BMVA Symposium: High-Performance Computing for Computer Vision

One Day BMVA symposium: BCS London, UK on May 22, 2019

Chairs: Giuseppe Ciaccio and Nicoletta Noceti - Università degli Studi di Genova

The BMVA Technical meeting "*HPC4CV – High-Performance Computing for Computer Vision*" has been held in London on May, 22nd, 2019, and chaired by Giuseppe Ciaccio and Nicoletta Noceti, from the University of Genova (Italy). During the day, three keynotes and six talks were given, with question-and-answer sessions after each presentation.

The meeting rooted in the observation that in the last decades, Computer Vision has made significant advances in a high number of tasks and that state-of-art methodologies are now achieving astonishing results. This has been made possible thanks to the lucky conjunction of two main factors: on the one hand, the increasing ease in data acquisition enables the collection of large datasets, on the other the advances in computing technologies allow to manage the data and process them, with new high-performance platforms and paradigms that are gaining momentum. While solving a good variety of vision tasks, these advances also generate new research questions, while shedding light on "old" tasks considered intractable until a few years ago. The aim of the meeting was then to discuss the interplay between Computer Vision and High-Performance Computing, highlighting challenges in Computer Vision that can be better dealt with these cutting-edge technologies, and examples of working systems.

The meeting started at 9:30 am, with one of the chairs, Nicoletta Noceti, opening the sessions and introducing the main motivations of the meeting while identifying the expected outcomes of the day, among the others the hope of triggering discussions and curiosity, establishing networking between different research domains.

The scientific contributions opened with the first keynote speech, given by Prof. **Tae-Kyun Kim**, head of the Computer Vision and Learning Lab. of the Imperial College in London. Prof. Kim discussed Machine Learning for 3D vision, presenting advances made by his research group in the estimation of the 3D pose of objects and hands, crucial to the success of modern human-machine interfaces. While discussing their approach based on a hierarchical machine learning architecture combining convolutional neural network with decision forests, he highlighted the computational demand and the possible benefits of modern computing technologies.

After the first keynote, the talks of contributing speakers started. **Oscar Rahnama** (University of Oxford and FiveAI) opened, with a talk on depth estimation from stereo pairs. He observed that the task requires very intense computation, achieving real-time performance leveraging only GPUs that

require large amounts of power. He proposed a solution based on the use of FPGA-based platforms, for the real-time computation of depth maps in power-constrained applications, also discussing a heterogeneous CPU-FPGA platform to accelerate those tasks whose memory access patterns and complexity prevent their efficient deployment on the FPGA. The speaker discussed performance but also insights accumulated with regards to efficient algorithm design for the platform.

After the coffee break, **Matteo Poggi** (University of Bologna) continued the discussion on depth maps but focusing on its estimation from single images rather than pairs. He concentrated on the deep learning framework and started observing that, although providing state-of-art accuracy, this family of methods also requires power-hungry dedicated hardware to achieve real-time performance, with the consequence of being unsuited for a broad category of applications with constrained hardware resources or power consumption. The solution he proposed is based on a family of compact and effective Convolutional Neural Networks that leverage self-supervision and can run in real-time on a broad set of embedded or consumer devices.

The morning session was closed by the second keynote speech, given by **Juan Gómez Luna**, a research associate at the ETH in Zürich. He discussed the emergence in the last decade of GPGPUs devices as an opportunity for Computer Vision workloads to improve performance and energy efficiency by orders of magnitudes. After having presented an overview of his past efforts in such a domain, he introduced a novel paradigm to cope with the bottleneck represented by need of moving data between memory and compute units. In the so-called Processing-in-Memory (PIM) paradigm, computing systems move from a compute-centric to a memory-centric design, where compute units are placed closer to where the data resides. Juan concluded the talk with some recent research activities for enabling the adoption of PIM in future computing systems.

The afternoon sessions were opened by the third, and last, keynote speech. **Marta Betcke**, a lecturer at the University College of London, presented two common approaches to combine deep learning architectures, and in particular Convolutional Neural Networks, with model-based reconstruction techniques, with application to the problem of image reconstruction in photoacoustic tomography. She specifically discussed the time needed by the computation, considerably shorter than for well-established non-linear regularisation methods, as the Total Variation. Also, she highlighted how the techniques adopted can offer the opportunity of exploiting prior knowledge through training on a class of images with expected characteristics.

Alessio Tonioni (University of Bologna) presented his approach to perform unsupervised and continuous online adaptation of a deep stereo network that allows the preservation of an appropriate accuracy in any environment. In this way, he specifically addressed the problem of end-to-end deep convolutional networks that notably suffer when exposed to scenarios severely different from the training set. To allow for the computational feasibility of the approach, he introduced a new lightweight and effective architecture, called Modularity Adaptive Network (MADNet) that independently trains sub-portions of the network as soon as new data are available. He shows the very good performance of his method from both the efficiency and the accuracy standpoints.

The contribution of **Bolin Pan**[1] (University College of London) has been focused on photoacoustic tomography, with an investigation on Curvelet representations. Her interest was on the reconstruction of the photoacoustic image and the dynamics in space and time, coupling the first with sparsity-constrained motion estimation models based on optical flow. She demonstrated how the use of compressed sensing techniques allows for the acceleration of the photoacoustic reconstruction based on the sparsity in the Curvelet frame for static inversion from subsampled photoacoustic data. Next, she presented a joint variational approach that incorporates Curvelet sparsity in the photoacoustic image domain and spatiotemporal regularisation via optical flow constraint to achieve improved reconstruction results.

Optical flow was also one main ingredient of the work proposed by **Nargiza Djurabekova** (University College of London). She started observing that novel conebeam CT scanners allow imaging in both seated and upright positions which are both, to a different degree, weight-bearing, leading to considerable enhancement of diagnostically relevant information. Due to the lack of a precise biomechanical model, the main issue is how to interpret the functional weight-bearing motion from these static scans, which would be immensely helpful in performing correction surgeries. In her talk, Nargiza presented an approach for the reconstruction of a dynamic phantom of the foot and ankle structure considering a variational framework including the optical flow and total variation constraints to compensate for the sparse measurements.

Elena Nicora (University of Genova) concluded the afternoon sessions, with a talk centred on the possible helpfulness of High-Performance Computing in the robotics domain, where its application is still in its infancy. To this purpose, she delineated a vision on current and future robotics tasks, with particular emphasis on companion robots, highlighting trends and desiderata. She identified several different computationally demanding tasks, for which cutting edge technologies may be beneficial, foreseeing possible interesting future lines of research.

The chair Nicoletta Noceti closed the meeting, wrapping-up the main contributions.

In memory of Giuseppe Ciaccio, open mind, brilliant scientist.

Nicoletta Noceti
University of Genova
Nicoletta.noceti@unige.it

---

[1] The talk has been given by Marta Betcke, supervisor of Bolin Pan, for the impossibility of the speaker to join the meeting.