ABSTRACT:- The high speed FIR Filter designs are presented using the concept of faithfully rounded truncated multipliers. Truncated multipliers offer significant improvements in area, delay, and power. We jointly consider the optimization of bit width and hardware resources without sacrificing the frequency response and output signal precision. Non uniform 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. Comparisons with previous FIR design approaches show that the proposed designs achieve the best area and power results. This design is simulated using Modelsim6.4b and synthesized using XILINX ISE10.1 software’s.

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 as

\[ y[n] = \sum_{i=0}^{M-1} a_i x[n-i]. \]

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, 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 coefficients \( a_i \).

Fig: Structures of linear-phase even-order FIR filter direct form

In the transposed form in Fig, the operands of the multipliers in the MCM module are the current input signal \( x[n] \) and coefficients. The results of individual constant multiplications go through structure adders (SAs) and delay elements. In the past decades, there are many papers on the designs and implementations of low-cost or high-speed FIR filters. In order to avoid costly multipliers, most prior hardware implementations of digital FIR filters can be divided into two categories: multiplier less based and memory based.
Multiplier less-based designs realize MCM with shift-and add operations and share the common sub operations using canonical signed digit (CSD) recoding and common sub expression elimination (CSE) to minimize the added cost of MCM. Mostly area savings are achieved by jointly considering the optimization of coefficient quantization and CSE. Most multiplier less 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 is large. 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 the SAs. In 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).

II. FIR FILTER DESIGN AND OPTIMIZATION

A digital FIR filter design consists of three steps and they are finding the filter order and coefficient, coefficient quantization, and hardware optimization. In the first step, the filter order and its corresponding coefficients are determined in order to satisfy the frequency response specifications. Then in the second step, coefficient quantization is performed to quantize the coefficients to finite bit width accuracy. Finally, hardware optimization is done to reduce the area and power consumption of hardware structure.

Hardware optimization is done by optimizing the design of multipliers used in FIR filter structure which occupies more area and consumes more power. Generally the design of parallel multipliers consists of following steps, i.e., partial product generation, partial product reduction, final carry propagate addition. In the first step, the partial products are generated from multiplicand and multiplier. In the partial product reduction step, the generated partial products are compressed and reduced to two partial products and then they are finally added using final carry propagate adder.

Already two famous reduction techniques available and they are Dadda tree and Wallace tree reduction. In Dadda reduction, whenever compression operation is required the reduction is performed whereas in Wallace tree reduction, the partial product bits are always compressed.

There are two types of reduction schemes adopted in this paper called scheme-1 and scheme-2 reductions in order to reduce the number of half adders used in each column and also to perform flexible column-by-column reduction. By combining scheme-1 and scheme-2 reduction in the design of truncated multiplication, the area cost can be minimized.

Fig: Three stages in digital FIR filter design and implementation

Fig: Structures of linear-phase even-order FIR filters transposed form
III. TRUNCATED MULTIPLIER DESIGN

Truncated multiplication reduces power consumption by computing only the most significant bits of product result and most common approach used for truncation is physical reduction of partial product bit matrix. The main objective of improved faithfully rounded truncated multiplier design is to compute $P$ most significant bits of the product with a total truncation error of not more than $1ulp$, where $1ulp = 2^{-P}$. In the Existing design of truncated multiplication, there are three processes to remove the unnecessary Partial Product Bits (PPBs) and they are: deletion, truncation, and rounding. Two rows of partial product bits are set undeletable because they will be removed at the subsequent process of truncation and rounding. The two white dots at level 1 of reduction and the two white dots at level 2 of reduction are not generated during the process of compression with FAs or HAs.

Thus, two simplified versions of the full adder (FA) and half adder (HA) cells are introduced, i.e., full adder and half adders without the sum output bits. For column $T$, the carry bit only is needed to be generated (to column $T + 1$) for the last full adder compression because the sum output bit will be discarded during the rounding process. For example, the FA white dot (the sum output bit) at level 4 of reduction.

For column $T + 1$ to $M + N$, although Scheme 1 is adopted to determine whether an HA is needed or not, we actually do not compress the column height to one because this compression will cause ripple carry. Indeed, at the last level of the reduction process, some column, for example column $i$, has a height of three, and the remaining columns beyond this specific column, i.e., columns $i + 1$, $i + 2$, . . . etc have a column height of two, as shown in level 4 of Fig. Finally compressed partial product bits are added using final carry propagate adder.

The error ranges of deletion, truncation, and rounding before and after adding the offset constants are given as follows:
Fig: Overall FIR filter architecture using multiple constant multipliers/accumulators with faithfully rounded truncation (MCMAT)

The Architecture of MCMA with truncation (MCMAT) that removes unnecessary PPBs. The white circles in the L-shape block represent the undeletable PPBs. The deletion of the PPBs is represented by gray circles. After PP compression, the rounding of the resultant bits is denoted by crossed circles. The last row of the PPB matrix represents all the offset and bias constants required including the sign bit modifications.

<table>
<thead>
<tr>
<th>Filter</th>
<th>M</th>
<th>Mlep</th>
<th>B</th>
<th>EWL</th>
<th>fstop</th>
<th>fnypf</th>
<th>Astop(dB)</th>
<th>Anypf(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>LP</td>
<td>25</td>
<td>28</td>
<td>11</td>
<td>10</td>
<td>0.15</td>
<td>0.25</td>
<td>0.09</td>
</tr>
<tr>
<td>B</td>
<td>LP</td>
<td>59</td>
<td>64</td>
<td>15</td>
<td>12</td>
<td>0.03</td>
<td>0.07</td>
<td>0.20</td>
</tr>
<tr>
<td>C</td>
<td>HP</td>
<td>131</td>
<td>131</td>
<td>19</td>
<td>17</td>
<td>0.40</td>
<td>0.57</td>
<td>0.10</td>
</tr>
</tbody>
</table>

V. SIMULATION AND SYNTHESIS RESULT

VI. CONCLUSION

This brief has presented low-cost FIR filter designs by jointly considering the optimization of coefficient bit width and hardware resources in implementations. 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.

REFERENCES


