The Johns Hopkins University  
Whiting School of Engineering  
Department of Electrical and Computer Engineering

Activity Recognition Using Interpretable Representations and Structured Dynamical Models

PhD Proposal Seminar by 
Lingling Tao  
Graduate Research Assistant (Dr. René Vidal/BME)  
Electrical and Computer Engineering

Abstract:
The human visual system is exquisitely sensitive to an enormous range of human movements. With no apparent effort, we can distinguish between simple motions (left leg up vs. right hand down), and actions (taking gun out vs. shaking hands). In sharp contrast, the capabilities of state-of-the-art action recognition technologies greatly lag those of humans. In a typical action recognition task, one needs to infer the activity from the image, video, or 3D data. However, most computer vision approaches to action recognition are based on the analysis of 2D image appearance and 2D image motion from images or videos. These systems have achieved only partial success in recognizing a handful of simple actions, such as walking vs. running. Approaches that rely on 3D information can extract a skeleton from each frame of the video sequence. However, these skeletons are often extracted in a frame by frame basis or by using an avatar to represent the 3D body kinematics, thereby neglecting the 3D body dynamics and the temporal relationship among different actions.

In our view, the performance of existing activity recognition approaches seems to be reaching a plateau, which is still well below human performance. We believe that this is because of a number of fundamental weaknesses of the “features + classifiers” approach, as well as the lack of feature representations that can capture discriminative human body movement.

This research proposal talk discusses how to develop highly interpretable, structured models of human motion that capture both 2D and 3D information about the dynamics of human movements, as well as spatial and temporal relationships among moving body parts, and actions. A novel feature representation named Moving Poselets is proposed for action recognition based on body-part movement. I will also talk about modeling the temporal structures of action series, combining the temporal model with Moving Poselets feature, and future directions on extending Moving Poselets to video domain.

Thursday, April 30, 2015 
3p.m. 
Latrobe 120

FOR DISABILITY INFORMATION CONTACT: Janel Johnson (410) 516-7031 janel.johnson@jhu.edu