POLITECNICO DI TORINO
Dipartimento di Automatica e Informatica

PhD in Computer and Control Engineering

XXIX cycle

Advisor
*Gianpiero Cabodi*

# Visual Analysis for Embedded System

PhD Candidate: *Denis Patti*

## 1. Abstract

Computer vision is a scientific discipline of computer science that includes the algorithms able to extract information from images or video sequences. Computer vision techniques are used in several industrial areas, such as the automotive field. Computer vision can be split into several applications such as object recognition within an image or motion analysis of a video track.

My purpose consists of finding new approaches of visual search for embedded system using GPU (Graphic processing unit). Computer vision algorithms often work at pixel level and therefore GPU architectures are ideal for these kinds of problems, because they allow applications to execute more operations in parallel than a CPU (SIMD which means Single Instruction Multiple Data).

## 2. Visual Control

Industrial and robotic controllers have to execute various complex independent tasks repeatedly in real time. In order to implement these algorithms with non-linear equations, massive computational power is required in a motion control system. In this project, inverse kinematics algorithm is selected as a test algorithm to measure performance of General Purpose Graphics Processing Unit (GPGPU) used to control a Robotic Arm with two joints. The performances are compared with other microcontroller in order to evaluate the gain obtained in terms of time [1].

The microcontrollers used for the comparison are FPGA and Arduino Due.

Execution speeds of these controllers are compared with NVIDIA Quadro K2200 GPU programmed with CUDA Parallel Computing Architecture.

Results validated that using computational power of GPU, execution time of large independent tasks is significantly decreased as shown in Fig. 1.

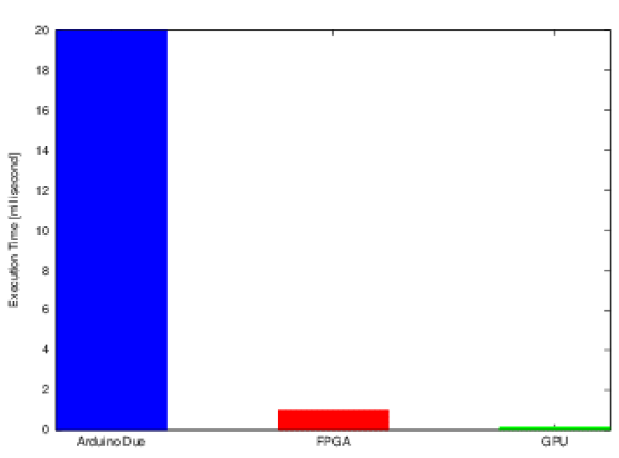

Figure 1: Comparison embedded device

## 3. Feature Extraction

Recent developments in embedded processors have enabled heterogeneous computing on mobile devices using open-access general-purpose computing languages. Following the MPEG CDVS standard, this work proposes an efficient feature computation phase, completely implemented on embedded devices supporting the OpenCL framework [2]. This work is based on MPEG-CDVS standard showed in Fig 2.

I worked on the CDVS detector and its design for multicore parallel GPUs. The original procedure has been adjusted using algorithmic choices and implementation details to target the intrinsic characteristics of the embedded platforms selected. The GPGPU implementation is compared with the ALP keypoint detector with the CPU based implementation of the CDVS standard. I collected data on different mobile GPUs which demonstrate that GPU solution is up to 7x faster than the CPU version. To sum up, one of the main feature of this new algorithm is to be fast enough to be able to open new visual search scenarios exploiting entire real-time on-board computations with no data transfer.

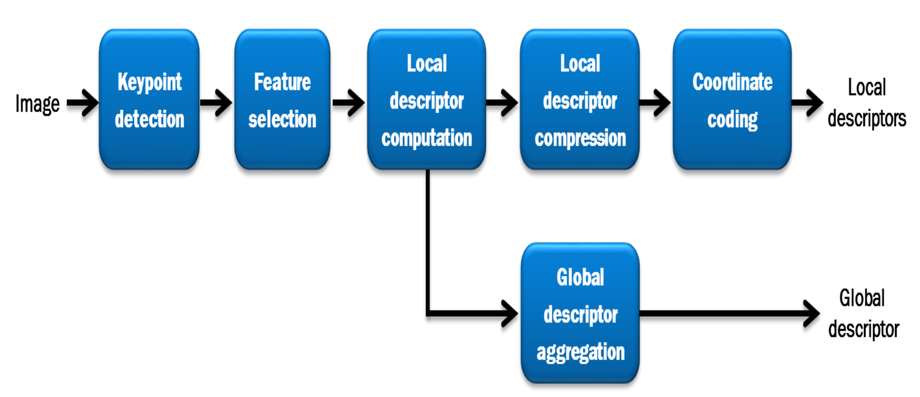


Figure 2: CDVS Model

## 4. Image Classification

Deep convolutional neural networks achieve state-of-the-art performance in image classification. The computational and memory requirements of such networks are however huge, and that is an issue on embedded devices due to their constraints. Most of this complexity sources from the convolutional layers, and in particular from the matrix multiplications they entail.

This work proposes a complete approach to image classification providing common layers used in neural networks [3].

Namely, proposed approach relies on a heterogeneous CPU-GPU scheme for performing convolutions in the transform domain.

CUDA-based implementation of proposed approach is evaluated over three different image classification networks on a Tegra K1 CPU-GPU mobile processor. Experiments show that the presented heterogeneous scheme boasts a 50x speedup (as shown in Table 1) over the CPU-only reference and outperforms a GPU-based reference by 2x while slashes the power consumption by nearly 30%.

Perspectively, the presented approaches can further benefit of most recent advances in embedded architectures.

Recent trends are in fact to resort to smaller filters to speed up the computationally intensive convolutional task, which can be Winograd-accelerated. Structure of the complete architecture can also be optimized using unified memory architecture of recent GPUs where extra memory transfers can be avoided by defining a more memory efficient scheme.

| Model | N° Layers | CPU time | GPU time |
|---|---|---|---|
| AlexNet | 5 | 10,20 s | 0,411 s |
| OverFeat Model | 8 | 418,04 s | 8,30 s |
| ResNet-34 | 34 | 248,36 s | 4,93 s |

Table 1: CPU vs GPU comparison

## 5. Conclusion

The results obtained in these different activities show that modern embedded system can have an important role in the future of automation control and computer vision.

## 6. Selected References

1. Rizvi Tahir et al., *Comparison of GPGPU based Robotic Manipulator with other Embedded Controllers*, DAS Conference, May 19-21 2016, Suceava
2. Massimo Balestri et al, *Accurate and Efficient Visual Search on Embedded System*, International conference on advanced in computing, communication and information technology, May 26-27 2015, Birmingham
3. Rizvi Tahir et al., *GPGPU Accelerated Deep Object Classification on a Heterogeneous Mobile Platform*, Electronics 5(4), 88, 2016