Learning to Approximate Computing at Run-time

Paulo Garcia, Mehryar Emambakhsh, Andrew Wallace

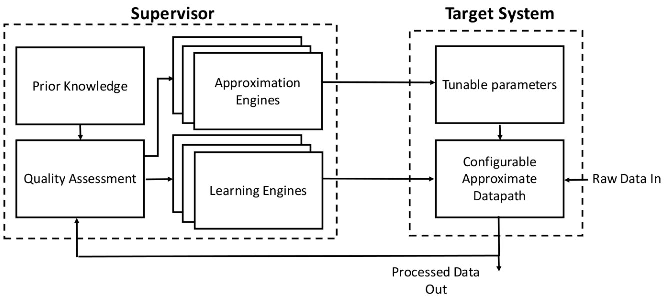

Fig. 1: Block diagram

***Abstract*—Intelligent sensor/signal processing systems are increasingly constrained by tight power budgets, especially when deployed in mobile/remote environments. Approximate computing is the process of adaptively compromising over the accuracy of a systems output in order to obtain higher performance for other metrics, such as power consumption or memory usage, for applications resilient to inaccurate computations. It is, however, usually statically implemented, based on heuristics and testing loops, which prevents switching between different approximations at run-time. This limits approximation versatility and results in under- or over-approximated systems for the specific input data, causing excessive power usage and insufficient accuracy, respectively. To avoid these issues, this paper proposes a new approximate computing approach by introducing a supervisor block embedding prior knowledge about runtime data. The target system (i.e., signal processing pipeline) is implemented with configurable levels and types of approximations [1]. Data processed by the target system is analysed by the supervisor and the approximation is updated dynamically, by using prior knowledge to establish a confidence measure on the accuracy of the computed results. Moreover, by iteratively evaluating the output, the supervisor block can learn and subsequently update tunable parameters, in order to improve the quality of the results. Our approach also envisions switching between multiple approximation and learning engines at run-time. We detail and evaluate this approach for tracking problem in computer vision. Results show our approach yields promising trade-offs between accuracy and power consumption.**

***Index Terms*—field programmable gate array (FPGA), optimisations, power, image processing, dataflow**

I. Introduction

blah blah

P. Garcia, M. Emambakhsh and A. Wallace are with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, U.K. E-mail: {p.garcia, a.m.wallace}@hw.ac.uk.

Manuscript received XXX XX, XXXX; revised XXX XX, XXXX.

II. Background and related work

A. Approximate computing

B. Prior knowledge in sensing

III. Prior knowledge for runtime approximations

IV. Proof of concept: extended Kalman filter

V. Experimental results

VI. Conclusions and future work

Acknowledgment

We acknowledge the support of the Engineering and Physical Research Council, grant references EP/K009931/1 (Programmable embedded platforms for remote and compute intensive image processing applications) and EP/K014277/1 (MOD University Defence Research Collaboration in Signal Processing).

References

[1] V. Vassiliadis, K. Parasyris, C. Chalios, C. D. Antonopoulos, S. Lalis, N. Bellas, H. Vandierendonck, and D. S. Nikolopoulos, "A programming model and runtime system for significance-aware energy-efficient computing," *SIGPLAN Not.*, vol. 50, no. 8, pp. 275–276, Jan. 2015. [Online]. Available: http://doi.acm.org/10.1145/2858788.2688546