International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099

VOL 2 ISSUE 1 (2017) PAGES 77 - 82

Received: 1/09/2017. Published: 26/09/2017

INEXACT FLOATING POINT ADDER FOR IMAGE PROCESSING APPLICATION

R.Keerthiga, A.Syed Mazhar

Department of Electronics and Communication Engineering

Dhanalakshmi Srinivasan Institute of Technology, Samayapuram, Tamil Nadu, India

Abstract

Power has become a key constraint in current nanoscale integrated circuit design due to the increasing demands for mobile computing and a low carbon economy. As an emerging technology, an inexact circuit design offers a promising approach to significantly reduce both dynamic and static power dissipation for error-tolerant applications. Although fixed-point arithmetic circuits have been studied in terms of inexact computing, floating-point arithmetic circuits have not been fully considered although they require more power. In this paper, the first inexact floating-point adder is designed and applied to high dynamic range (HDR) image processing. Inexact floating-point adders are proposed by approximately designing an exponent subtractor and mantissa adder. Related logic operations including normalization and rounding modules are also considered in terms of inexact computing. Two HDR images are processed using the proposed inexact floating-point adders to show the validity of the inexact design. HDR-VDP is used as a metric to measure the subjective results of the image addition. Significant improvements have been achieved in terms of area, delay and power consumption. Comparison results show that the proposed inexact floating-point adders can improve power consumption and the power-delay product by 29.98% and 39.60%, respectively.

Keywords- Inexact computing, Floating-point adders, Low power, High-dynamic-range image

I. Introduction

With advancements in the development of digital integrated circuits, power consumption has increased dramatically. Power has become a key design constraint due to the high demand for mobile computing and a low carbon economy. Traditional designs apply fully accurate computing to all types of applications. However, error-tolerant applications (such as those perceived by humans) do not require full accuracy, so it is possible to perform useful computation with inexact circuits. In these cases, inexact computing is a very attractive approach to save power, area and achieve higher performance when compared to conventional accurate designs.

The arithmetic unit is the core of a processor, and its power largely determines the power of the whole processor. Recent research on inexact fixed-point adders has shown that inexact processing hardware with a relative error of 7.58% can be 15 times more efficient in terms of speed, area, and energy product than an accurate chip [1]. Inexact chips are smaller, faster and consume less energy. Although fixed-point arithmetic circuits have been studied in terms of inexact computing, floating-point arithmetic circuits that are much more power-hungry have not been fully considered. The floating-point format offers a high dynamic range for computationally

VOL 2 ISSUE 1 (2017) PAGES 77 - 82
Received: 1/09/2017. Published: 26/09/2017

intensive applications; floating-point adders and multipliers are commonly used in DSP systems. However, application to embedded DSP systems is limited due to the high power consumption.

In this paper, adder designs are studied as a starting point for inexact floating-point arithmetic. Several inexact designs are proposed and verified with application to high dynamic range images. A subjective visual difference predictor metric is used to measure the image addition results. In addition, a methodology is proposed for designing inexact floating-point arithmetic circuits.

II. Preliminaries

A. Review

Although the inexact design of fixed-point adders has been extensively studied [1-4], little research has been conducted on inexact floating-point arithmetic design. A low power design of a floating-point multiplier was investigated by Tong et al. which involves truncating hardware [5]; the rounding unit was found to require almost half of the hardware of an exact floating-point multiplier. Therefore, the rounding unit is a candidate for removal to save power, similar to an inexact design. A probabilistic floating-point multiplier was proposed by Gupta et al. [6] as an energy-efficient design. However, to the best of the authors' knowledge, there has been no research to date on an inexact floating-point adder design, which has a more complex structure than a floating-point multiplier. Therefore, in the next sections design techniques to achieve an inexact floating-point adder design are discussed.

B. Floating-Point Adder and Architecture

A generic floating-point adder architecture includes exponent comparison, mantissa swap, and alignment, mantissa addition, normalization and rounding of the mantissa, as shown below. Two operands are first unpacked, in which a 23-bit mantissa is added to the hidden bit. The addition of floating-point numbers involves comparing two 8-bit exponents and adding two 24-bit mantissa; the exponents are first evaluated to find the larger number. The mantissa is then swapped according to the exponent comparison and aligned to have an equal exponent before they are added in the mantissa adder.

Following the addition, the normalization is completed by left shifting with the number of leading zeros. Therefore, leading zero detection is a key step of normalization. Rounding the normalized result is the last step before storing it; special cases (such as overflow, underflow, zero) are also detected and represented by flags. The accurate single precision floating-point adder architecture is shown in Fig.1.

III. Design Of Inexact Floating-Point Adders

A simple design of an inexact floating-point adder involves replacing the mantissa adder and exponent adder with approximate adders. However, as the floating-point adder architecture is substantially different from a fixed-point adder, related logic.

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 2 ISSUE 1 (2017) PAGES 77 - 82
Received: 1/09/2017. Published: 26/09/2017

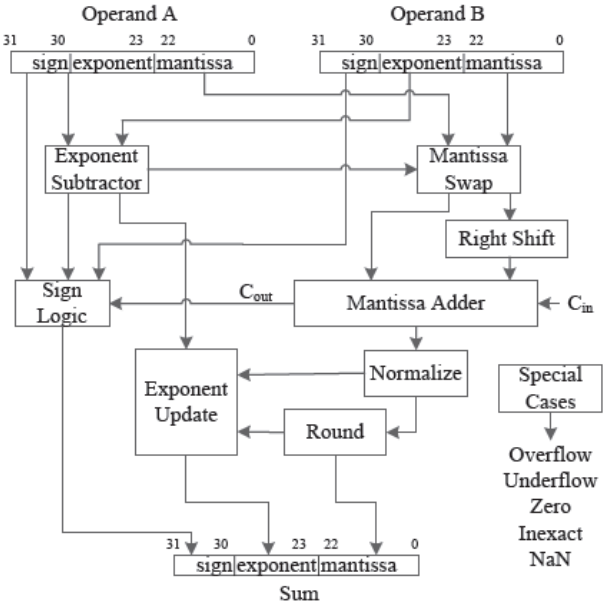


Fig. 1: The accurate single precision floating-point adder architecture

A. Exponent Subtractor

The exponent is dominant in the floating-point format because it determines the dynamic range. Due to its importance in the number format, an approximate design of the exponent subtractor must be carefully considered. The results can be significantly affected by applying the approximate design to only the least significant bit (LSB) of the exponent subtractor.

B. Mantissa Adder

The mantissa is less significant than the exponent; therefore, it is more appropriate to consider using an inexact mantissa adder in an inexact floating-point adder. The mantissa adder is also larger than the exponent subtractor, so it offers a larger design space. The mantissa adder is modified to an inexact design in the proposed approach. Different inexact fixed-point adders have been proposed and can be used in the mantissa adder, such as lower-part-OR adders (LOA) [2], approximate mirror adders [3] and approximate XOR/XNOR-based adders [4].

C. Normalization and Leading Zero Counting

Normalization is necessary to ensure that the additional results are in the correct range; the sum or difference may be too small and a multi-bit left shift may be required. A reduction of the exponent is also necessary. The normalization is performed based on a leading zeros/ones counter that determines the required number of left shifts. As the mantissa adder is not exact for the least significant n bits, the detection of the leading zeros can also be simplified in an approximate design; therefore, an approximate leading zero detection logic can be used. An approximate fixed-point added structure is given below in fig 2.

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 2 ISSUE 1 (2017) PAGES 77 - 82
Received: 1/09/2017. Published: 26/09/2017

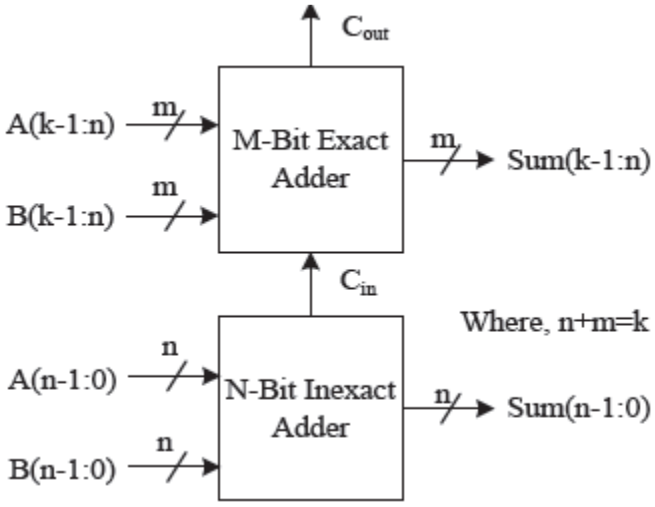


Fig. 2 An approximate fixed-point adder structure

D. Rounding

A proper rounding maintains three extra bits (i.e. guard bit, round bit and sticky bit); however, the adder may require another normalization and exponent adjustment after the rounding step. It is clear that the hardware overhead for rounding is significant. However, it does not affect the results of the inexact addition as the lower significant n bits are already inexact. Therefore, rounding can likely be ignored in approximate floating-point adders and they can be designed based on the above discussion and as outlined in the architecture given below in Fig. 3.

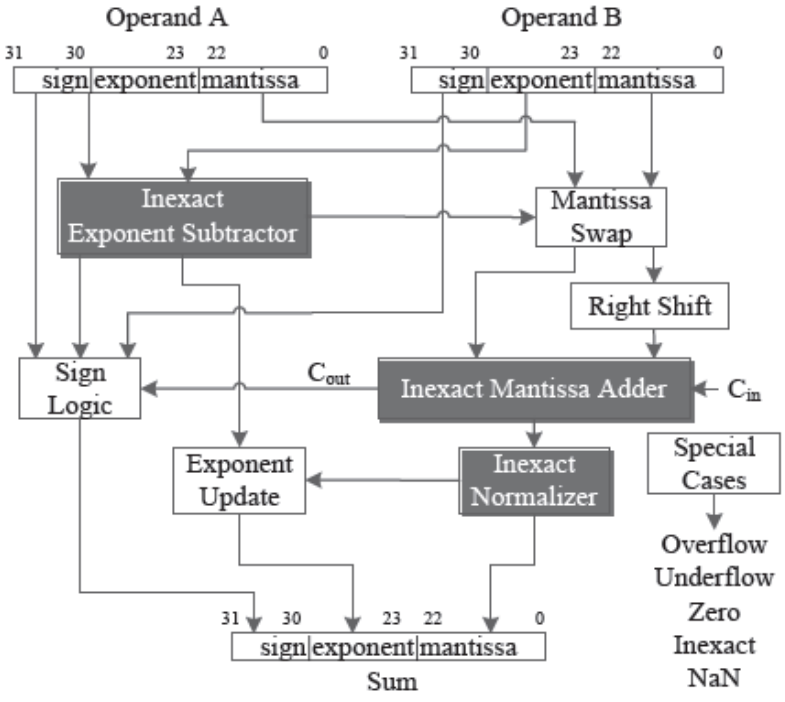
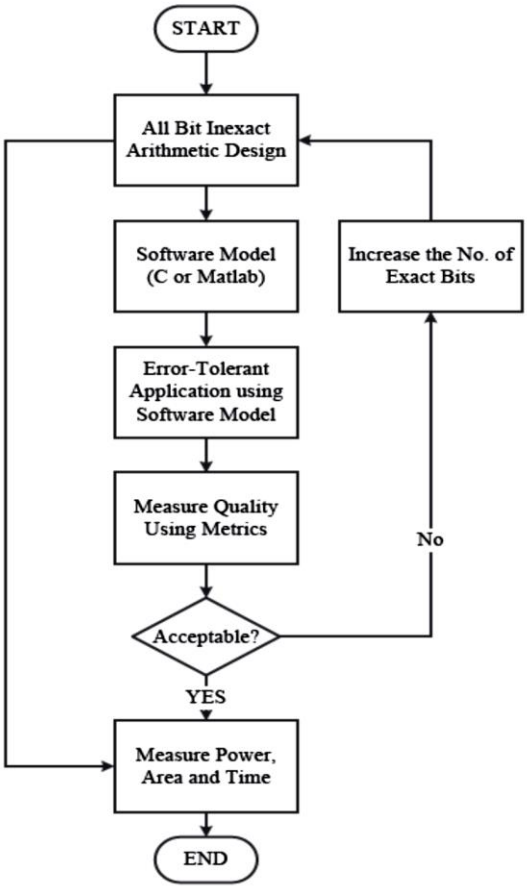


Fig. 3 The inexact single precision floating-point adder architecture

International Journal of Global Engineering (IJGE) E- ISSN: 2456-3099
VOL 2 ISSUE 1 (2017) PAGES 77 - 82
Received: 1/09/2017. Published: 26/09/2017

IV. Inexact Circuit Design Methodology

An extreme case of an inexact mantissa adder is a design in which all bits in the adders are inexact. The accuracy can then be improved by using more exact adders in the mantissa adder until the accuracy requirement is met for a specific application. A design procedure for an inexact floating-point adder is shown below.



The design methodology for an inexact floating-point adder.

An inexact floating-point adder design can be modeled using a hardware description language such as VHDL or Verilog. Then, the design is simulated and synthesized to assess if the preliminary performance results such as area, time and power dissipation of the inexact design are suitable for a specific application. The approximate design must be remodeled in software for the targeted error-tolerant applications. The results are then measured to meet the desired metrics and the accuracy can be adjusted by increasing the number of exact bits in the adders and changing the related logic accordingly; hence, the desired inexact design can be achieved through this iterative process.

V. Conclusion

Inexact floating-point adder designs have been investigated in this paper. Approximate designs of the mantissa adder and exponent adder are proposed and consideration is given to the related

VOL 2 ISSUE 1 (2017) PAGES 77 - 82

Received: 1/09/2017. Published: 26/09/2017

normalization and rounding logic. A design procedure has also been proposed to guide the general inexact design of arithmetic circuits. Two extreme cases for the inexact design of floating-point adders have been studied. The first design uses an all-bit inexact mantissa adder. The second design uses an inexact LSB in the exponent subtraction. Both designs are applied to high dynamic range images and the results show that both inexact floating-point adders are very low power designs, require a small area and offer higher performance than their equivalent exact designs, and thus, are suitable for high dynamic image applications. It is shown that the exponent part is a dominant part of the floating-point number format and provides a small design space for inexact design compared to the mantissa adder. Related logic such as rounding also plays an important role in inexact designs.

References

[1] A. Lingamneni, etc. "Algorithmic methodologies for ultra-efficient inexact architectures for sustaining technology scaling." Proc. ACM Int. Conf. Computing Frontiers, pp.3-12, 2012.

[2] H. Mahdiani, A. Ahmadi, S. Fakhraie, and C. Lucas, “Bio-inspired imprecise computational blocks for efficient VLSI implementation of soft-computing applications,” IEEE Trans. Circuits Syst. I, Reg. Papers, vol. 57, pp. 850-862, 2010.

[3] V. Gupta, D. Mohapatra, S. Park, A. Raghunathan, and K. Roy, “IMPACT: IMPrecise Adders for Low-Power Approximate Computing,” Proc. Int. Symp. Low Power Electronics and Design (ISLPED), pp. 1-3, 2011.

[4] Z. Yang, A. Jain, J. Liang, J. Han and F. Lombardi, “Approximate XOR XNOR-based Adders for Inexact Computing”, Proc. 13th IEEE Conf. Nanotechnol. (IEEE-NANO), pp.690-693, 2013

[5] J. Y. Tong, D. Nagle, and R. Rutenbar, “Reducing power by optimizing the necessary precision/range of floating-point arithmetic”, IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 8, pp. 273-286, 2000.

[6] A. Gupta, S. Mandavalli, V. Mooney, K. Ling, A. Basu, H. Johan, and B. Tandianus, “Low power probabilistic floating-point multiplier design”, Proc. IEEE Comput. Soc. Annu. Symp. VLSI, pp. 182-187, 2011.