AN ADAPTIVE AND INTEGRATED MULTIMODAL SENSING AND PROCESSING FRAMEWORK FOR LONG RANGE MOVING OBJECT DETECTION AND CLASSIFICATION

Tao Wang, CUNY Graduate Center
Senior Software Engineer, Video Understanding and Exploitation
BAE Systems, Technology Solutions, All-Source Exploitation

Monday November 05, 2012, NAC 7/225, 12:30 pm – 1:30 pm

Abstract:

In applications such as surveillance, inspection and traffic monitoring, long-range detection and classification of targets (vehicles, humans, etc) is a highly desired feature for a sensing system. A single modality will no longer provide the required performance due to the challenges in detection and classification with low resolutions, noisy sensor signals, and various environmental factors due to large sensing distances. Multimodal sensing and processing, on the other hand, can provide complementary information from heterogeneous sensor modalities, such as audio, visual and range sensors. However, there is a lack of effective sensing mechanisms and systematic approaches for sensing and processing using multimodalities. In this thesis, we described a systematical framework for Adaptive and Integrated Multimodal Sensing and Processing (thereafter, the AIM-SP framework) that integrates novel multimodal long-range sensors, adaptive feature selection and learning-based object detection and classification for achieving the goal of adaptive and integrated multimodal sensing and processing. Based on the AIM-SP framework, we have made three unique contributions. First, we have designed a novel multimodal sensor system called Vision-Aided Automated Vibrometry (VAAV), which is capable of automatically obtaining visual, range and acoustic signatures for moving object detection at a large distance. Second, multimodal data, acquired from multiple sensing sources, are integrated and represented in a Multimodal Temporal Panorama (MTP) for easy alignment and fast labeling. Accuracy of target detection can be improved using multimodalities. Further, a visual reconstruction method is developed to remove occlusions, motion blurs and perspective distortions of moving vehicles. With various types of features extracted on aligned multimodal samples, we made our third contribution on feature modality selection using two approaches. The first approach uses multi-branch sequential-based feature searching (MBSF) and the second one uses boosting-based feature learning (BBFL).

Bio-sketch

Tao Wang is currently a Senior Software Engineer at BAE Systems, working on Video Understanding and Exploitation. He received his B.S. degree in Computer Science from Stony Brook University, New York, in 2002, and the M.Eng. degree in Civil Engineering from Cornell University, New York, in 2004. Tao is now a PhD candidate at the Graduate Center of City University of New York (expected graduation: Fall 2012). Since 2006, he has been a research assistant in the City College Visual Computing Laboratory, working on multimodal sensor design, data integration, and video surveillance.