Evaluation of PhD Thesis
HARDWARE ACCELERATION OF OBJECT DETECTION IN IMAGES

Author of Thesis: ing. Petr Musil
SUPERVISOR Prof. Dr. Ing. Pavel Zemčik

Evaluator: Ing. Jiri Kadlec, kadlec@utia.cas.cz

The thesis presents research and technology development contributions of ing. Petr Musil to the state-of-the-art in the topic of visual object detection in FPGA. His research work was focused on fast and powerful object detectors with low demands on resources. Such detectors could be applied mainly in transport, industry or security. One of the applications of the detector is be demonstrated on the task of detecting license plates for parking control in residential zones. The benefits of his implementation results are shown in comparison to current technologies.

Ing. Petr Musil is proposing methods for optimizing object detection on FPGAs. His main focus is on detectors using boosted soft cascades of classifiers with local image features as weak classifiers. He managed to upgrade sequential evaluation of weak classifiers with parallelization by evaluation of several independent image positions simultaneously. He also introduced novel approach for multi-scale object detection with no need for external memory. Using these methods he demonstrated, that it is possible to design an object detector based on soft cascade deployed in programmable hardware with resulting precision comparable to the state-of-the-art, with real-time performance, with lower power consumption and less computing resource demands comparing to existing ones.

His advances in comparison to the state of the art in this field of research are:

- better detection performance among boosted classifier – multi-scale face detection on Full HD (1920×1080 pixels) video at 60 fps.
- better detection performance in processed detection windows per clock cycle among all hardware detectors
- better performance/resources ratio
- better accuracy in face detection on CMU dataset
- comparable accuracy in unconstrained licence plate detection in comparison to the easier aligned licence plates detection

Ing. Petr Musil is co-author of three papers where the research results and omplemented technology is described. These papers are parts of the thesis.

High performance FPGA object detector: Hardware prototype, FPL 2013. Paper introduce an architecture of an engine for high-performance multi-scale detection of objects in videos based on WaldBoost training algorithm. The key properties of the architecture include the processing of streamed data and low resource consumption. The engine is implemented in FPGA and that it can process 640×480pixel video streams at over 160 fps without the need of external memory. Main contribution of ing. Petr Musil in this paper is design of a memory structure for efficient reading of image values per block for effective evaluation of a weak classifier. He suggested on-the-fly multi-scale detection and implemented a detector simulator in C language to verify detector properties. He also implemented and tested the detector in the VHDL language.

Cascaded Stripe Memory Engines for Multi-Scale Object Detection in FPGA, TCSVT 2019. Evolution of the previous paper witch expands performance and usability. FPGA detector can process a stream of image data so that it stores a narrow stripe of the input image and its scaled versions and uses a detector unit which is efficiently pipelined across multiple image positions within the memory. Paper shows how to process images with up to 4K resolution at

high framerates using cascades engine. As a detector algorithm use boosted soft cascade with simple image features that require only pixel comparisons and look-up tables; therefore, they are well suitable for hardware implementation. Main contribution of ing. Petr Musil in this paper is the proposal of the cascade connection of detectors.He modified the detector hardware implementation in VHDL and created a tool to generate the optimal distribution for the cascade of detectors. The properties were tested by conducting a set of experiments on various cascade detector configurations.

Unconstrained License Plate Detection in FPGA, submitted to VEHITS3. This paper shows the practical use of the previous detector in traffic application on the task of detecting unconstrained License Plate. To detect and localize license plates is use multiple sliding window detectors based on simple image features, each tuned to a certain range of projections. On a large dataset is detection rate 98%. Main contribution of Ing Petr Musil in this paper (this paper is still under review) is the development a modification of the previously published detector for processing several different classifiers. He also trained selected standard CNN and ACF detectors to compare their accuracy with developed detectors.

Ing. Petr Musil is co author of another three papers:

- *Single-Loop Approach to 2-D Wavelet Lifting with JPEG 2000 Compatibility*, SBACPADW4 2015
- *High Dynamic Range Video Concepts, Technologies and Applications*, Real-Time HDR Video Processing and Compression Using an FPGA, 2016
- *True HDR camera with bilateral filter based tone mapping*, SCCG5 2017

These three publications do not directly contribute to the scientific goal of the dissertation. However, technologically they add to the options of using object detection in images. In some applications, for example, in bad light conditions such as sharp backlight, it is advantageous to combine object detection with HDR image processing to improve accuracy. The platforms created in these publications were used for the experimental work with object detection in this dissertation.

It is clear, that the presented scientific contribution of Ing. Petr Musil contributed to the technical goal to create an object-in-image hardware detector usable in practice. I see significant benefit is the scalability of performance and relatively low resource consumption. Achieved results indicate, that it becomes possible to implement detector, which can process the FullHD video at 60 fps on FPGA with a current price of approximately 100 USD. The proposed hardware detector is also the first presented solution for the detection of all-sized objects, at 4K resolution. The proposed detectors are expected to be utilised in smart cameras in industrial, transport or security applications. Detectors can be used for the detection of faces, pedestrians, products, licence plates.

Based on this evaluation I recommend this thesis for successful defence.

Prague 17.4 2021

Ing. Jiri Kadlec, CSc