



**UNIVERSITY OF CENTRAL FLORIDA**  
CENTER FOR RESEARCH IN COMPUTER VISION

**Dr. Ajay Divakaran**  
**SRI International**

---

**“Comprehensive Human State Modeling and Its Applications”**  
**Thursday, April 2, 2015 · 11:00AM · HEC 450**



**ABSTRACT**

We present a suite of multimodal techniques for assessment of human behavior with cameras and microphones. These techniques drive the sensing module of an interactive simulation trainer in which the trainee has lifelike interaction with a virtual character so as to learn social interaction. We recognize facial expressions, gaze behaviors, gestures, postures, speech and paralinguistics in real-time and transmit the results to the simulation environment which reacts to the trainee's behavior in a manner that serves the overall pedagogical purpose. We will describe the techniques developed and results, comparable to or better than the state of the art, obtained for each of the behavioral cues, as well as identify avenues for further research. Behavior sensing in social interactions poses a few key challenges for each of the cues including the large number of possible behaviors, the high variability in execution of the same behavior within and across individuals and real-time execution. Furthermore, we have the challenge of appropriate fusion of the multimodal cues so as to arrive at a comprehensive assessment of the behavior at multiple time scales. We will also discuss our approach to social interaction modeling using our sensing capability to monitor and model dyadic interactions. We will present a video of the demonstration of the end to end simulation trainer.

Please see the following links for related demonstrations:

► [MIBADemo January2015 FinalDraft - YouTube](#); [SRI Social Interaction Modeling](#); [SRI Master Trainer Demo - YouTube](#)

**BIOGRAPHY**

Ajay Divakaran, Ph.D., is a Program Director and leads the Vision and Multi-Sensor group in SRI International's Vision and Learning Laboratory. Divakaran is currently the principal investigator for a number of SRI research projects. His work includes multimodal modeling and analysis of affective, cognitive, and physiological aspects of human behavior, interactive virtual reality-based training, tracking of individuals in dense crowds and multi-camera tracking, technology for automatic food identification and volume estimation, and audio analysis for event detection in open-source video. He has developed several innovative technologies for multimodal systems in both commercial and government programs during the course of his career. Prior to joining SRI in 2008, Divakaran worked at Mitsubishi Electric Research Labs for 10 years, where he was the lead inventor of the world's first sports highlights playback-enabled DVR. He also oversaw a wide variety of product applications for machine learning. Divakaran was named a Fellow of the IEEE in 2011 for his contributions to multimedia content analysis. He developed techniques for recognition of agitated speech for his work on automatic sports highlights extraction from broadcast sports video. He established a sound experimental and theoretical framework for human perception of action in video sequences as lead-inventor of the MPEG-7 video standard motion activity descriptor. He serves on Technical Program Committees of key multimedia conferences, and served as an associate editor of IEEE Transactions on Multimedia from 2007 to 2010. He has authored two books and has more than 100 publications to his credit, as well as more than 40 issued patents. Divakaran received his M.S. and Ph.D. degrees in electrical engineering from Rensselaer Polytechnic Institute. His B.E. in electronics and communication engineering is from the University of Jodhpur in India.