

The Energy Efficient Functional Unit for Fully Optimized DSP Accelerator Architecture Manipulating Carry-Save Arithmetic



Ch.M.M Komali

Dept. of Electronics and Communication Engineering,
Baba Institute of Technology and Sciences,
Visakhapatnam, AP, India.



B.V.R Gowri

Dept. of Electronics and Communication Engineering,
Baba Institute of Technology and Sciences,
Visakhapatnam, AP, India.

Abstract

DSP accelerators are hardware modules appended to a processor core remotely to improve the accomplishment and usefulness of computationally concentrated DSP functions. Area particular hardware designs shapes perfect speeding up as far as execution and force, yet their unbendable data ways lead to expanded silicon many-sided quality. In adaptable DSP accelerators agent functional computational unit (FCU) are joined to enhance execution, decrease energy consumption and to give adaptable data ways. The design misuses carry save (CS) number juggling to empower quick affixing of added substance and multiplicative operations. Be that as it may, the carry save enhancement approaches have constrained effect on data flow graph (DFG) overwhelmed by duplications. In any case, research exercises have demonstrated the math improvements at more noteworthy reflection levels contrasted with auxiliary circuit one significantly impact on the data way execution. CS representation keeps on being extensively usual to plan quick number juggling circuits as a result of its characteristic advantage of disposing of the huge convey proliferation chains. Hardware increasing speed keeps on being shown an extremely encouraging usage system for digital signal processing (DSP) space. Rather than receiving a solid application-particular incorporated circuit outline approach, inside this brief, we show a

composition accelerators agent design made out of adaptable computational models that offer the execution of a major gathering of operation layouts present in DSP popcorn portions. It can likewise be utilized as a part of different picture processing applications.

Keywords: Carry-Save (CS), Datapath Synthesis, Flexible Accelerator, Digital Signal Processing (DSP), Hardware Acceleration.

I. Introduction

The consolidation of heterogeneity through specific hardware accelerators enhances execution and diminishes energy consumption. Advanced inserted frameworks target high-complete application space names requiring productive usage of computationally serious digital signal processing (DSP) functions. Numerous researchers have recommended utilizing area particular coarse-grained reconfigurable accelerators agents, to have the capacity to build ASICs' adaptability without significantly bargaining their execution [1]. Despite the fact that application-specific integrated circuits (ASICs) make up the perfect increasing speed arrangement with regards to execution and force, their firmness brings about lifted plastic many-sided quality, as different instantiated ASICs are important to quicken different popcorn pieces. Elite adaptable datapaths happen to be proposed to effectively outline or fastened methods

situated in the underlying data-flow graph (DFG) of the portion. The formats of complex blinded systems are by and large expelled from the bit's DFG or determined by a predefined conduct layout library. Outline decisions around the accelerators agent's datapath very effect its productivity. Existing makes coarse-grained recon-figural datapaths predominantly misuse design level advancements, e.g., lifted direction level parallelism. The area particular engineering era and change the kind and amount of calculation models accomplishing a customized outline structure. Adaptable designs were proposed abusing ILP and operation affixing [2]. Of late, Ansaloni et al. received forceful operation anchoring to permit the calculation of whole sub expressions utilizing numerous ALUs with heterogeneous number-crunching highlights. These reconfigurable designs bar math improvements all through the building combination and consider them restricted to the inside circuit structure of primitive parts, e.g., adders, all through the rationale blend. Be that as it may, research exercises have demonstrated the math advancements at more noteworthy deliberation levels contrasted with basic circuit one impressively impact on the datapath execution. Timing-driven advancements as indicated by carriesave (CS) number juggling were completed in the distribute Register Transfer Level (RTL) plan stage. Normal sub expression disposal in CS calculations can be utilized to advance straight line DSP circuits [3]. Verma et al. created change strategies around the application's DFG to make best utilization of CS number juggling earlier the specific datapath combination. These CS improvement approaches target unyielding datapath, i.e., ASIC, executions. Of late, Xydis et al. proposed a versatile engineering blending the ILP and pipelining methods utilizing the CS-mindful operation affixing. Be that as it may, the greater part of the previously mentioned arrangements highlight a characteristic impediment, i.e., CS streamlining is limited to combining just increases/subtractions. A CS to parallel transformation is set before every operation cap is not quite the same as expansion/subtraction, e.g., augmentation, along these lines, apportioning various CS to twofold

changes that intensely debases execution in light of tedious convey propagations[3]. Inside this brief, we exhort a higher-execution compositional arrangement for that union of adaptable hardware DSP accelerators agents by blending enhancement systems from both engineering and number-crunching measures of deliberation. We present a versatile datapath engineering that adventures CS upgraded formats of tied methodology. The recommended design contains adaptable computational models (FCUs), which let the execution of a major gathering of operation layouts present in DSP popcorn parts. The recommended accelerators agent engineering gives normal increases as high as 61.91% in territory delay item and 54.43% in energy consumption in contrast with condition of-workmanship adaptable datapaths, maintaining proficiency toward layered advancements.

II. Related Work

The significant part of the task advancement division considers and completely overview all the required requirements for building up the undertaking. Before building up the instruments and the related outlining it is important to decide and review the time element, asset necessity, labor, economy, and organization quality. Once these things are fulfilled and completely reviewed, then the following stride is to decide about the product particulars in the individual framework, for example, what kind of working framework the venture would require, and what are all the vital programming are expected to continue with the following stride, for example, building up the apparatuses, and the related operations. Numerous scientists have proposed in the past frameworks, the utilization of space particular coarse-grained reconfigurable accelerators agents so as to expand ASICs' adaptability without essentially trading off their execution. Superior adaptable data ways have been proposed to effectively outline or tied operations found in the underlying data-flow graph (DFG) of a portion. The formats of complex affixed operations are either extricated specifically from the piece's DFG or indicated in a predefined behavioral layout library. Outline choices on the accelerators agent's data way very effect its effectiveness. Existing

takes a shot at coarse-grained reconfigurable data ways chiefly abuse engineering level advancements, e.g., expanded direction level parallelism (ILP).

A. An adaptable and energy-effective coarse-grained reconfigurable engineering for portable frameworks

Two of the most vital configuration issues for cutting edge handheld gadgets are remote systems administration and the processing of interactive media. Both applications depend intensely on computationally escalated digital signal processing calculations. Programmable models that keep pace with the expanding execution necessities turn out to be increasingly control hungry. This is risky for a battery fueled cell phone, since it has just a restricted measure of energy accessible. On the other hand, devoted models are excessively resolute, making it impossible to keep pace with changing gauges and capabilities. A cell phone requires superior, adaptability and (energy-) proficiency. These negating prerequisites should be adjusted in the framework engineering of a cell phone. In this paper a heterogeneous design of space particular processing tiles is proposed. The point of convergence is the coarse-grained reconfigurable engineering of the Montium processing tile, which is intended to execute digital signal processing calculations energy proficiently.

B. ADRES: An engineering with firmly coupled VLIW processor and coarse-grained reconfigurable grid

The coarse-grained reconfigurable designs have points of interest over the conventional FPGAs as far as postponement, territory and arrangement time. To execute whole applications, a large portion of them join a guideline set processor (ISP) and a reconfigurable network. In any case, very little consideration is paid to the mix of these two sections, which brings about high correspondence overhead and programming trouble. To address this issue, we propose a novel design with firmly coupled long direction word (VLIW) processor and coarse-grained reconfigurable framework. The points of interest

incorporate rearranged programming model, shared asset costs, and diminished correspondence overhead. To endeavor this design, our already created compiler system is adjusted to the new engineering. The outcomes demonstrate that the new engineering has great execution and is exceptionally compiler-accommodating

C. An elite data way to synthesize DSP parts

An elite data way to actualize digital signal processing (DSP) parts is presented in this paper. The data way is acknowledged by an adaptable computational part (FCC), which is an unadulterated combinational circuit and it can execute any 2 times 2 layout (group) of primitive assets. Along these lines, the data way's execution profits by the intra segment binding of operations. Because of the adaptable structure of the FCC, the data way is executed by a little number of such parts. This takes into account direct associations among FCCs and for abusing entomb segment fastening, which further enhances execution. Because of the comprehensiveness and adaptability of the FCC, basic and productive calculations perform booking and official of the data flow graph (DFG).

III. Digital Signal Processor

A. Architecture Overview

Our FU has been designed and implemented for a parallel asynchronous DSP named CADRE [5]. CADRE was proposed to be a minimum power consumption DSP whilst meeting the performance requirements of next generation cellular phones. However, the simulation results show that the power dissipation is still on the high side. In [8], the power dissipation of CADRE was analyzed and approximately 50% of the overall power consumption was found to be dissipated in the FU. Therefore, the new FU in this research will replace on the original design. The original philosophy of the architecture is described to help the reader understand our FU easily. CADRE was implemented by exploiting four-way parallelism as this appears to be optimal for power reduction, Chandrakasan and Brodersen [6]. This is based on the premise that area can be traded for

increased speed because silicon area is rapidly becoming less expensive. Most of the DSP activity can be characterized by frequent repetition of fixed instruction sequences. So, the instruction encoding which determines the selection and passage of data for each operation can be predetermined and stored in advance in a configurable memory which is located locally to each FU. These encodings can then be recalled with a compressed instruction. Because the configuration memories are RAMs, this allows reconfiguration at any point in execution. In addition, CADRE the encodings could be expanded within the FUs.

This top-level architecture has been adapted for the current research work because time is too limited to fully implement the CADRE design. However, we still keep the major advanced feature such as four-way parallelism to give high throughput. Meanwhile, a new FU has been designed with its configuration memory. The new system consists of four FUs connected together with a global bus and a pair of FUs are connected locally. The input data of each FU will be directly provided from on-chip RAM blocks with the output data being kept in another RAM block. The encoded top-level instruction is stored in a program memory. It contains controllable to enable the FU and accumulator writeback plus a 5-bit address which accesses the configuration memory associated with each FU; the functional unit instructions are stored in advance into each configuration memory of the FUs.

B. Instruction Set

The instruction set of the proposed FUs has two sets of instructions: computation and data movement and these can occur concurrently.

Computational instructions:

These instructions contain arithmetic and logical operations; the arithmetic instructions include distance, normalization, shift, ADD, SUB, MPY and MAC (multiply and accumulate). The output destination of the operation can be 1 of 4 accumulator registers (AccA to AccD) inside the FU.

Data Movement Instructions:

These instructions process the data movement in or out of the FU. These instructions include an accumulator register to accumulator register transfer within the FU itself, an accumulator register to an accumulator register in another FU, or a movement of data to the output RAM.

IV. Flexible Accelerator Architecture

An optimized accelerator architecture for DSP using modified booth is shown in figure1. The architecture mainly consists of flexible computational units (FCUs). Each FCU operates directly on CS operands and produces data in the same form for the direct use of intermediate result. The number of computational unit is determined at the design time based on instruction level parallelism and area constraints imposed by the designer. Each FCU can be configured to operate based on a set of operation templates. The most suitable FCU is selected with the help of multiplexer. Register bank is mainly used to store the intermediate results and sharing values of operands among the FCUs[6]. The different types of templates are shown in figure1. Each template may consist of an adder/subtractor module and a multiplier section. The multiplier here used is a modified booth multiplier. Booth algorithm for multiplication is a simple method in which multiplication is carried out with repeated addition operation. To overcome the main limitations of booth algorithm modified booth algorithm is used. Modified booth multiplier[7] is used to perform high speed multiplication using modified booth algorithm. The main advantage is that we can reduce the number of partial products to half. The detailed diagram of multiplier is shown in figure 1.

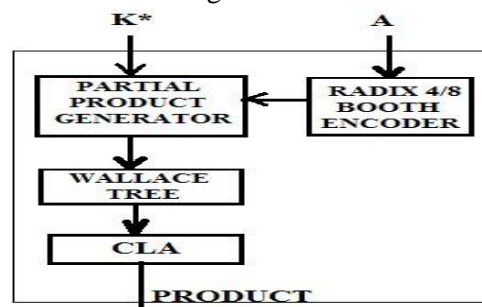


Figure 1: Modified Booth Encoded Multiplier

The Radix8 booth encoder within the multiplier performs the process of encoding the multiplicand based on the multiplier bits. It will compare three bits at a time based on the overlapping technique. For the purpose of generation of the partial products partial product generator is used. For large multipliers the performance of the modified booth algorithm is limited. For that purpose booth encoding together with the Wallace tree structure can be used. Wallace tree adders are used in high speed designs to produce two rows of partial products that can be added in the final stage. Critical path and the number of adders are reduced as compared to the parallel adders. The speed, area and the power consumption of the multipliers are directly proportional to the efficiency of the compressors.

A carry save adder is a type of digital adder used in the computer architecture to compute the sum of three or more n bit numbers in binary. It differs from other digital adders in the sense that it outputs two numbers of the same dimension as that of the inputs, one which is a sequence of the sum bit and the other which is a sequence of carry bits. This kind of carry save adder trees are used for high speed implementation of multiple operand addition. The adder/ subtractor module is mainly designed with the help of carry select adders instead of using the ripple carry adders. Thus the carry propagation in the ripple carry adder can be completely overcome by using the carry select adder. This will significantly reduce the delay and enhances the performance.

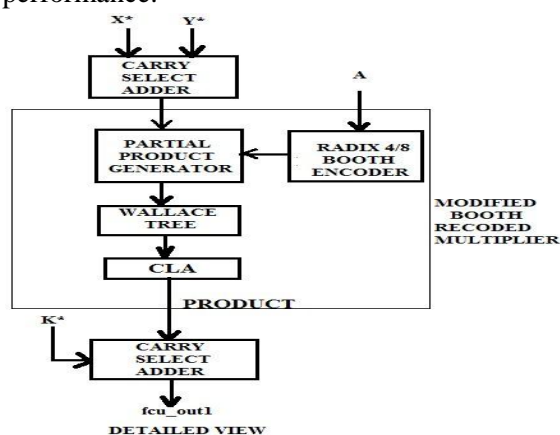


Figure 2: Detailed view of computational unit 1

V. PROPOSED SYSTEM

DSP accelerators performance can be increased by the optimization of both in architectural and arithmetic level. Flexible data path architecture that exploits carry-save optimized templates which comprises flexible computational unit. In arithmetic level the carry- save arithmetic eliminates the intermediate carry propagate adder which is used to convert carry save to binary format. The Fig.1. shows the block diagram of flexible computational unit used in DSP Accelerator. The flexible computational unit mainly consists of a compressor, mux's, a multiplier unit and finally an Adder. The detailed view of the basic building block is shown above: The two's complement 4:2 CS adder produces two outputs based on the input carry signal value. if the value of carry is 0 then the $N^* = X^* + Y^*$ otherwise $N^* = X^* - Y^*$. The MUX 1 determines whether A is multiplied by N^* (1) or K^* (2). The MUX 2 specifies among K^* (1) and N^* (2) which is to be added with the multiplication product. The multiplexer MUX 3 accepts the output of MUX 2 and its ones complement and makes an outputs based on the equation 4.1 if the signal CL3 is 0 or it produces an output based on the equation 4.2 if the CL3 is 1. The multiplier comprises a CS-to-MB module, which exploits a recent recoding technique, the 17-bit P^* in its respective Modified Booth digits with minimal carry propagation. The multipliers product consists of 17 bits.

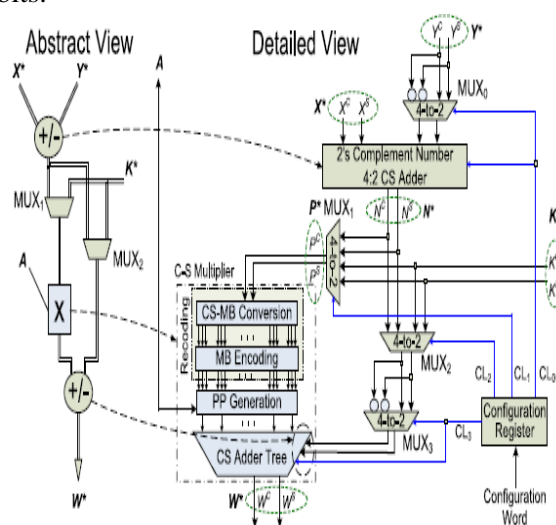


Fig.3. Proposed FCU

The compensating method truncation technique is used for reducing the error imposed at the products accuracy of the multiplier. Since all the FCU inputs consist of 16 bits and it does not produce overflows, the 16 most significant bits of the 17-bit W^* are inserted in the appropriate FCU when requested.

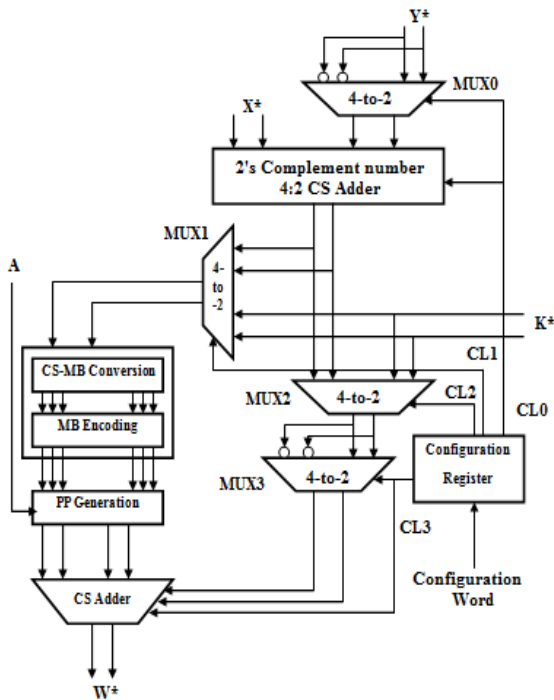


Fig.4. Detailed view of FCU

The FCU enables template operation chaining by fused add-multiply operator;

$$W^* = A(X^* + Y^*) + K^* \quad (1)$$

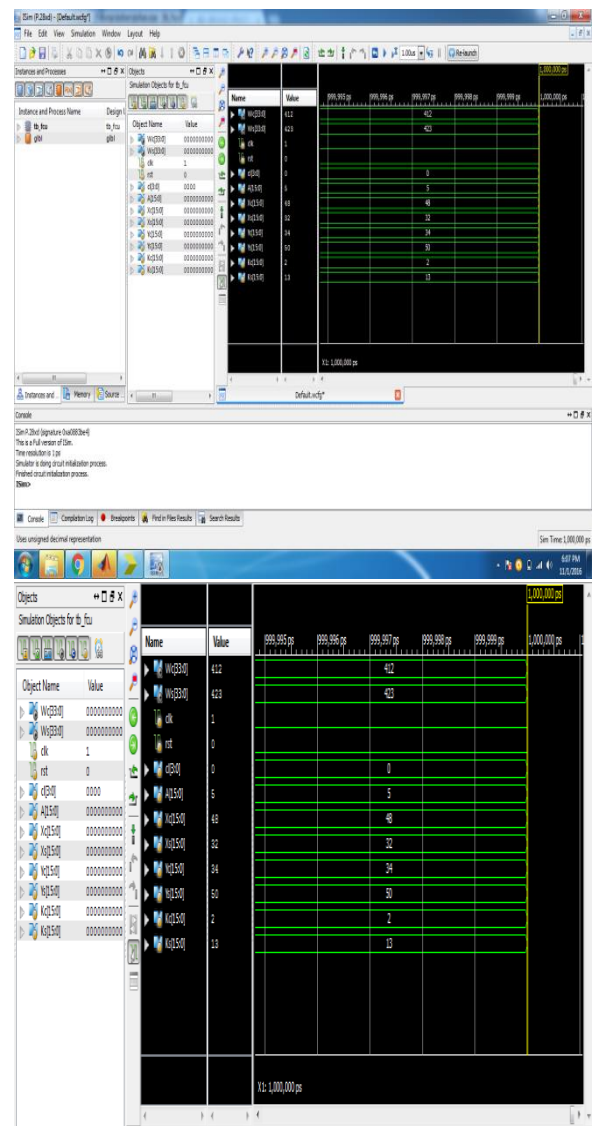
$$W^* = A + K^* + (X^* + Y^*) \quad (2)$$

For increasing the performance of the system we can optimize the add-multiply operator. The existing recoding schemes are on the basis of complex manipulations in bit level. These recoding schemes are implemented by dedicated circuits in gate-level in this case. In order to get more efficient implementation of the fused Add-Multiply (FAM) unit, direct recoding of the sum of two numbers in its modified booth (MB) form leads to a more efficient. The efficient design of FAM operators can be done by the direct shaping of modified booth form of sum of the two numbers to

sum to modified form(S-MB). The S-MB algorithm is simple in structure and can be modified in order to apply either in signed or unsigned numbers, which can comprise odd or even number of bits. Three alternative schemes of the proposed S-MB approach. That is by using various building blocks such as conventional, signed-bit Full Adders (FAs) and Half Adders (HAs).

VI. Results and Discussions

The results of the carry save adder is implemented using the VLSI tool. The advantages such power leakage and complexity of the system is reduced by using the carry save adder in FCU. The simulation results are shown below.



VII. Conclusion

In this brief, we introduced a flexible accelerator architecture that exploits the incorporation of CS arithmetic optimizations to enable fast chaining of additive and multiplicative operations. The proposed flexible accelerator architecture is able to operate on both conventional two's complement and CS-formatted data operands, thus enabling high degrees of computational density to be achieved. Theoretical and experimental analyses have shown that the proposed solution forms an efficient design trade off point delivering optimized latency/area and energy implementations.

References

- [1] P. Ienne and R. Leupers, Customizable Embedded Processors: Design Technologies and Applications. San Francisco, CA, USA: Morgan Kaufmann, 2007.
- [2] P. M. Heysters, G. J. M. Smit, and E. Molenkamp, "A flexible and energy-efficient coarse-grained reconfigurable architecture for mobile systems," *J. Supercomput.*, vol. 26, no. 3, pp. 283–308, 2003.
- [3] B. Mei, S. Vernalde, D. Verkest, H. D. Man, and R. Lauwereins, "ADRES: An architecture with tightly coupled VLIW processor and coarsegrained reconfigurable matrix," in *Proc. 13th Int. Conf. Field Program. Logic Appl.*, vol. 2778. 2003, pp. 61–70.
- [4] M. D. Galanis, G. Theodoridis, S. Tragoudas, and C. E. Goutis, "A high-performance data path for synthesizing DSP kernels," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 25, no. 6, pp. 1154–1162, Jun. 2006.
- [5] K. Compton and S. Hauck, "Automatic design of reconfigurable domainspecific flexible cores," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 16, no. 5, pp. 493–503, May 2008.
- [6] S. Xydis, G. Economakos, and K. Pekmestzi, "Designing coarse-grain reconfigurable architectures by inlining flexibility into custom arithmetic data-

paths," *Integr., VLSI J.*, vol. 42, no. 4, pp. 486–503, Sep. 2009.

[7] S. Xydis, G. Economakos, D. Soudris, and K. Pekmestzi, "High performance and area efficient flexible DSP datapath synthesis," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 3, pp. 429–442, Mar. 2011.

[8] G. Ansaloni, P. Bonzini, and L. Pozzi, "EGRA: A coarse grained reconfigurable architectural template," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 19, no. 6, pp. 1062–1074, Jun. 2011.

Author Details

Ch.M.M Komali pursuing her M.Tech in the department of Electronics and Communication Engineering, Baba Institute of technology and Sciences, Visakhapatnam, A.P., India. Affiliated to Jawaharlal Nehru Technological University, Kakinada, Approved by AICTE, NEW DELHI. She obtained her B.Tech(ECE) from JNTUK college of Engineering.

B.V.R Gowri, M.Tech, working as Assistant Professor in the department of Electronic and Communication Engineering, Baba Institute of technology and Sciences, Visakhapatnam, A.P., India. Affiliated to Jawaharlal Nehru Technological University, Kakinada. Approved by AICTE, NEW DELHI. Her area of interest in Digital Signal Processing, Pulse & Digital Circuits