Hardware-Efficient JPG Decompression on an Embedded Multi-Core picoMIPS Architecture

??????????????????????

***Abstract*—This paper reviews and summarizes recent energy saving embedded processor design techniques and JPEG decompression techniques on multicore architectures. Several researches are given to evaluate heterogeneous and reconfigurable embedded many-core system. After it, a brief introduction of the picoMiPS architecture is given. Finally, a research about a novel JPEG decoding scheme for heterogeneous architectures is critical reviewed.**

***Index Terms*—Embedded processors, MIPS architecture, JPEG decompression, hardware efficient design.**

I. Introduction

The energy and area efficient of embedded processor becomes more and more important to meet the requirement of related electronics. For instance, consider mobile electronics. Their feature of small size requires an area efficient embedded processor. Besides, as mobile electronics are usually powered by batteries, an energy efficient processor is required to ensure the battery life. Also, high performance such as powerful graphics capability is one of the most important requirements. Therefore, new techniques on embedded processor should be studied to meet these requirements. Compared to single core, multi-core structure is necessary to meet the demands' trend of increasing a processor's speed. However, power consumption and dissipation becomes a problem which should be concerned due to the additional cores [6]. Heterogeneous and reconfigurable embedded many-core system, such as the ARM big.LITTLE architecture [1], can improve the energy efficiency and maintain the high performance at the same time. Homogeneous multi-core systems have only the same cores whereas heterogeneous multi-core systems may have different types of cores [2]. Through using flexible interconnect, power gating and software control within each core, reconfigurable architectures can achieve heterogeneity itself.

The picoMIPS architecture is inspired by MIPS. It is a variable-architecture Reduced Instruction Set Computer (RISC) microprocessor. In order to maximize area efficiency, only the necessary datapath elements will be contained in each implementation. Besides, it can implement many of energy-efficient technologies. Through implementing a heterogeneous architecture, the picoMIPS can process multiple different applications. A reconfigurable architecture can also be implemented to create a general purpose but application specific processor effectively.

The JPGE format is the compression standard for digital images. JPEG images are widely used in the popular websites such as Google, Facebook and Baidu [3]. JPEG decoding is computationally expensive which consists of Huffman decompression, dequantization, IDCT, image upsampling and YCbCr to RGB colour space conversion [4]. By using a heterogeneous architecture consisting of a CPU and an OpenCL programmable GPU, it can achieve up to 95% of the theoretically attainable maximal speedup [5].

Manuscript received April 6, 2016.

X.Liu is with the Electronics and Computer Science Department, University of Southampton, UK (e-mail: {xl10g15@ecs.soton.ac.uk).

II. Heterogeneous And Reconfigurable Multicore Architectures

This section presents more details about the heterogeneous and reconfigurable architectures, which aims at achieving a greater performance, energy efficiency and area efficiency balance. Heterogeneous and reconfigurable technologies aim at increasing the level of the balance [2].

The introduction of multi-core architecture increases the power consumption of the processor. Besides, the gap between average power and peak power has widened with the increasing of the level of core integration [6]. Several metrics are used to evaluate the energy efficiency of a system such as performance per watt [7], average and peak power or energy-delay product [8].

A. Heterogeneous Multi-core Architectures

A heterogeneous multi-core architecture consists of cores different size and complexity deigned to complement each other according to performance and energy efficiency [9]. Small cores in the system will be used to process simple and small tasks. This mainly affects the energy efficiency of the system. Large cores will be implemented to process complex and large tasks. This mainly affects the performance of the system. Varying methods, such as asymmetric core sizes [10], custom accelerators [11], varied caches sizes [12], and heterogeneity within each core [13], can be used to create a heterogeneous architecture.

The ARM big.LITTLE architecture [1] is one of the most

successful examples of heterogeneous architectures. The system is designed with the dynamic usage patterns of modern smart phones in mind. Therefore, there are periods of high intensity processing followed by longer periods of low intensity processing [14]. High intensity tasks will be handled by the A15 processor to increase the system's quantity and meet the performance requirements. Low intensity tasks will be handled by A7 processor to improve energy efficiency and save battery life.

Compared to heterogeneous multi-core architectures, all the cores of homogeneous multi-core architectures are the same. It is demonstrated that heterogeneous multi-core architecture can provide a significantly higher performance than a homogeneous multi-core architecture for a given fixed circuit area [15]. A heterogeneous multi-core architecture can match each application to the related core which is the best one suited to meet its performance requirements. Power consumption and heat dissipation are the key challenges in high performance and complex processor designs currently. The single-ISA heterogeneous multicore processor can significantly reduce processor power dissipation [15].

B. *Reconfigurable Multi-core Architectures*

Reconfigurable architecture is another structure which can improve energy and area efficiency of processors. It can adjust the complexity and performance level according to the currently executing application. Through this way, it makes the system more flexible and adaptable.

The objective of reconfigurable architectures is to achieve both energy efficiency and high performance within the same processor. There are several reconfiguration techniques. One is dynamic partial reconfiguration [16], which is now offered by FPGA manufacturers such as Xilinx and Altera. Another one is composable and partitionable architecture [17], which is employed by some dynamically reconfigurable systems to provide adaptive granularity. Coarse grained reconfigurable array architecture [18] is also one of the reconfiguration techniques.

C. *Both Heterogeneous and Reconfigurable Architectures*

The combination of heterogeneous and reconfigurable is also possible and meaningful. For example, Dynamic Core Morphing (DCM) [19] is one of techniques using both a heterogeneous architecture and hardware reconfiguration. Although cores in the system all feature a baseline configuration, reconfiguration can trigger the re-assignment of high performance functional units to different cores to improve the execution. Compared to a homogeneous static architecture, it can increase 43% performance per watt gains and save 16% energy.

III. PICOMIPS

The MIP architecture is a reduced instruction set computer

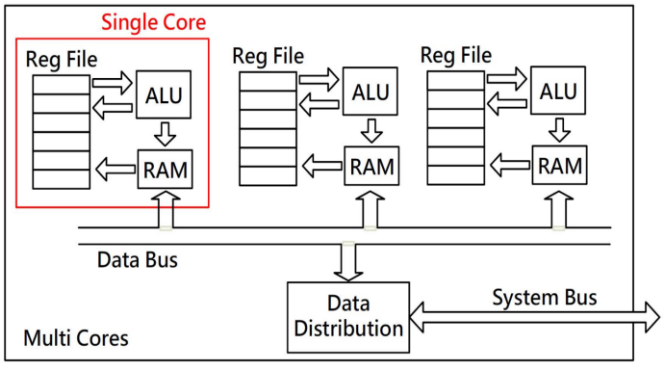

Fig. 1. A Multi-core implementation of the picoMIPS architecture, reprinted from [21]

(RISC) instruction set architecture (ISA). It is one of the simplest RISC architectures, whose characteristics are small hardware and low power consumption. The picoMIPS architecture is inspired by MIPS. It is a variable-architecture application-specific processor design methodology. Its main objective is to achieve extremely low energy consumption and it is usually implemented in a many-core small worker environment. Besides, it is a proposed architecture for embedded processors which usually run specific, fixed application types.

As a microprocessor architecture, the picoMIPS can implement many technologies to improve the energy efficiency and performance. Heterogeneous architectures can be implemented to allow the system process different application at the same time and reconfigurable architectures can be implemented to create application specific processors.

For a JPEG decompressor, processor based on the picoMIPS concept is more area efficient than a GPP, although its functionality is reduced [20]. An ASIC implementations have a much higher performance but cost more area. However, the hardware size is about four times compared with picoMIPS. Therefore, the picoMIPS represents a balance between the performance and area efficiency [20].

IV. JPEG DECOMPRESSION

The JPGE format is the compression standard for digital images. It is commonly used for images produced by digital photography. Therefore, JPEG images are widely used in the popular websites such as Google, Facebook and Baidu [3].

JPEG decoding is computationally expensive which consists of Huffman decompression, dequantization, IDCT, image upsampling and YCbCr to RGB colour space conversion [4]. Figure 2 shows the software architecture of libjpeg-turbo. A sub-class of Huffman codes are suitable for decoding multiple chunks of the encoded bitstream in parallel [21]. After Huffman decoding, a buffer of decoded data is transferred from the CPU to the GPU. The IDCT, upsampling and colour conversion kernels are invoked subsequently.

The discrete cosine transform (DCT) algorithm can be a

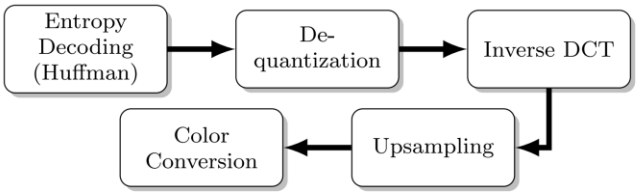

Fig. 2. JPEG decoder path [5]

stage in JPEG compression. The DCT and IDCT has been performed in a multi-core context through a picoMIPS [22]. The structure of the system is as shown in figure 1. The single core is a normal picoMIPS architecture. It is replicated 3 time. All of the cores in the system are connected to a data distribution module through a data bus. The block data is transferred to each core in turn. This design makes it possible to process multiple data in parallel. Therefore, the throughput of the system can be tripled theoretically.

The chrominance colour space is usually different from the luminance space during JPEG encoding. Hence, the sample rates of these colour spaces must be upsampled to the original size. The final stage of the JPEG decoder is to convert the YCbCr colour space to the RGB colour space according to related algorithm.

A novel JPEG decoding scheme for heterogeneous architectures consisting of a CPU and an OpenCL programmable GPU achieves up to 95% of the theoretically attainable speedup [5]. Heterogeneous JPEG image decompression can improve image viewing experience ranging from personal photos to very large image applications.

V. Conclusion

As a variable-architecture application-specific approach to embedded processor, the picoMIPS architecture has been presented. Compared with other architectures such as corresponding FPGA soft-cores and dedicated ASIC hardware, the picoMIPS represents a balance between the performance for area and energy efficiency benefits.

Besides, the heterogeneous and reconfigurable architectures have been introduced as the common energy efficient technologies. The development of the processor architectures has evolved from single core to homogeneous multi-core and into heterogeneous multi-core nowadays.

Normally, three main points should be considered when designing a processor related to this project, which are performance, power consumption and area efficiency. The picoMIPS architecture will be supportive to find a balance among all of them. Besides, by implementing the heterogeneous and reconfigurable architectures, the system will gain a further improvement.

References

[1] P. Greenhalgh, “big. LITTLE processing with ARM coretex-A15 & Cortex-A7,” ARM White Paper, 2011.

[2] R. Kumar, D. Tullsen, P. Ranganathan, N. Jouppi, and K. Farkas, “Single-ISA heterogeneous multi-core architectures for multithreaded workload performance,” in Computer Architectures, 2004. Proceedings. 31st Annual International Symposium on, 2004, pp. 64-75.

[3] Alexa Top 500 Global Sites. http://www.alex.com/topsites, retrieved April. 2016.

[4] J. Hing et al. Design, implementation and evaluation and evaluation of task-parallel JPEG decoder for the libjpeg-turbo library. *International journal of Multimedia and Ubiquitous Engineering*, 7(2), 2012.

[5] S. Wasuwee, H. Jingun, C. Seongwook, L. Yeongkyu, K. Shin, and B. Bernd, “Dynamic Partitioning-based JPEG Decompression on Heterogeneous Multicore Architectures,” 2014.

[6] C. Isci, A. Buyuktosunglu, C. Y. Chen, P. Bose, and M. Martonosi, “An Analysis of Efficient Multi-Core Global Power Management Policies: Maximizing Performance for a Given Power Budget,” in *Microarchitecture*, 2006. pp.347-358.

[7] R. Rodrigus and A. Annamalai, “Performance Per Watt Benefits of Dynamic Core Morphging in Asymmetric Multicores,” *in Parallel Architectures and Compilation Techniques (PACT)*, 2011.

[8] R. Kumar, and D. Tullsen, “Single-ISA heterogeneous multi-core architectures for multithreaded workload performance,” in *Computer Architecture*, 2004.

[9] R. Kumar, V. Zyuban, and D. Tullsen, “Interconnections in multi-core architectures: understanding mechanisms, overheads and scaling,” in *Computer Architecture*, 2005. ISCA ’05. *Proceedings. 32nd International Symposium on*, 2005, pp 400-420.

[10] D. . Woo, and H. –H. Lee, “Extending Amdahl’s Law for Energy Efficiency Computing in the Many-Core Era,” *Computer*, vol. 41, no. 12, 2008.

[11] H. M. Waidyasooriya, and M. Kameyama, “FPGA implementation of heterogeneous multicore platform with SIMD/MIMD custom accelerators,” in *Circuits and Systems* (ISCAS), 2012 IEEE *International Symposium* on, 2012, pp. 1339-1342.

[12] B. de Abreu Silva, L. Cuminato, and V. Bonato, “Reducing the overall cache miss rate using different cache size for Heterogeneous multi-core Processors,” 2012.

[13] A. Luckefahr and S. Mahlke, “Composite Cores: Pushing Heterogeneity Into a Core,” in Microarchitecture (MICRO). 2012 45th Annual IEEE/ACM International Symposium on, 2012, pp. 317-328.

[14] B. Jeff, “Advances in big.LITTLE Technology for Power and Energy Savings,” ARM, Tech. Rep., 2012.

[15] H. Abeer, “A comparative Study on Heterogeneous and Homogeneous Multiprocessors,” Report research, University of Jordan, Computer Engineering Department, 2009

[16] M. Huebner, S. Schuck, M. Kuhnle, and J. Becker, “New 2-Dimensional Partial Dynamic Reconfiguration Techniques for Real-time Adaptive Microelectronic Circuits,” *Proc. Of Emerging VLSI Technologies and Architectures, Karlsruhe, Germany*, 2006.

[17] B. Akesson, A. Hansson, and K. Goossens, “Composable resource sharing based on latency-rate servers,” in *Digital System Design. 12th Euromicro Conference on*, IEEE, 2009.

[18] Ahn, M., Yoon, J.W., Paek, Y., Kim, Y., Kiemb, M., Choi, K.: A spayial mapping algorithm for heterogeneous coarse-grained reconfigurable architectures. In: DATE ’06: *Proceedings of the Conference on Design, Automation and Test in Europe*, pp. 363-368 (2006).

[19] A. Das, R. Rodrigues, I. Koren and S. Kundu, “A Study on the Performance Benefits of Core Morphing in an Asymmetric Multicore Processor,” International Conference on Computer Design, 2010.

[20] L. Charles, and J.K. Tom, “Energy Efficient Multi-Core Processing,” University of Southampton, Electronics, Vol. 18, NO. 1, June 2014.

[21] S. T. Klein and Y. Wiseman. Parallel Huffman decoding with applications to JPEG files. Computer J., 46:487-497, 2003.

[22] G. Liu, “Fpga implementation of 2d-dct/idct algorithm using multicore picomips,” Master’s thesis, University of Southampton, School of Electronics and Computer Science, September 2013.