

An OpenCL-based Acceleration for Canny Algorithm Using a Heterogeneous CPU-FPGA Platform

Samah Rahamneh and Lina Sawalha

Electrical and Computer Engineering, Western Michigan University

Email: {samah.z.rahamneh, lina.sawalha}@wmich.edu

Abstract—Field programmable gate arrays (FPGAs) provide both performance and power benefits to heterogeneous systems. In this work, we used a closely-coupled CPU-FPGA heterogeneous system to accelerate Canny edge detector algorithm and compared the performance of the hybrid implementation with that of the optimized separate CPU and FPGA implementations.

Introduction: The diversity of workload characteristics stimulated the deployment of heterogeneous architectures to accommodate the disparity of applications [1]. Field Programmable Gate Arrays (FPGAs) have advantages over other accelerators because of their power and performance benefits [2]. Edge detection algorithms are among the widely used algorithms in image processing applications such as computer vision and image segmentation [3]. We propose a hybrid CPU-FPGA algorithm to accelerate Canny edge detector on a heterogeneous using OpenCL. We utilize a delay based Weighted Round Robin (WRR) algorithm to partition and distribute images between the CPU and the FPGA. We used Intel’s Hardware Research Acceleration Program (HARP) to implement the hybrid accelerator. HARP cluster nodes combine Intel Xeon processor and Arria 10 GX 1150 FPGA on a multi-packaged chip, connecting through PCIe and Quick Path (QPI) interconnects [4]. Our results shows increased system performance compared to CPU-only and FPGA only implementations.

Hybrid CPU-FPGA Acceleration for Canny Algorithm: The hybrid CPU-FPGA implementation aimed for increasing the CPU-FPGA processing overlap by allowing both to execute the same kernel on different parts of the image simultaneously. We first split the image into tiles as shown in Figure 1. The tiles were then distributed to the CPU and the FPGA using WRR algorithm in a way that reduces execution time and increases system throughput. The weights are assigned to both the CPU and the FPGA using CPU to FPGA ratio of the execution time of a single tile (2:1). This delay-aware split of tiles between the CPU and FPGA boosts system utilization through the simultaneous handling of different data sets (tiles) by both the CPU and the FPGA. This method can be used

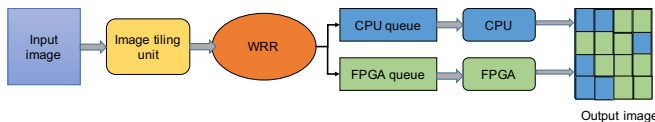


Fig. 1: Hybrid CPU-FPGA processing of images.

for many other heterogeneous architectures and algorithms to

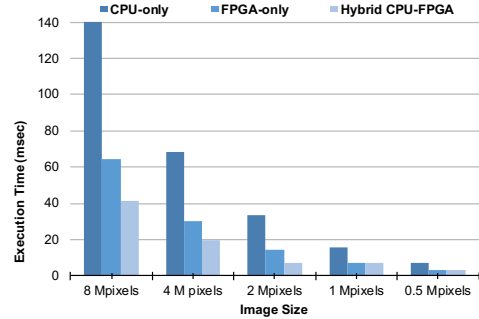


Fig. 2: Execution Time for CPU-only, FPGA-only, and CPU-FPGA Hybrid Implementations.

balance the load of the different architectures and enhance performance.

Experimental Results and Discussion: Figure 2 shows the performance achieved using our tile-based hybrid implementation over CPU and FPGA implementations for different image sizes. For example, using a two-megapixel (2MP) image, the speedup gained by the hybrid implementation is 4.8X over a CPU-only and 2.1X over a FPGA-only implementations. However, for a 0.5MP image, the hybrid implementation result in 2.1X speedup over the CPU and no noticeable speedup over the FPGA-only implementation. This is because as we tend to process small images, the CPU becomes a bottleneck and its execution time can become dominant over the FPGA’s execution time. The FPGA consumes its data while the CPU is still processing its part. The CPU bottleneck can be solved by assigning the FPGA a higher weight leaving the CPU with only a smaller portion of the image frame (smaller tiles).

Acknowledgment: This work was funded by the National Science Foundation Grant No. 1821691, and Intel Inc. equipment access.

REFERENCES

- [1] B. Varghese and R. Buyya, “Next generation cloud computing: New trends and research directions,” *Future Generation Computer Systems*, vol. 79, pp. 849–861, Feb. 2018.
- [2] P. K. Gupta, “Xeon+ fpga platform for the data center,” Presentation at the Fourth Workshop on the Intersections of Computer Architecture and Reconfigurable Logic, Aug. 2015.
- [3] L. Shao, R. Yan, X. Li, and Y. Liu, “From heuristic optimization to dictionary learning: A review and comprehensive comparison of image denoising algorithms,” *IEEE Transactions on Cybernetics*, vol. 44, no. 7, pp. 1001–1013, Jul. 2014.
- [4] admin, “Hardware Accelerator Research Program.” [Online]. Available: <https://software.intel.com/en-us/hardware-accelerator-research-program>