

Visual Understanding of Human Actions

Speaker: Prof. Hamed Pirsiavash
University of Maryland, Baltimore County
Monday, October 5, 2015
11AM- 12 noon, VSE Room 3507

Abstract

The aim of computer vision is to develop algorithms for computers to “see” and understand the world as humans do. Central to this goal is understanding human behavior; for instance, in order for a robot to interact with humans, it should understand our actions to produce the desirable response. As such, my work explores several directions in computationally representing and understanding human actions.

In this talk, I will focus on the problems of detecting actions and judging their quality. First, I will describe simple grammars for modeling long-scale temporal structure in human actions. Real-world videos are typically composed of multiple action instances, where each instance is itself composed of sub-actions with variable durations and orderings. Our grammar models capture such hierarchical structure while admitting efficient, linear-time parsing algorithms for action detection. The second part of the talk will describe our algorithms for going beyond detecting actions to actually judging how well they are performed. Our learning-based framework can both judge and provide feedback to performers to improve the quality of their actions.

Biography



Hamed Pirsiavash is an assistant professor at the University of Maryland Baltimore County (UMBC) since August 2015. Prior to that, he was a postdoctoral research associate at MIT working in Artificial Intelligence Laboratory and with Prof. Antonio Torralba. He obtained his PhD at the University of California Irvine under the supervision of Prof. Deva Ramanan. Hamed research is in the intersection of computer vision and machine learning, more specifically in understanding human actions.