An Efficient Digital FIR Filter Designs Based on Parallel Faithfully Rounded Truncated MCM/A

Miss. S. Anitha. P.G scholar , V.S.B Engineering College.

Abstract:- Finite impulse response (FIR) designs are presented using the concept of faithfully rounded truncated multipliers. System performance is generally determine by the performanceof multiplierWe jointly consider the optimization of bit width and hardware resources without sacrificing the frequency response and output signal precision. Nonuniform coefficient quantization with proper filter order is proposed to minimize total area cost.Multiple constant multiplication/accumulation in a direct FIR structure is implemented using an improved version of truncated multipliers. Here proposed vedic multiplication used for the improve the performance of the multiplication process.

***Index Terms*—Digital signal processing (DSP), faithful rounding,**
finite impulse response (FIR) filter, truncated multipliers,VLSI design.

I. INTRODUCTION

FINITE impulse response (FIR) digital filter is one of the fundamental components in many digital signal processing (DSP) and communication systems. It is also widely used in many portable applications with limited area and power budget.A general FIR filter of order M can be expressed asIn case of linear phase, the coefficients are either symmetric or antisymmetric with $a_i = a_{M-i}$ or $a_i = -a_{M-i}$.There are two basic FIR structures, direct form and transposed form, as shown in Fig. 1 for a linear-phase even-order FIR filter. In the direct form in Fig. 1(a), the multiple constant multiplication (MCM)/accumulation (MCMA) module performs the concurrent multiplications of individual delayed signals and respective filter coefficients, followed by accumulation of all the products. Thus, the operands of the multipliers in MCMA are delayed input signals $x[n - i]$ and coefficientsa_i. In the transposed form in Fig. 1(b), the operands of the multipliers in theMCM module are the current input signal $x[n]$ and coefficients. The results of individual constant multiplicationsgo through structure adders (SAs) and delay elements.,

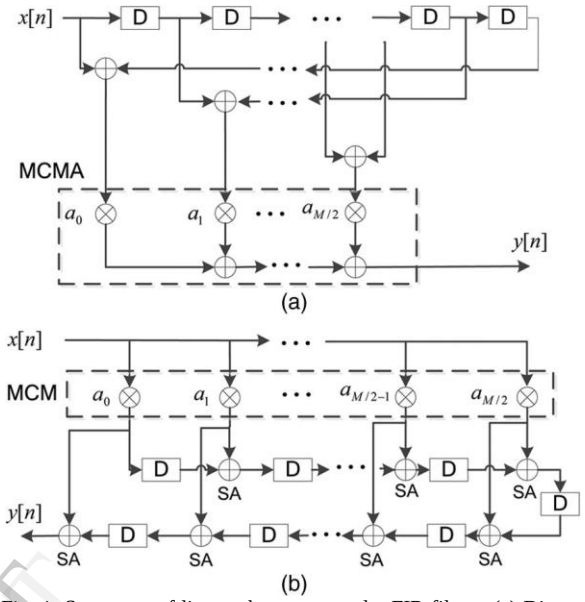


Fig. 1. Structures of linear-phase even-order FIR filters: (a) Direct form and (b) transposed form.

In the past decades, there are many papers on the designs and implementations of low-cost or high-speed FIR filters [1]–[13],[15]–[19]. In order to avoid costly multipliers, most prior hardware implementations of digital FIR filters can be divided into two categories: multiplierless based and memory based.Multiplierless-based designs realize MCM with shift-and add operations and share the common suboperations using canonical signed digit (CSD) recoding and common subexpression elimination (CSE) to minimize the adder cost of MCM [1]–[10]. In [18] and [19], more area savings are achieved by jointly considering the optimization of coefficient quantization and CSE. Most multiplierless MCM-based FIR filter designs use the transposed structure to allow for cross-coefficient sharing and tend to be faster, particularly when the filter order islarge. However, the area of delay elements is larger compared with that of the direct form due to the range expansion of the constant multiplications and the subsequent additions in theSAs. In [17], Blad and Gustafsson presented high-throughput (TP) FIR filter designs by pipelining the carry-save adder trees in the constant multiplications using integer linear programming to minimize the area cost of full adders (FAs), half adders (HAs), and registers (algorithmic and pipelined registers).

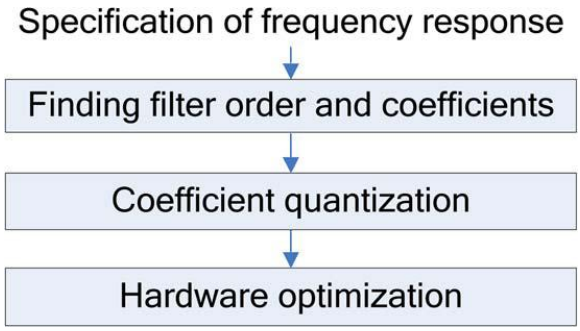


Fig. 2. Three stages in digital FIR filter design and implementation

Memory-based FIR designs consist of two types ofapproaches: lookup table (LUT) methods and distributed arithmetic(A) methods [11]–[13]. The LUT-based design storesin ROMs odd multiples of the input signal to realize the constant multiplications in MCM [11]. The DA-based approaches recursively accumulate the bit-level partial results for the inner product computation in FIR filtering [12], [13]. An important design issue of FIR filter implementation is the optimization of the bit widths for filter coefficients, which has direct impact on the area cost of arithmetic units and registers. Moreover, since the bit widths after multiplications grow, many DSP applications do not need full-precision outputs. Instead, it is desirable to generate faithfully rounded outputs where the total error introduced in quantization and rounding is no more than one unit of the last place (ulp) defined as the weighting of the least significant bit (LSB) of the outputs. In this brief, we present low-cost implementations of FIR filters based on the direct structure in Fig. 1(a) with faithfully rounded truncated multipliers.

The MCMA module is realized by accumulating all the partial products(PPs) whereunnecessary PP bits (PPBs) are removed without affecting thefinal precision of the outputs. The bit widths of all the filter coefficients are minimized using nonuniform quantization with unequal word lengths in order to reduce the hardware cost whilestill satisfying the specification of the frequency response.This brief is organized as follows. Section II discusses the nonuniform quantization and optimization of filter coefficients. Section III describes the PP generation and compression in the faithfully rounded MCMA module. Section IV compares the experimental results.experimental results.

II. Coefficient Quantization and Optimization

A generic flow of FIR filter design and implementation can be divided into three stages: finding filter order and coefficients,coefficient quantization, and hardware optimization, as shown in Fig. 2. In the first stage, the filter order and the corresponding coefficients of infinite precision are determined to satisfy the specification of the frequency response. Then, the coefficients are quantized to finite bit accuracy. Finally, various optimization approaches such as CSE are used to minimize the area cost of hardware implementations. Most prior FIR filter implementations focus on the hardware optimization stage.

After FIR filter operations, the output signals have larger bit width due to bit width expansion after multiplications. In many practical situations, only partial bits of the full-precision outputs are needed. For example, assuming that the input signals of the FIR filter have 12 bits and the filter coefficients are quantized to 10 bits, the bit width of the resultant FIR filter output signals is at least 22 bits, but we might need only the 12 most significant bits for subsequent processing.

ALGORITHM IN MATLAB:

In this brief, we adopt the direct FIR structure with MCMA because the area cost of the flip-flops in the delay elements is smaller compared with that of the transposed form. Furthermore, we jointly consider the three design stages in Fig. 2 in order to achieve more efficient hardware design with faithfullyrounded output signals.

Unlike conventional uniform quantization of filter coefficients with equal bit width, the nonuniform quantization technique with possibly different bit widths is adopted in this brief. Fig. 3 shows the pseudocode of the proposed quantization scheme.

Initially, subroutine **Parks_McClellan**() is used to find the filter order *M* for the given frequency response. Step 1 of uniform quantization starts with calling the MATLAB built-infunction **remez**() to find the coefficients for the FIR filter of order *M*. Then, we quantize the coefficients with enough bit and generate the set of uniformly quantized coefficients a_i with equal bit width *B*. The subroutine **freq_resp_satisfied**() checks if the frequency response is still satisfied after quantization.

```
Given specification of frequency response;
M = Parks_McClellan();  // Find FIR filter order M

// Step 1: Uniform Quantization
area_min = MAX;
while () {
   coeff = remez(M);    // Find FIR filter coefficients for given M
   B ← 0;
   while ( freq_resp_satisfied(B) = 0)
      B ← B+1;          // Quantize coefficients to B bits
   Select CSD or radix-4 Booth recoding with fewer nonzero digits;
   area = area_cost_estimate( B );  // estimate total area cost of FIR filter
   If (area < area_min) {
      area_min = area;
      M ← M+1;}
   else {
      M ← M−1;
      break; }
}

// Step 2: Non-uniform Quantization
N = { M/2      M is even  // N non-redundant coeffs. for linear phase FIR
    { (M+1)/2  M is odd
B_i ← B, i = 0,1,···,N−1;
while () {
   reduction = 0;
   for (i=0 to N-1)
      if ( freq_resp_satisfied( B_i−1 ) = 1) {
         B_i ← B_i−1;
         reduction = 1; }
   if (reduction = 0 )  break;  // no further bit-width reduction is possible
}

// Step 3: Coefficient Fine-tune
while () {
   reduction = 0;
   for (i=0 to N-1) {
      a_i ← a_i + 2^−B_i ;      // increase each coefficient value by 1 ulp
      if ( freq_resp_satisfied( B_i−1 ) = 1) {
         B_i ← B_i−1;
         reduction = 1;
         continue; }     // continue fine-tune for next coefficient
      a_i ← a_i − 2^−B_i ;      // decrease each coefficient value by 1 ulp
      if ( freq_resp_satisfied( B_i−1 ) = 1) {
         B_i ← B_i−1;
         reduction = 1; }
   }                              // end of for loop for every coefficients
   if (reduction = 0 )  break;  // no bit width reduction is possible
}
```

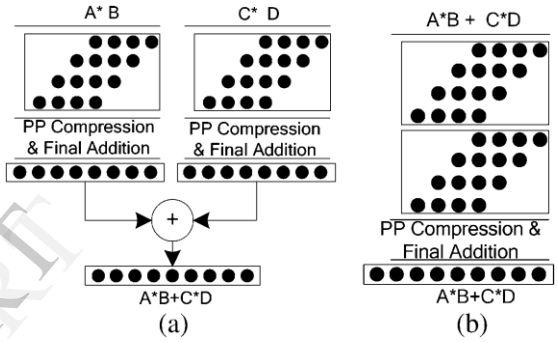
Fig. 3. Proposed algorithm of coefficient quantization and fine tuning.

After coefficient quantization, we perform recoding to minimizethe number of nonzero digits. In this brief, we consider CSD recoding with digit set of $\{0, 1, -1\}$ and radix-4 modifiedBooth recoding with digit set of $\{0, 1, -1, 2, -2\}$ and select theone that results in smaller area cost.While most FIR filter designs use minimum filter order,we observe that it is possible to minimize the total area byslightly increasing the filter order. Therefore, the total area of the FIR filter is estimated using the subroutine **area_cost_ estimate**() using the approach in [20]. Indeed, the total number of PPBs in the MCMA is directly proportional to the number of FA cells required in the PPB compression because a FA reduces one PPB.

After Step 1 of uniform quantization and filter order optimization, the nonuniform quantization in Step 2 gradually reduces the bit width of each coefficient until the frequency response is no longer satisfied.

Finally, we fine-tune the nonuniformly quantized coefficients by adding or subtracting the weighting of LSB of each coefficient and check if further bit width reduction is possible. Using the algorithm in Fig. 3, we can find the filter order M and the nonuniformly quantized coefficients that lead to minimized area cost in the FIR filter implementation.

III. PP Truncation and Compression



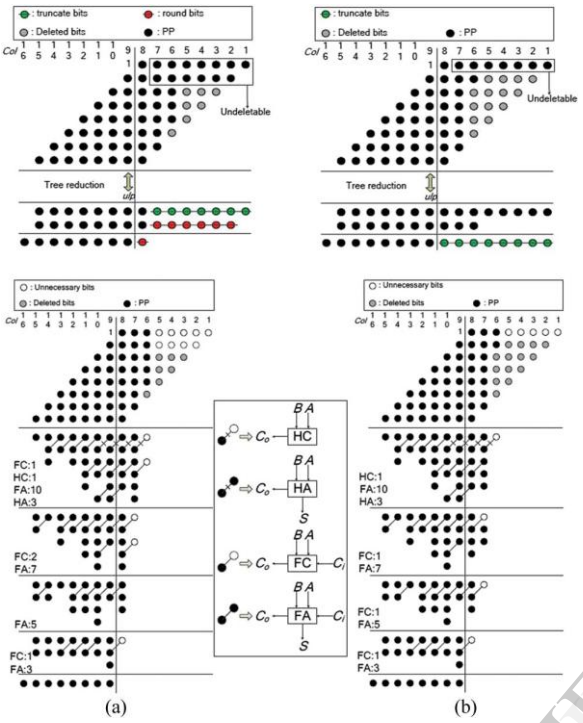
The FIR filter design in this brief adopts the direct formin Fig. 1(a) where the MCMA module sums up all the products $\hat{a}_i \times x[n-i]$. Instead of accumulating individual multiplication for each product, it is more efficient to collect all the PPs into a single PPB matrix with carry-save addition to reduce the height of the matrix to two, followed by a final carry propagation adder. Fig. 4 illustrates the difference of individual multiplications and combined multiplication for $A \times B + C \times D$.

In order to avoid the sign extension bits, we complement the sign bit of each PP row and add some bias constant using the property $\bar{s} = 1 - s$, where s is the sign bit of a PP row, All the bias constants are collected into the last row in the PPB matrix.

The complements of PPBs are denoted by white circles with overbars. In the faithfully rounded FIR filter implementation, it is required that the total error introduced during the arithmetic operations is no larger than one ulp. We modify a recent truncated multiplier design in [14] so that more PPBs can be deleted, leading to smaller area cost. Fig. 6 compares the two approaches. In [14], the removal of unnecessary PPBs is composed of three processes: deletion, truncation, and rounding.

Two rows of PPBs are set undeletable because they will be removed at the subsequent truncation and rounding.

TRUNCATION & ROUNDING:



In this brief, we propose an improved version of the faithfullyrounded truncated multiplier design as shown in Fig. (b).

TABLE I
SPECIFICATIONS OF THE THREE FIR FILTERS UNDER CONSIDERATION

Filter		M	M_{opt}	B	EWL	f_{pass}	f_{stop}	A_{pass} (dB)	A_{stop} (dB)
A	LP	25	28	11	10	0.15	0.25	0.09	46
B	LP	59	64	15	12	0.02	0.07	0.20	60
C	HP	121	121	19	17	0.40	0.37	0.10	80

a single row of PPBs is made undeletable (for the subsequent rounding), and the PPB elimination consists of only deletionand rounding. The error ranges of deletion and rounding in the
improved version are as follows:

$-\text{ulp} \le E\_D \le 0 < 1/2$
$\text{ulp} \le E_D = E\_D + 1/2$
$\text{ulp} \le 1/2$
$-\text{ulp} < E\_R \le 0 - 1/2$
$\text{ulp} < E_R = E\_R + 1/2$
$\text{ulp} \le 1/2 < \text{ulp}$
$-\text{ulp} < E = (E_D + E_R) \le \text{ulp}.$

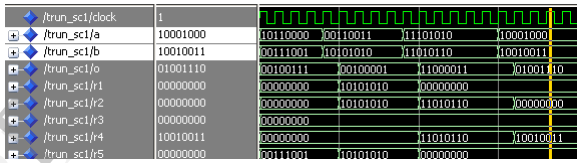
Since the range of the deletion error in the improved version is twice larger than that in [14], more PPBs can be deleted, leading to smaller area in the subsequent PPB compression

IV. EXPERIMENTAL RESULTS AND COMPARISONS

We implemented three FIR filters with the specificationsgiven in Table I [15], [16]. M is the original filter order while M_{opt} is the filter order with optimized total area using the method in Fig. 2. B denotes the number of fractional bits for uniformly quantized coefficients with filter order M_{opt},
EWL is the effective word length without counting the leading sign bits, f_{pass} and f_{stop} are the passband and stopband edge frequencies normalized to one, and A_{pass} and A_{stop} denote the corresponding peak-to-peak ripples.

SIMULATION RESULTS:

TRUNCATION SCHEME-1:



CONVENTIONAL MULTIPLIER:

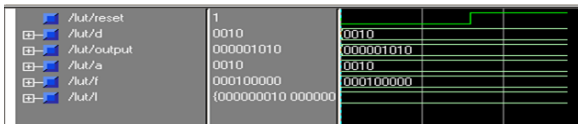
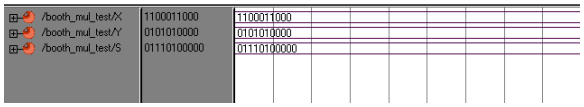


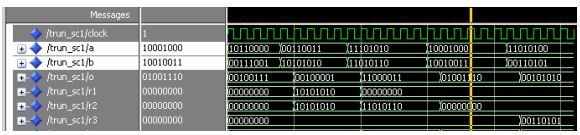
TABLE:

MULTIPLIERS	AREA~	POWER~
Scheme-1	1678	55pw
Scheme-2	1306	42pw

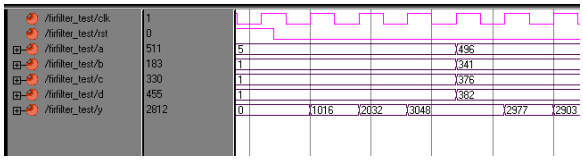
TRUNCATION SIGNED MULTIPLIER:



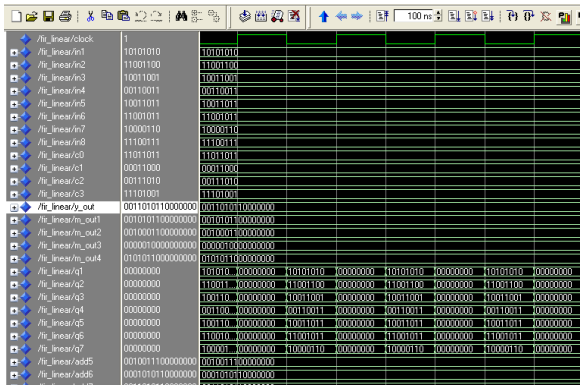
TRUNCATION-2 UNSIGNED MULTIPLIER:



3 TAP FIR FILTER:



LINEAR PHASE FIR:



SYNTHESIS RESULTS OF FILTER A WITH 12-BIT INPUT AND OUTPUT SIGNALS

Filter A (28-tap LP)	MCM (um^2)	SAs (um^2)	DFFs (um^2)	Area (um^2)	Delay (ns)	TP (M data/s)	Power (mw)	AP/TP	Struct.
CSD/Booth	4527	9737	9655	23919	4.78	209	1.15	2.38	Trans.
	13460		5030	18490	6.75	148.	0.91	2.05	Direct
NRSCSE[1]	3151	9737	9655	22543	4.78	209	1.10	2.14	Trans.
MBPG[2]	3141	9737	9655	22533	4.78	209	1.14	2.22	Trans.
LUT [11]	16155	9737	9655	35547	4.77	210	1.66	5.08	Trans.
1D DA [12]	3146		5526	8672	4.83	17	0.36	3.32	Direct
MCMA	10161		5030	15191	6.58	152	0.90	1.63	Direct
MCMA_opt	9277		5030	14307	6.31	159	0.81	1.32	Direct
MCMAT_I	7716		5030	12746	6.35	158	0.71	1.04	Direct
MCMAT_II	7460		5030	12490	6.35	158	0.70	1	Direct

Proposed work:

Although most prior designs are based on the transposed form, we observe that the direct FIR structure with faithfully rounded MCMAT leads to the smallest area cost and power consumption.

Using vedic multiplication , also reduce the hard ware computation.

CONCLUSION:

This brief has presented low-cost FIR filter designs by jointly considering the optimization of coefficient bit width and hardware resources in implementations. Multiplier is the key component of many high performance systems, using this truncation parallel multiplier area ,complexity, power are reduced.

REFERENCES:

1.H.-J. Ko and S.-F. Hsiao, "Design and application of faithfully rounded sand truncated multipliers with combined deletion, reduction, truncation and rounding," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 58, no. 5, pp. 304–308, May 2011.

2.F. Xu, C. H. Chang, and C. C. Jong, "Contention resolution algorithms forcommon subexpression elimination in digital filter design," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 52, no. 10, pp. 695–700, Oct. 2005.

3. I.-C. Park and H.-J. Kang, "Digital filter synthesis based on an algorithm to generate all minimal signed digit representations," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 21, no. 12, pp. 1525–1529, Dec. 2002.

4 C.-Y. Yao, H.-H. Chen, T.-F. Lin, C.-J. J. Chien, and X.-T. Hsu, "A novel common-subexpression-elimination method for synthesizing fixed-point FIR filters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 11, pp. 2215–2221, Sep. 2004.

5. O. Gustafsson, "Lower bounds for constant multiplication problems," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 54, no. 11, pp. 974–978, Nov. 2007.

6. A. Blad and O. Gustafsson, "Integer linear programming-based bit-leveloptimization for high-speed FIR filter architecture," *Circuits Syst. SignalProcess.*, vol. 29, no. 1, pp. 81–101, Feb. 2010.

7. F. Xu, C. H. Chang, and C. C. Jong, "Design of low-complexity FIRfilters based on signed-powers-of-two coefficients with reusable commonsubexpressions," *IEEE Trans. Comput.-Aided Design Integr. CircuitsSyst.*, vol. 26, no. 10, pp. 1898–1907, Oct. 2007.

8. Y. J. Yu and Y. C. Lim, "Design of linear phase FIR filters in subexpressionspace using mixed integer linear programming," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 10, pp. 2330–2338, Oct. 2007.

9. K. C. Bickerstaff, M. Schulte, and E. E. Swartzlander, Jr., "Reduced areamultipliers," in *Proc. Int. Conf. Appl.-Specific Array Processors*, 1993,pp. 478–489.

10. P. K. Meher, "New approach to look-up-table design and memory-based realization of FIR digital filter," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 57, no. 3, pp. 592–603, Mar. 2010.