Abstract:

In far-field visual surveillance, one of the key tasks is to monitor activities in the scene. Through clustering motion patterns of objects people can understand typical activities in a scene, detect abnormal activities and activities of interest, and learn the models of semantically meaningful scene structures.

Many existing activity analysis approaches in visual surveillance are ad hoc, relying on predefined rules or simple probabilistic models, which prohibits them from modeling complicated activities. In this talk, I will present a novel unsupervised framework to learn motion patterns from complicated and large scale data sets using nonparametric hierarchical Bayesian models. It jointly models simple and complicated activities/interactions at different hierarchical levels. When modeling complicated activities, hierarchical Bayesian models structure dependency among a large number of variables to avoid the overfitting problem. Various constraints and knowledge can be nicely added into a Bayesian framework as priors. So our framework is flexible to model activities in different applications and under different camera settings. When the number of clusters is not well defined in advance, our nonparametric approaches can learn it driven by data. Under this framework, we learn the models of activities and interactions in crowded scenes, where it is difficult to track objects because of frequent occlusions. We analyze activities observed in multiple camera views without tracking objects across camera views. The topology of multiple camera views is assumed to be unknown and arbitrary. The models of activities can be dynamically updated over time.

Our framework has also been successfully applied to a medical imaging application where some issues similar to learning motions patterns arise. Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) can visualize and quantify the organization of white matter in the brain. People connect local diffusion measurements to global fibers and cluster fibers into anatomically meaningful bundles. It is called tractography segmentation. Results on multiple large scale surveillance and DT-MRI data sets will be presented.

About the Speaker:

Xiaogang Wang received the BS degree in electrical engineering and information science from the University of Science and Technology of China in 2001, and the M. Phil. degree in information engineering from the Chinese University of Hong Kong in 2003. He is currently a PhD student in the Computer Science and Artificial Intelligence Laboratory at the Massachusetts Institute of Technology. His research interests include computer vision, machine learning and medical imaging, especially in the areas of visual surveillance, face recognition, object recognition and tractography segmentation.