Walid WALID
PhD Programme in Electrical, Electronics and Communications Engineering
Department of Electronics and Telecommunications, Politecnico di Torino
Supervisors: Prof. Maurizio MARTINA, Prof. Guido MASERA

VLSI Architectures for Video Processing and RISC-V

In the internet and smartphone era, video conferencing and live streams are a big part of information transmission. Video processing involves many domains, from high-quality video transmissions to real-time decision-making systems. The usual software-based models and serial approach running on single-issue cores are slow. Hence, there is a need for speed up to make it compatible with real-time applications. The video processing tasks have parallelism that can be exploited for this speedup via the hardware acceleration. Thus, this work is aimed towards achieving this task. In this aspect, two algorithms were targeted for hardware acceleration. In addition, the processor domain is also explored in order to select cores for video processing.

The first algorithm is in the domain of object tracking. Tracking objects in real-time is a crucial step in image processing. Thus the hardware real-time object tracker for several image resolutions should provide good accuracy alongside meeting the frame rate requirements. The sophisticated algorithms employed in the state-of-the-art are computationally demanding for embedded architectures. An algorithm with good performance and high parallelism is discriminative scale space tracking. The parallelism can be exploited for speedup with hardware acceleration. The first part of this work proposes hardware architectures for the major blocks to attain the real-time requirements of a discriminative scale-space tracker on FPGA. The main contributions in this work are the improvements in the architecture for the core mathematical functions in the algorithm, including the discrete Fourier transform, QR factorization, and singular value decomposition. For 320×240 image resolution, the proposed architecture obtains a mean 25.38 fps.

The second algorithm is in the domain of video encoding. In recent years, many devices have been using videos for communication. With the adaptation of high-quality video and image content, the need for more memory space is increased. It also demands more data compression from the advancing video coding techniques. These demands led to the development of AOMedia Video 1 or AV1 by the Alliance for Open Media. AV1 is a royalty-free codec, thus allowing it to be used in many devices without paying extra cost. This work also studies the AV1 model and selects the Wiener filter as a complex block requiring acceleration. Wiener filter is an in-loop

restoration tool in AV1. This work presents a separable symmetric normalized architecture for its implementation. The proposed design allows $100 \times$ reduction in processing time and $5 \times$ speedup in mega samples per second concerning existing works in literature.

The thesis's third part concerns the processor core for video processing applications. Many Instruction Set Architectures are available but often closed and incremental Instruction Set Architectures. On the contrary, RISC-V is the new open and royalty-free Instruction Set Architecture developed at UC Berkeley in 2010. Apart from being an open architecture, RISC-V is a modular Instruction Set Architecture. Starting from the base Instruction Set Architectures, any other extension can be implemented as per need without implementing all the previous ones. This modularity allows RISC-V to be adaptable for video processing tasks. This work first identifies the core requirement for video processing tasks based on which a processor can be selected. Next, it explores the RISC-V cores in literature based on the selected parameters. Afterward, it compares them and selects those most suitable for video processing. In addition, the work also extends a RISC-V core in literature called LEN5 for video processing.