

Séminaire MIDI : Florent Perronnin

19 Mai 2011, 13:30 – 15:00

Titre

Fisher Vectors for Large-Scale Image Classification

Date et salle

Jeudi 19 mai 2011, 13h30, à l'ENSEA, salle 384

Abstract

The Fisher kernel is a generic framework which combines the strengths of generative and discriminative approaches to pattern recognition. In a nutshell, it consists in characterizing a sample by its deviation from a generative probability model. This gradient vector — which we call the Fisher Vector (FV) — can be subsequently used as input to any discriminative classifier.

During this talk, I will discuss the application of this principle to the problem of image classification with an emphasis on large-scale problems. I will relate the FV to other image representations and show especially that it can be understood as a generalization of the popular bag-of-visual-words (BOV). I will show that it is highly scalable and report results on a subset of 9 million images and 10 thousand classes of ImageNet. Compared to the state-of-the-art on this dataset (a system based on spatial pyramid BOV representations) we report a significant 160% relative improvement: 16.7% accuracy versus 6.4%.