UNIVERSITY OF CENTRAL FLORIDA

CENTER FOR RESEARCH IN COMPUTER VISION

FINAL ORAL EXAMINATION

OF

GUANG SHU

B.S., HUAZHONG UNIVERSITY OF SCI. & TECH., 2006
M.S., SHANGHAI JIAOTONG UNIVERSITY, 2009

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(COMPUTER ENGINEERING)

Tuesday, November 4, 2014, 10:00 A.M.
101 Harris Corporation Engineering Center

DISSERTATION COMMITTEE

Professor Mubarak Shah, *Chairman*
Professor Ladislau Boloni
Professor Jun Wang
Professor Mingjie Lin
Professor Kiminobu Sugaya

OUTLINE OF GRADUATE STUDIES

Major: **Computer Engineering**

Computer Vision	Tappen
Advanced Computer Vision	Shah
Computer Vision Systems	Shah
Design and Analysis of Algorithms	Wocjan

SELECTED PUBLICATIONS (H-index: 5; Total Citations: 147)

- Improving an Object Detector and Extracting Regions using Super-pixels, G. Shu, A. Dehghan and M. Shah, in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2013.
- Semi-supervised Learning of Feature Hierarchies for Object Detection in a Video, Y. Yang, G. Shu and M. Shah, in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2013.
- Part-based Multiple-Person Tracking with Partial Occlusion Handling, G. Shu, A. Dehghan, O. Oreifej, E. Hand and M. Shah, in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2012.
- A Two-Stage Reconstruction Approach for Seeing Through Water, O. Oreifej, G. Shu, T. Pace, and M. Shah, in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2011.
- Gait recognition based on dynamic region analysis, X. Yang, Y. Zhou, T. Zhang, G. Shu and J. Yang, in Signal Processing, 2008.

DISSERTATION

HUMAN DETECTION, TRACKING AND SEGMENTATION IN SURVEILLANCE VIDEO

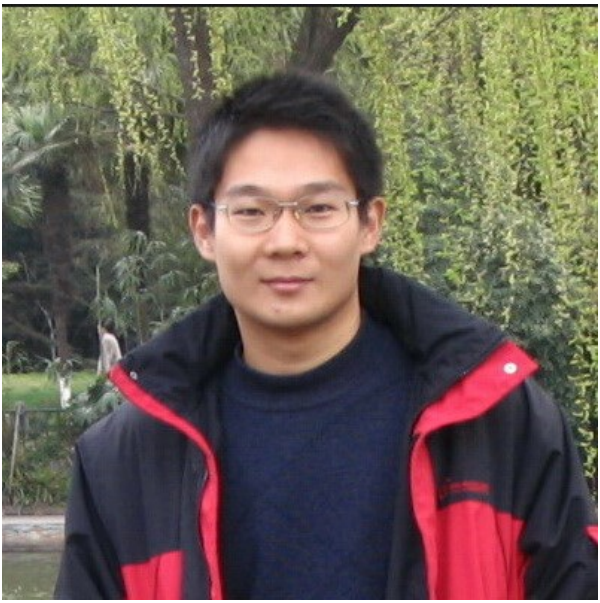
Video Surveillance is becoming increasingly important for public security and home safety. Computer vision algorithms for human detection, tracking, segmentation, action recognition etc play key roles in modern video surveillance systems.

This dissertation deals with the means of improving significantly the current state of human detection, tracking and segmentation based on learning scene-specific information in a video. Firstly, we propose a novel method for human detection by iteratively refining the output of a generic detector. In our approach, a DPM human detector is employed to collect detection examples; a support vector machine (SVM) classifier is trained using superpixel-based Bag-of-Words (BoW) feature and is used to validate the detections. The BoW feature encodes useful color features of the scene, hence it can handle the problem of appearance variation.

Given robust human detection, we propose a robust multiple-human tracking framework using a deformable part-based model (DPM). Our approach learns part-based person-specific SVM classifiers which capture articulations of moving human bodies with dynamically changing backgrounds. With the part-based model, we handle occlusions in both detection and tracking stages. In the detection stage, we select the subset of parts which maximizes the probability of detection, which significantly improves detection performance in crowded scenes. In the tracking stage, we distribute the score of the learned-person classifier among its corresponding parts, which allows us to track partial occlusions, and prevent the performance of the classifiers from being degraded.

Next, in order to obtain precise boundaries of humans, we propose a novel method for multiple human segmentation in video by incorporating human detection and part-based detection potential into a multi-frame optimization framework. In the first stage, we separate superpixels corresponding to a human and background by using the part-based detection potentials and minimizing an energy function using Conditional Random Field (CRF). In the second stage, the spatio-temporal constraints of the video is leveraged to build tracklet-based Gaussian Mixture Models (GMM) , and the boundaries are smoothed by multi-frame graph optimization.

Finally, we have developed an efficient real time tracking system, NONA, for high -definition surveillance video. We implement the system using a multi-threaded architecture (Intel Threading Building Blocks (TBB)), which executes video ingestion, tracking and video output in parallel. We employ a Fast Fourier Transform based normalized cross correlation as the core tracking algorithm for efficiency. We also incorporate Adaptive Template Scaling to handle the scale change, and Local Frame Differencing to resolve challenging issues such as occlusion and cluttered backgrounds.



GUANG SHU

1983	Born in Liaocheng, China.
2006	B.S., Huazhong University of Sci. & Tech., China.
2009	M.S., Shanghai Jiaotong University., China.
2009	Research Intern, Sharp Corp., Shanghai, China.
2009-14	Ph.D., University of Central Florida, Orlando, FL.
2014	Research Intern, Sighthound Inc., Orlando, FL.
2015-	Research Scientist, Sighthound Inc., Orlando, FL.