e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

FIR IMPLEMENTATION USING FILTER STRUCTURES APPROACH WITH FIXED POINT CONVERSION TECHNIQUE

Daljeet Kaur Khanduja[*1]

[*1]Professor In Mathematics, Sinhgad Academy Of Engineering, Kondhwa, Pune, Maharashtra, India.

DOI : https://www.doi.org/10.56726/IRJMETS45534

ABSTRACT

The design and implementation of FIR filter banks for multirate analysis and synthesis are examined in this paper using a least-Pth norm approach and floating-point arithmetic. However, in order to minimize cost and energy, they are typically implemented using fixed-point arithmetic. This paper discusses the conversion of floating points to fixed points for software implementation while taking the filter structure into consideration. Our strategy aims to identify the fixed-point specification that reduces the amount of time required to execute code for a given accuracy requirement. In our approach, the processor architecture is considered, and the code generation process is integrated with the floating-point to fixed-point conversion process. The arithmetic complexity is then decreased while still adhering to design restrictions by employing various realization structures. The current work aims to reduce the number of memory accesses while boosting the speed of filter computation. As a result, the system's processing speed and space utilization time have both decreased. The effectiveness of two different filter architectures is evaluated and compared. A signal to noise ratio (SNR) of between 50 and 60 dB is observed from a variety of signal samples, which is suitable for signal analysis and synthesis purposes.

Keywords: P^{th} Norm, DSP, FIR, Fixed Point Conversion, SNR.

I. INTRODUCTION

The algorithms used in communication systems and digital signal processing are typically specified with infinite precision (IP) operations at the beginning. Floating-point computations in digital computers have high precisions, these operations are commonly referred to as floating-point in the literature, particularly for circuit designers [1-3]. On the other hand, digital implementations of these techniques rely on limited precision (LP) or finite precision (FP) approximations. Binary numbers, including 2's complement and binary unsigned numbers, are the most widely used fixed-point realizations for finite precision number systems in hardware implementation systems [4–9]. Overflow on the most-significant bit (MSB) occurs when a number cannot be represented by a certain fixed point data type with finite word length. or quantization on the least-significant-bit (LSB) or both take place.

To meet the requirements for cost and power consumption, fixed-point architectures are used to implement digital signal processing algorithms that were originally specified and created with floating-point data types. Fixed-point architectures have narrower memory and bus widths, which significantly reduces cost and power usage. Additionally, the mantissa and exponent are more difficult to process using floating-point operators. As a result, the area and latency of floating-point operators are higher than those of fixed-point operators. In this situation, the application specification needs to be fixed-point transformed. The manual conversion procedure is a laborious and error-prone task that extends the time needed for development. Several studies [10] have demonstrated that this manual conversion can account for up to 30% of the overall deployment time. So, in order to speed up development, automatic floating-to-fixed-point conversion procedures are needed.

One of the essential building elements in many applications of digital signal processing (DSP) is the FIR filter. The direct and cascade FIR filter structures are the most frequently employed in digital filter construction; the former is simpler and thus easier to build. However, due to the huge dynamic range of the coefficients, the direct version is typically more sensitive to the effects of coefficient quantization in fixed point implementation. On the other hand, the cascade form reduces dynamic range and hence decreases sensitivity, but the realization is more challenging because it requires scaling the coefficient and proper section ordering to prevent overflow

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

and reduce round off noise. The cascade form structure is robust to quantization and roundoff noise, but the direct form structure has less hardware complexity

II. DESIGN FORMULATION

A) The P^{th} Norm and Infinity Norm

A nonlinear phase design that minimizes any norm from 2 (minimum error energy) to infinity (minimax/equiripple error) is provided by least P^{th} norm. The resulting design approach is exceedingly versatile, computationally effective, and enables the realization of a wide range of trade-offs. It also allows for the direct application of numerous combinations of restrictions. A (near) minimax design is produced using this method by minimizing a weighted error function without constraints, where the weighting function is maintained during the optimization process and the power p is assumed to be an even integer.

B) Design Concept of Filter Design With Optimal Approach

Let $H_d(w)$ be the desired frequency response over a frequency range Ω which is a compact subset of $[0,\pi]$, and $H(w)$ be the m^{th} - order transfer function given by

$$H(w) = h_0 \frac{1 + b^T f}{1 + a^T f} \quad (1)$$

$$f = [z^{-1}\ z^{-2} \ldots\ z^{-m}]^T \quad (2)$$

$$a = [a_1\ a_2\ \ldots\ a_m]^T \quad (3)$$

$$b = [b_1\ b_2\ \ldots\ b_m]^T$$

And $z = e^{jw}$ with w the normalized frequency, that is $0 \le w \le \pi$. The filter design problem is to find real valued M dimensional vector $(M = 2m+1)$

$$X = [h_0\ \ a_1\ \ \ldots a_m\ \ b_1 \ldots\ b_m]^T \quad (4)$$

that minimizes the P^{th} norm objective function

$$E(x) = \int_{\Omega} |H(w) - H_d(w)|^p dw \quad (5)$$

Where $p \ge 2$ is an even integer and that the corresponding filter (5) is stable.

The least P^{th} norm algorithm can be used to design FIR/ IIR filters. The following equation shows the frequency response of an IIR filter with N zeros and M poles.

$$H(w) = \frac{B(w)}{A(w)}$$

$$= \frac{\sum_{n=0}^{N} b(n) e^{-jwn}}{1 + \sum_{n=1}^{M} a(n) e^{-jwn}} \quad (6)$$

Where $B(w)$ is the Fourier transform of the forward coefficients.

Where $A(w)$ is the Fourier transform of the backward coefficients.

$b(n)$ is the set of forward coefficients

$a(n)$ is the set of reverse coefficients

When M equals zero, the IIR filter reduces to a FIR filter. Usually a (0) is normalized to 1 as shown in equation (6).

The least P^{th} norm algorithm uses complex approximation or magnitude approximation to create the design. The following equation is the complex approximation.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

$$\|E\|_p = \left(\sum_{i=0}^{L-1} \left(W(i) |H(w_i) - D(w_i)| \right)^p \right)^{\frac{1}{P}}$$

Where W(i) is a positive weight at the i^{th} frequency point

H is the response of the designed filter

D is the target response

L is the number of frequency points used to perform the calculation

p is the P^{th} norm

C) Implementation

The incorporation of filter banks for signal analysis and synthesis in order to potentially achieve improved time processing is the first stage in the methodology covered in this study. The actual implementation of the method mentioned in this study is shown in the block diagram in figure (1).

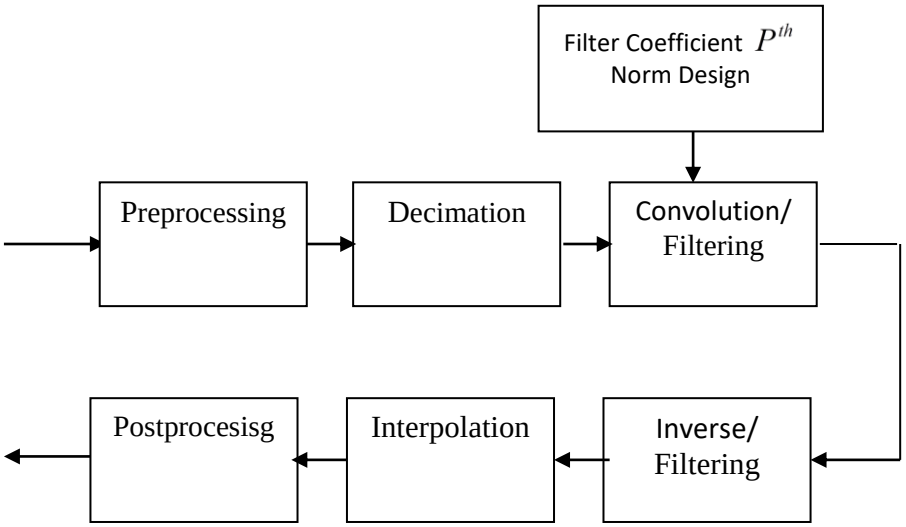

Fig.1 Block diagram of signal Using P^{th} norm filter design

- **Input signal wave:**

The input wave is low pass filtered and pre-emphasized with an 8 KHz sampling rate.

- **Preprocessing:**

Various digital filters, including Low Pass Filters (LPF) and Notch Filters to eliminate the 50Hz. Line frequency impact, are used in this block to remove undesirable noise from the data.

- **Decimation**

The decimation factor assists in improving the band level data when dealing with multi-rate analysis systems so that the analysis part can be verified with norm level design. For this experiment, we consider decimation by 2, 4, etc.

- P^{th} **Norm filter design:**

FIR filters typically deal with linearity and stability, but when it comes to optimality parameters, the P^{th} norm enters the picture. The specifics of our design for the P^{th} norm filter, which captures the filter coefficients, are covered in section IV for norm.

- **Filtrating/convolution**

The primary objective of filtering is to increase the higher frequencies in order to flatten the spectrum.

- **Inverse filtering:**

Analysis characteristics are rearranged in relation to the five-band system to retrieve the original contents from the feature vector for signal generation purposes. Our ability to improve signal parameters is aided by the inverse filtering outcome.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

- **Interpolation:**

The result of the inverse filtering is up sampled in interpolation. It's a step in the right direction for recovering the original signal parameters.

- **Post Processing**

Final tuning is carried out in the post processing block.

D) Performance Evaluation

To apply the design approach, MATLAB scripts were created [11]. For every simulation, signals in *.wav file were sampled at an 8 KHz rate. Section 4 presents the outcomes for various configurations. On a mathematical comparison of the unprocessed and processed data, objective measurements are built.

Most objective quality evaluations express signal quality in terms of a numerical distance metric. The most used technique for assessing signal quality is the signal to noise ratio. The ratio of signal to noise, expressed in decibels, is used to compute it.

$$SNR = 10\log_{10}\left(\frac{\sum_n s^2(n)}{\sum_n \left[s(n) - s\hat{}(n)\right]^2}\right)$$

where $s(n)$ is the clean signal and $s\hat{}(n)$ is the processed signal.

III. REALIZATION OF FIR DIGITAL FILTERS

In general, a causal FIR system is described by the difference equation; $y[n] = \sum_{k=0}^{M} b_k x[n-k]$

Or equivalently by the system function $H(z) = \sum_{k=0}^{M} b_k z^{-k}$

The impulse response of the FIR system is identical to the coefficients $\{b_k\}$, that is

$$h[n] = b_n, \quad 0 \le n \le M$$
$$= 0, \quad \text{otherwise}$$

The direct form realization follows immediately from the non-recursive difference equation or by the convolution summation

$$y[n] = \sum_{k=0}^{M} h[k]x[n-k]$$

These algorithms' structure is described as having a "Direct form structure," which is a repeating delay-and-add style. The algorithm (filter) function for FIR (Finite Impulse Response) filters is computed using "n" computation coefficients and a limited number of "n" delay taps on a delay line. FIR filters are most frequently created using the aforementioned structure, which is non-recursive and repeated delay-and-add. Every sample of fresh and present value data is required by this structure.

- **Floating-to-fixed-point conversion methodology**

The idea of a Q point for a fixed point number is presented for the purposes of this essay. This labeling practice is as follows:

Q[QI].[QF]

where [QI] denotes the number of integer bits and [QF] denotes the number of fractional bits.

The number of integer bits [QI] plus the number of fractional bits [QF] yield the total number of bits used to represent the number. Sum [QI] + [QF] is known as the Word Length (WL). A floating-point number must be seen as having two independent components, the integer content and the fractional content, in order to be represented in fixed point. The integer range of a floating-point variable, or its Min to Max range, determines how many bits are needed to represent the integer part of the number in an algorithm. The processes for converting from floating point to fixed point are

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

- Determine the minimum and maximum number for a given floating point number.
- Based on the minimum and maximum value, choose the Q format.
- Use the above Q format to scale the number.
- Verify the saturation and accuracy.
- Finally convert to fixed point number.

Since fixed points operate more quickly than floating point, there are several benefits to switching from floating point to fixed point. Furthermore, floating point variables required more hardware to store, making them expensive in terms of hardware cost, but fixed-point memory blocks are very inexpensive for storing data. Floating point offers us a huge dynamic range while fixed point provides a relatively narrow dynamic range. While floating point is employed in bigger application areas where more accuracy is required, fixed point architecture is generally suitable for portable application devices where minimal precision is acceptable.

IV. RESULTS

This section presents the findings for the five-band scheme in general. The first instance [11] employs a single signal source and a bank of five band filters with a system delay of 54 samples.

Table 1: Specifications of some signal samples.

Signal Name	Tap order	SNR	Elapsed time
S1	54	55.73	0.016
S2	54	56.05	0.032
S3	54(low delay)	55.77	0.031
S4	54(low delay)	59.63	0.078
S5	54 (low delay)	65.19	0.016
S6	54(cosine modulated)	11.82	0.016
S7	54(cosine modulated)	16.04	0.531
S8	54(cosine modulated)	11.71	0.016
S9	54	34.24	0.015
S10	54	27.06	0.015

With a variety of input signals that included signal frequency and sampling frequency variables, we tested the filter structure algorithm. P^{th} norm filter shows the optimal characteristics in terms of filter design aspects like tap length etc., the elapsed time for filtering with the signals indicated in Table 1.

With an overall system time of 0.016 sec and filter coefficients, or 54 tap length, it is possible to observe a good range of SNR values above 50-60dB.

Table 2: Total Delay and Order calculations

Structure	Tap Length	Delays	Multiplier	Additions
Direct Form	55	55	56	55
	96	96	97	96
Cascade	55	220	220	330
	96	384	384	576

Table 2 shows the overall delay and multiplications that take place with which we need to design the P^{th} norm filter with various structuring forms which help us to understand the hardware requirement for the system.

In Table 2 it is observed that cascade structure requires a greater number of multipliers that is why computation order O(n) of this system is more.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

V. CONCLUSION

In this paper, a least P^{th} norm algorithm is investigated using floating point arithmetic for the design and implementation of FIR filter banks for multirate analysis and synthesis. This algorithm produces the best nonlinear phase designs that can minimize any norm from 2 (minimum error energy) to infinity (minimax/equiripple error). The resulting design process is particularly adaptable and enables the direct imposition of a wide range of constraint combinations as well as the realization of a wide range of trade-offs. Fixed-point arithmetic is necessary for effective application implementation. As a result, a method for converting from floating point to fixed point has been discussed. This method reduces the code execution time while maintaining correctness. The DSP architecture is taken into consideration in order to optimize the fixed-point specification in comparison to earlier techniques. Comparing this method to one that relies on simulation, the optimization time is indeed much reduced. We can infer from the tabulated data in the tables above that the norm filter has applications in the signal analysis and synthesis period. In 0.016 seconds, an SNR between 50 and 60 dB is attained from diverse signal samples. Thus, we draw the conclusion that the system will operate consistently with the norm filter and be suitable for signal generation and analysis.

VI. REFERENCES

[1] C. Shi, and R. W. Brodersen, "Floating-point to fixed-point conversion," IEEE Trans. Signal Processing. 2004

[2] C. Shi, and R. W. Brodersen, "An automated floating-point to fixed-point conversion methodology," Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Processing, Vol. 2, pp. 529-532, April 2003.

[3] C. Shi, "Statistical method for floating-point to fixed-point conversion," 2002, Master Thesis, Department of EECS, Univ. of California, Berkeley. (Advisor: Robert W. Brodersen).

[4] A. V. Oppenheim, and R. W. Schafer, with J. R. Buck. Discrete-Time Signal Processing. 2nd ed., Prentice Hall, 1999, ch. 6.

[5] L. B. Jackson. Digital filters and signal processing: with MATLAB exercises, 3rd ed. Boston : Kluwer Academic Publishers, 1996

[6] S. S. Haykin. Adaptive filter theory. 3rd Edition. Prentice Hall, 1996.

[7] R.M. Gray, and D.L. Neuhoff, "Quantization" IEEE Trans. Inform. Theory, vol. 44,No.6,pp.2325-2383,Oct. 98.

[8] D. A. Patterson, and J. L. Hennessy, Computer Organization & Design—the Hardware/software interface, 2nd ed., Morgan Kaufmann, 1998, ch. 4.193

[9] C. Fang, T. Chen, and R. A. Rutenbar, "Floating-point error analysis based on affine arithmetic," Proc. IEEE Int. Conf. on Acoust., Speech, and Signal Processing., vol. 2, pp. 561-564, Apr. 2003.

[10] T. Grotker, E. Multhaup, and O. Mauss, "Evaluation of HW/SW tradeoffs using behavioral synthesis," in Proceeding of 7th International Conference on Signal Processing Applications and Technology (ICSPAT'96), pp. 781–785, Boston, Mass, USA, October 1996.

[11] Time Domain Filter Bank Analysis. A New Design theory. Kambiz Nayebi, Thomas P.Barnwell, Mark J.T. Smith. IEEE Trans. Signal Processing.pp.1412-1429, Vol 40. No.6.