Visual Object Recognition - Traditional Methods Along with Deep Learning Approaches

Jan Flusser

Institute of Information Theory and Automation, Czech Republic

flusser@utia.cas.cz

The talk falls into the area of visual Artificial Intelligence (AI), particularly to image recognition by deep networks. In AI applications such as surveillance systems, autonomous robots, unmanned vehicles, drones, etc., cameras and other visual sensors form the “eyes” of the system while image recognition algorithms substitute the visual cortex of the brain. The key requirement is a continuous (possibly real-time) analysis of the visual field and, in that way, preparing the basis for decision and next action planning. The visual analysis may comprise scene segmentation, detection of objects and persons of interest, recognition of their identity and their behaviour, and even prediction of their next actions. In this talk, we focus on the recognition part, where the image/object is classified as a member of one of the pre-defined classes. Current convolutional networks work with inefficient pixel-wise image representation, which does not provide almost any invariance. This leads to the use of very large training sets and to massive augmentation. We propose to decompose intra-class variances into two degradation operators where one of them (image rotation, scaling, blurring, etc.) can be mathematically modelled by a superposition integral with a transformation of the coordinates. We further propose to design hybrid network architectures that use both pixel-level and newly developed high-level invariant image representations such that the high-level representation eliminates the influence of modelable degradations. This leads to a substantial reduction of the training data without sacrificing the recognition rate. The hybrid architectures could define a new standard in image-oriented networks.

14th Conference on DATA ANALYSIS METHODS FOR SOFTWARE SYSTEMS

23