Title: Vision and The Eye: From Perceptual Evidence to Large-Scale Computational Models

Speaker: Cristian Sminchisescu, Lund University

Web: http://www.maths.lth.se/people/math-csu/

Abstract

Artificial intelligence and visual recognition are rapidly becoming transformative technologies through the widespread availability of digital cameras, computation power and big data. Nowadays big data is increasingly more image data. Most recent progress in computer-based visual recognition heavily relies on machine learning methods trained using large scale datasets annotated by humans, according to a general view of relevant visual `tasks' like image classification, visual object detection, semantic segmentation, or 3d pose reconstruction. While such data has made advances in model design and evaluation possible, it does not necessarily provide insights into those intermediate levels of computation, or deep structure, perceived as ultimately necessary in order to design reliable computer vision models that could operate in the real world, unconstrained. This is noticeable in the accuracy of state of the art systems trained with such annotations, which still lag significantly behind human performance in similar tasks. Nor does the existing data makes it immediately possible to exploit insights from a working system - the human eye - to derive potentially better features, models or algorithms.

In this talk I will present perceptual and computational insights resulted from the analysis of large-scale human eye movement datasets collected in the context of visual recognition tasks. I will show that computational attention models (fixation detectors, scan-paths estimators, weakly supervised object detector response functions and search strategies) can be learned from human eye movement data, and can produce state of the art results when used in end-to-end automatic visual recognition systems. I will conclude by reviewing some of the main mathematical challenges in making future progress and discuss several promising new research directions in the area of deep layered networks and reinforcement learning.

Bio

Cristian Sminchisescu is a Professor in the Department of Mathematics, Faculty of Engineering at Lund University, working in computer vision and machine learning. He has obtained a doctorate in computer science and applied mathematics with focus on imaging, vision and robotics at INRIA, France, under an Eiffel excellence doctoral fellowship of the French Ministry of Foreign Affairs, and has done postdoctoral research in the Artificial intelligence Laboratory at the University of Toronto. He holds a Professor equivalent title at the Romanian Academy and a Professor rank, status appointment at Toronto, and advises research at both institutions. During 2004-07, he has been a Faculty member at the Toyota Technological Institute, a philanthropically endowed computer science institute located at the University of Chicago, and later on the Faculty of the Institute for Numerical Simulation in the Mathematics Department at Bonn University.

Cristian Sminchisescu is a member in the program committees of the main conferences in computer vision and machine learning (CVPR, ICCV, ECCV, ICML, or NIPS), an Area Chair for ICCV, ACCV and CVPR during 2007-15, a Program Chair for ECCV 2018, and a member of the Editorial Board (Associate Editor) of IEEE Transactions for Pattern Analysis and Machine Intelligence (PAMI). He has offered tutorials on 3d tracking, recognition and optimization at ICCV and CVPR, the Chicago Machine Learning Summer School, the AEFRAI Vision School in Barcelona, the Computer Vision summer school at ETH in Zurich, and Visum in Porto. Over time, his work has been funded by the US

National Science Foundation, the Romanian Science Foundation, the German Science Foundation, the Swedish Science Foundation, the European Commission under a Marie Curie Excellence Grant, and recently, the European Research Council under an ERC Consolidator Grant. Cristian Sminchisescu's research interests are in the area of computer vision (articulated objects, 3d reconstruction, segmentation and recognition) and machine learning (optimization and sampling algorithms, structured prediction and kernel methods).