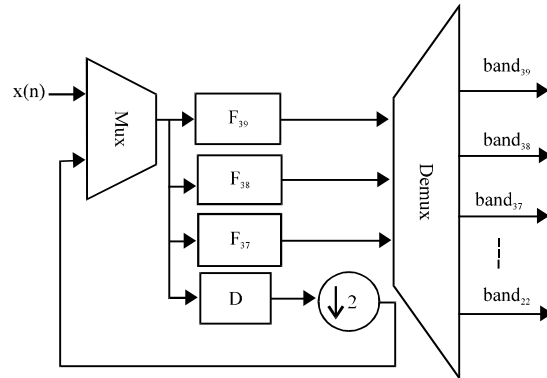

Fig. 7: Filter bank with recursive structure
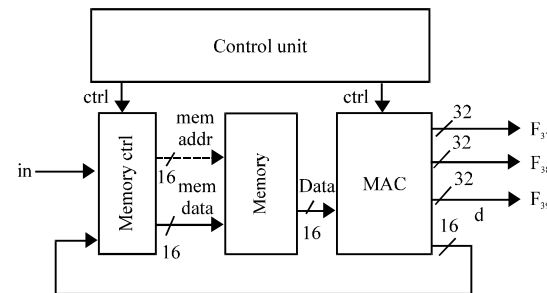

Fig. 8: Recursive Pyramid algorithm
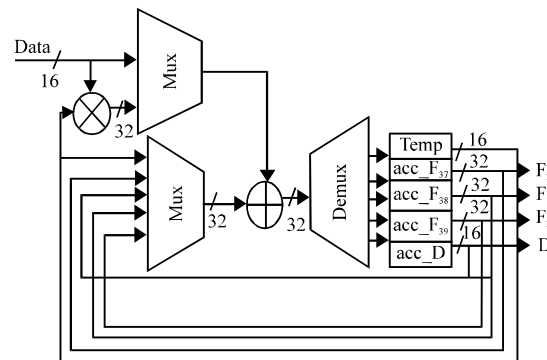

Fig. 9: Architecture of filter bank

Fig. 10: Datapath of the MAC unit

the tmp register and the results are, respectively accumulated into the corresponding accumulator registers: acc_F_{37}, acc_F_{38}, acc_F_{39} and acc_D.

Hence, after performing the computations of each tap, the four accumulators, acc_F_{37}, acc_F_{38}, acc_F_{39} and acc_D would contain the outputs of the filters F_{37}, F_{38}, F_{39} and D, respectively.

LOW POWER TECHNIQUES

The power consumption is indeed a critical issue for hearing aids. Here, researchers focus on minimizing the dynamic power which accounts for the major part of the total power consumption in the simulation. Multipliers and adders account for most of the power consumed by the device. A multiplier with low power consumption is chosen for the design by considering various existing multipliers. Instead of using conventional adders in the design of multiplier, minority function based full adders are used which has less number of transistors when compared to a conventional CMOS full adder. This results in power and area minimization of the multiplier design.

Multipliers: Computational performance of a hearing aid system is limited by its multiplication performance and since multiplication dominates the FIR filter operation time. Good multipliers are in extremely high demand due to its speed and low power consumption. To reduce dynamic power consumption, a suitable low power multiplier design is opted.

Array multiplier is an efficient layout of a combinational multiplier. In array multiplier, consider two binary numbers A and B, of m and n bits. There are mxn summands that are produced in parallel by a set of mxn AND gates. Though, array multiplier demands optimum power and components, the delay is large. Due to large area requirement and delay, it is not suitable for low power systems.

The modified booth multiplier scans three bits at a time to reduce the number of partial products. The two bit from the present pair and a third bit from the high order bit of an adjacent lower order pair. Modified Booth algorithm is widely used to implement multiplication in various applications. It provides high performance than other multiplication algorithms. The method of booth recording reduces the numbers of adders and hence the delay required to produce the partial sums by examining three bits at a time (Table 1). Wallace tree consists of carry save adders arranged as shown in the Fig. 11.

A fast process for multiplication of two numbers was developed by Wallace. The basic idea in Wallace multiplier is that all the partial products are added at the same time instead of adding one at a time. The n bit CSA consists of disjoint Full Adders (FAs). It consumes 3 bit

Table 1: Modified booth encoding

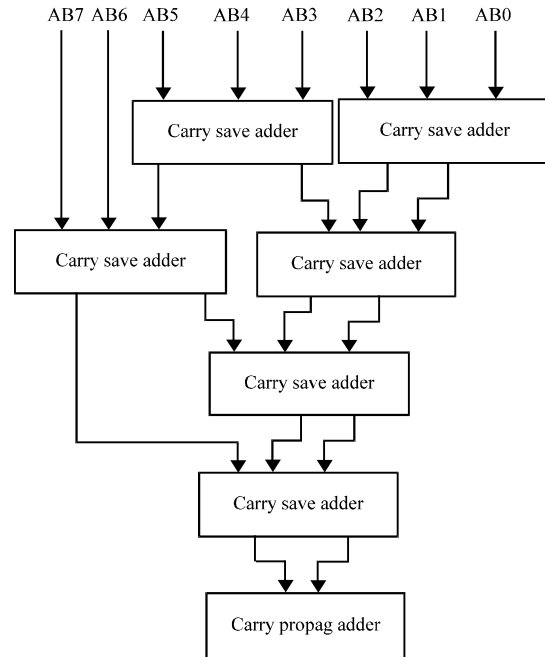| b_{2i+1}	b_{2i}	b_{2i-1}	Operation	Neg_i	Two_i	One_i
0	0	0	0	0	0	0
0	0	1	+y	0	0	1
0	1	0	+y	0	0	1
0	1	1	+2y	0	1	0
1	0	0	-2y	1	1	0
1	0	1	-y	1	0	1
1	1	0	-y	1	0	1
1	1	1	0	0	0	0

Fig. 11: Wallace tree multiplier

input vectors and produces n bit sum S and carry C vectors. Consequently, the CSA has the same propagation delay as only one FA delay and the delay is constant for any values of n. For sufficiently large n, the CSA implementation becomes much faster. This speeds up the multiplication process.

Adder circuit: Low power full adders (Shams *et al.*, 2002; Rouholamini *et al.*, 2007) are required to implement the multipliers in filter in order to reduce the power consumption of hearing aid design. Comparing to the XOR implementations of full adder cells, majority based full adders (Ibrahim *et al.*, 2008) are more reliable and robust. Moreover, the bridge style circuits by sharing transistors can operate faster and are smaller when compared to the conventional circuits but when designing the whole circuit of full adder with this style increases the number of transistors which results in more power consumption, delay and area wastage. The functionality of the full adder cell is based on the following equations:

$$\overline{Sum} = \overline{Cout}(A+B+ABCin)+ABCin \qquad (4)$$

$$\overline{Cout} = Minority \ (ABCin) \qquad (5)$$

Thus, the cout function can be implemented using minority circuit from which the sum function can be determined.

The minority function acts as follows: if the number of '0's becomes greater than the number of '1's at the input, the output will be '1's. Minority is a function of odd number of inputs. The full adder cell (MBFA) (Navi *et al.*, 2011) is designed by using a 3 input minority circuit, followed by a bridge style structure as shown in Fig. 12. The truth table of a full adder, shown in Table 2, authenticates the basic equations of the design. The minority function acts as follows: if the number of '0's becomes greater than the number of '1's at the input, the output will be '1's. Minority is a function of odd number of inputs. The design uses 12 transistors and is based on majority-NOT gates implemented with high-performance

CMOS bridge circuit. This design uses transistors called bridge transistor and sharing transistors of different paths to generate new paths from supply lines to circuit outputs. The bridge style design is completely symmetric providing regularity and better performance than other CMOS design styles.

Using the bridge circuit which leads to the reduction of delay and power consumption of the full adder cell it also increases the robustness and reliability of the circuit. Using this adder in multipliers reduces the power consumption, area and delay of multipliers.

Clock gating: The clock gating is one of the mostly used low power technique at system level. Non active hardware units are shut down to save power. Disabling the clock signal (clock gating) in inactive portions of the chip is a useful approach for power dissipation reduction as shown in Fig. 13.

Clock gating can be applied to different hierarchical levels. This technique is widely used for power reduction. Independent blocks when clock gated, it greatly affects power savings because gating larger blocks achieves in higher power savings in 'off' clock cycle. Considering the datapath of the MAC unit, it consists of four accumulator acc_F_{37}, acc_F_{38}, acc_F_{39} and acc_D. These accumulator registers consume significant power indeed. Since, the register dissipates power at each transition of the clock signal, it would waste power if its value need not be updated. Disabling the clock signal when module is inactive reduces power consumption of the entire circuit by about 15%.
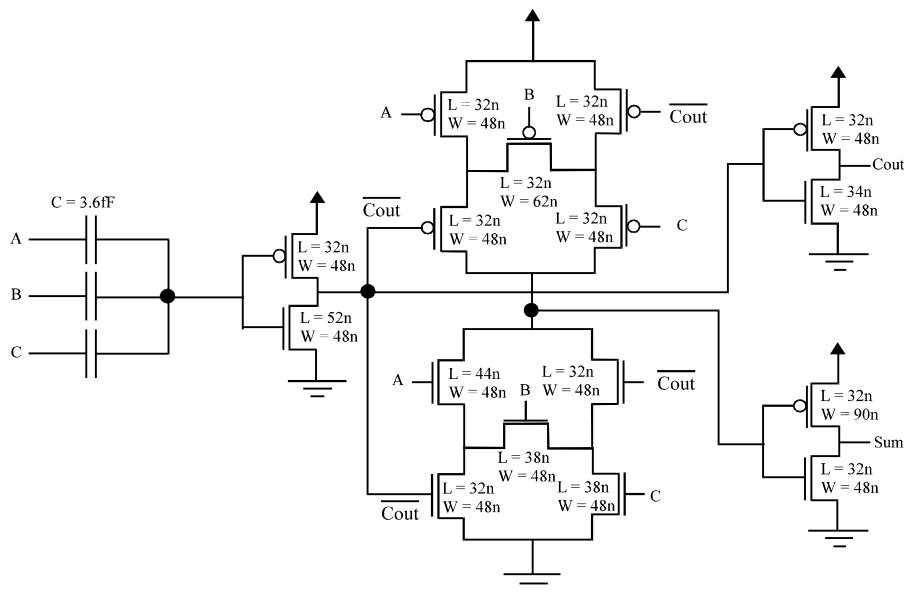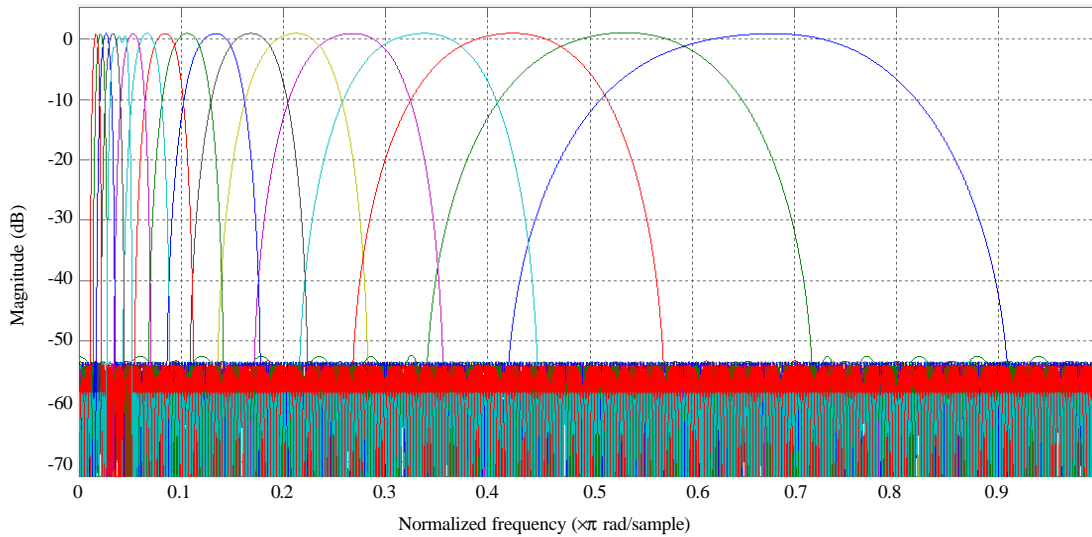
Table 2: Minority based full adder cell truth table

A	B	Cin	$\overline{Cout}$	Minority (A, B, Cin)	$\overline{Sum}$
0	0	0	1	1	1
0	0	1	1	1	0
0	1	0	1	1	0
0	1	1	0	0	1
1	0	0	1	1	0
1	0	1	0	0	1
1	1	0	0	0	1
1	1	1	0	0	0

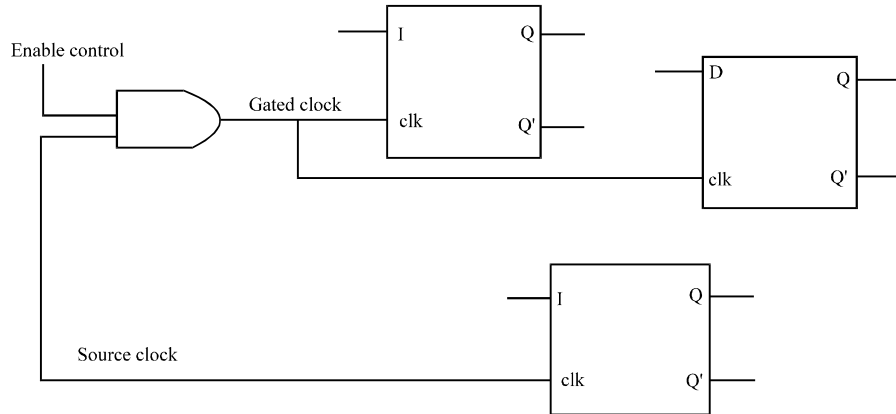Fig. 12: MBFA

Fig. 13: Clock gating structure

Fig. 14: Magnitude response of eighteen parallel filters

Table 3: Comparison results

Performance of multiplier	Array	Array (MBFA)	Modified booth	Modified booth (MBFA)	Booth Wallace tree	Booth Wallace tree (MBFA)
No. of cells	967	824	859	781	869	806
Total area (µm²)	10947	8531	9531	7942	9662	8137
Leakage power (mW)	0.0458	0.0194	0.0355	0.0264	0.0357	0.0516
Dynamic power (mW)	0.3773	0.2521	0.3021	0.2303	0.3140	0.2102
Total power (mW)	0.4232	0.2715	0.3375	0.2567	0.3497	0.2618
Power (with clock gating) (mW)	0.3864	0.2513	0.3062	0.2315	0.3157	0.2407
Propagation delay (psec)	10406	8287	9487	7952	7555	7369

RESULTS AND DISCUSSION

The auditory compensation contains filter bank which is implemented using 1/3 octave multi rate FIR filter bank with systematic coefficient. Figure 14-16 show the magnitude response of filters F_{39}, F_{38}, F_{37} and D designed. The 18 parallel filters require 3329 multipliers where it consumes more power during implementation. Multi-Rate algorithm requires only 135 multipliers as compared to straightforward method which is 4% computational complexity.

Using Multi-Rate algorithm power consumption is also reduced. The performance of different multipliers have been compared in Table 3. It is obvious that using MBFA instead of full adder results in significant reduction in power, area and delay.
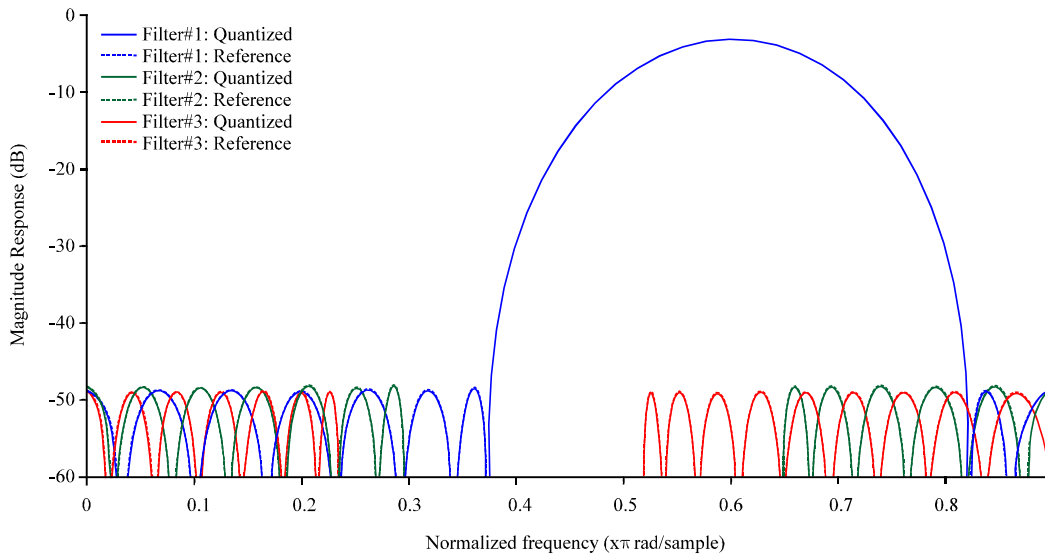
280

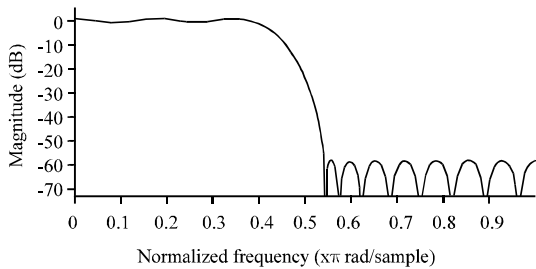Fig. 15: Magnitude response of 39, 38 and 37th filters

Fig. 16: Magnitude response of D filter

Using Modified Booth algorithm power consumption of MAC unit gest reduced. Instead of using booth using booth wallace tree will reduce the power consumption and also reduce the propagation delay so the speed of the filter will increases. Using this method, power and propogation delay gets reduces as compared to Array Multiplier Method.

CONCLUSION

In this study, a multi-rate architecture to implement the ANSI S1.11 filter bank with FIR filters for digital hearing aids is proposed where the computations can be greatly reduced with the multi-rate processing. Also, a systematic flow has been proposed to optimize the FIR coefficients for the multi-rate filter banks to minimize their orders. So, the proposed architecture needs only 4% of additions and multiplications of a straight forward method. Using some power optimization techniques researchers can further reduce power during

implementation. This method helps to design better hearing aids with reduced computation complexity and 30% lower power consumption.

REFERENCES

ANSI, 2004. Specification for octave-band and fractional-octave-band analog and digital filters. ANSI S1.11-2004, Standards Secretariat Acoustical Society of America, February, 2004.

Chong, K.S., B.H. Gwee and J.S. Chang, 2006. A 16-channel low power non-uniform spaced filter bank core for digital hearing aid. IEEE Trans. Circ. Syst. II: Express Briefs, 53: 853-857.

Ho, M.S. and S.T. Young, 2004. Building a platform of hearing aid's compensative strategy for Mandarin. Proceedings of the 18th International Congress on Acoustics, April 4-9, 2004, Kyoto, Japan, pp: 831-834.

Ibrahim, W., V. Beiu and M.H. Sulieman, 2008. On the reliability of majority gates full adders. IEEE Trans. Nanotechnol., 7: 56-67.

Kuo, Y.T., T.J. Lin, Y.T. Li, W.H. Chang and C.W. Liu, 2007. Design of ANSI S1.11 filter bank for digital hearing aids. Proceedings of 14th IEEE International Conference on Electronics, Circuits and Systems, December 11-14, 2007, Marrakech, pp: 242-245.

Kuo, Y.T., T.J. Lin, Y.T. Li, W.H. Chang and C.W. Liu, 2010. Design and implementation of low power ANSI S1.11 filter bank for digital hearing aids. IEEE Trans. Circ. Syst. I: Regular Papers, 57: 1684-1696.

Lozano, A. and A. Carlosena, 2003. DSP-based implementation of an ANSI S1.11 acoustic analyzer. IEEE Trans. Instrum. Meas., 52: 1213-1219.

Navi, K., H.H. Sajedi, M.H. Moaiyeri, A. Jalali and O. Kavehei, 2011. High speed full adder based on minority function and bridge style for nanoscale. Integr. VLSI J., 44: 155-162.

Rouholamini, M., O. Kavehei, A. Mirbaha, S. Jasbi and K. Navi, 2007. A new design for 7:2 compressors. Proceedings of the ACS/IEEE International Conference on Computer Systems and Applications, May 13-16, 2007, Amman, pp: 474-478.

Shams, A.M., T.K. Darwish and M.A. Bayoumi, 2002. Performance analysis of low power 1 bit CMOS full adder cells. IEEE Trans. Very Large Scale Integr. Syst., 10: 20-29.