Airborne Video Surveillance and Camera Networks

Dr. Mubarak Shah
Center for Research in Computer Vision
University of Central Florida
Orlando, FL 32816

Video surveillance and monitoring is one of most active areas of research in Computer Vision. The main steps in a video surveillance system include: detection and categorization of objects of interest in video (e.g. people, vehicles), tracking of those objects from frame to frame, and recognition of their activities, behavior and patterns. In this talk, first I will present our method for tracking thousands of objects in high resolution, low frame-rate, and multiple camera aerial videos. The proposed algorithm avoids the pitfalls of global minimization of data association costs and instead maintains multiple local associations for each track. In contrast with 1-1 correspondence constraints of bipartite graph matching and multiple hypotheses tracking algorithms, the proposed method allows representation of object state in terms of *many to many* data associations *per track*. Next, I will present recently developed method for Persistent Multi-Frame Multi-Object Detection in Wide Area Aerial Videos employing Fully Convolutional Deep Neural Networks. The network is designed to accept multiple video frames at a time as the input and yields detection results for all objects in the temporally center frame. This multi-frame approach yield far better results than its single frame counterpart. Next, I will introduce tracking in network of multiple non-overlapping cameras and present a unified three-layer hierarchical approach for solving tracking problems employing a constrained dominant sets clustering (CDSC) technique, a parametrized version of standard quadratic optimization. Finally, I will present our approach for inferring motion of objects that are invisible to all cameras in a multiple camera setup. As opposed to methods for learning relationships between disjoint cameras, we take the next step to actually infer the exact spatiotemporal behavior of objects while they are invisible. Given object trajectories within disjoint cameras' FOVs (field-of-view), we introduce constraints on the behavior of objects as they travel through the unobservable areas that lie in between. These constraints include vehicle following (the trajectories of vehicles adjacent to each other at entry and exit are time-shifted relative to each other), collision avoidance (no two trajectories pass through the same location at the same time) and temporal smoothness (restricts the allowable movements of vehicles based on physical limits).