KEYNOTE TALK

Monday, December 12, 2016

8:30 AM – 9:30 AM / (Ballrooms 313 & 316)

ISVC 2016: 12th International Symposium on Visual Computing

Dynamic Scene Understanding for Autonomous Driving

James Rehg

Georgia Institute of Technology

Abstract

Autonomous navigation in complex urban environments requires the ability to recover a holistic model of the dynamic scene as it unfolds in time. Dynamic scene models include a 3D semantic map of the structures in the urban environment, such as buildings, roadway, and sidewalks, which define the substrate for the dynamic behaviors of vehicles and pedestrians. This static model is then augmented with a complete description of the dynamic scene elements via target detection and multi-target tracking. I will describe five recent advances which support the construction of dynamic scene models. Our starting point is a method for recovering a 3D volumetric semantic map of a static scene from monocular video. Given a sparse point cloud from structure-from-motion, we recover a dense 3D volumetric map through a global CRF optimization and the use of structural priors. The recovery of scene semantics is enabled by a novel and fast method for semantic video segmentation which uses feature optimization to enforce temporal continuity. We then describe a deep model for object detection in video frames which can serve as the starting point for multitarget tracking. Our detection architecture was the best-performing single model detector in the 2016 Microsoft COCO Challenge. We present a multiple-hypothesis tracking approach to multi-target tracking which incrementally processes video frames and uses efficient on-line appearance classification to differentiate object instances. We close by showing recent experimental results from our AutoRally project, a 1:5 scale open source vehicle platform which supports research in high-speed autonomous driving. This is joint work with Drs. Evangelos Theodorou and Panos Tsiotras and Ph.D. students Paul Drews, Brian Goldfain, Chanho Kim, Abhijit Kundu, Yin Li, and Grady Williams.

Speaker Bio-Sketch: James M. Rehg (pronounced "ray") is a Professor in the School of Interactiv Computing at the Georgia Institute of Technology, where he is Director of the Center for Behavioral Imaging and co-Director of the Computational Perception Lab (CPL). He received his Ph.D. from CMU in 1995 and worked at the Cambridge Research Lab of DEC (and then Compaq) from 1995-2001, where he managed the computer vision research group. He received an NSF CAREER award in 2001 and a Raytheon Faculty Fellowship from Georgia Tech in 2005. He and his students have received best student paper awards at ICML 2005, BMVC 2010, Mobihealth 2014, and Face and Gesture 2015, and a 2013 Method of the Year Award from the journal Nature Methods. Dr. Rehg serves on the Editorial Board of the Intl. J. of Computer Vision, and he served as the Program co-Chair for ACCV 2012 and General co-Chair for CVPR 2009, and is serving as Program co-Chair for CVPR 2017. He has authored more than 100 peer-reviewed scientific papers and holds 25 issued US patents. His research interests include computer vision, machine learning, robot perception and mobile health. Dr. Rehg was the lead PI on an NSF Expedition to develop the science and technology of Behavioral Imaging, the measurement and analysis of social and communicative behavior using multi-modal sensing, with applications to developmental disorders such as autism. He is currently the Deputy Director of the NIH Center of Excellence on Mobile Sensor Data-to-Knowledge (MD2K), which is developing novel on-body sensing and predictive analytics for improving health outcomes. See www.cbs.gatech.edu and md2k.org for details.