



UNIVERSITY OF CENTRAL FLORIDA
CENTER FOR RESEARCH IN COMPUTER VISION

FINAL ORAL EXAMINATION

OF

AMIR ROSHAN ZAMIR
M.S., UNIVERSITY OF CENTRAL FLORIDA, 2013

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
(ELECTRICAL ENGINEERING)

Wednesday, March 12, 2014, 9:00 A.M.
101 Harris Corporation Engineering Center

DISSERTATION COMMITTEE

Professor Mubarak Shah, *Chairman*
Professor Rahul Sukthankar
Professor Sumit K. Jha
Professor Mingjie Lin
Professor Sasan Fathpour

OUTLINE OF GRADUATE STUDIES

Major: Electrical Engineering

<i>Courses:</i> Computer Vision	Tappen
Advanced Computer Vision	Shah
Computer Vision Systems	Shah
3D Computer Vision	Foroosh
Machine Learning	Wiegand
Image Processing	Foroosh

SELECTED PUBLICATIONS

1. **Image Geo-localization Based on Multiple Nearest Neighbor Feature Matching Using Generalized Graphs**, A. Roshan Zamir and M. Shah, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2014.
2. **Robust Refinement of GPS-Tags Using Random Walks with an Adaptive Damping**, A. Roshan Zamir, S. Ardeshir and M. Shah, in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
3. **Video Classification Based on Generalized Maximum Co-occurrence Cliques**, S. Modiri, A. Roshan Zamir, and M. Shah, in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
4. **Visual Business Recognition - A Multimodal Approach**, A. Roshan Zamir, A. Dehghan and M. Shah, in *ACM International Conference on Multimedia (ACM MM)*, 2013.
5. **GMCP-Tracker: Global Multi-object Tracking Using Generalized Minimum Clique Graphs**, A. Roshan Zamir, A. Dehghan and M. Shah, in *European Conference on Computer Vision (ECCV)*, 2012.
6. **City Scale Geo-spatial Trajectory Estimation of a Moving Camera**, Gonzalo Vaca-Castano, A. Roshan Zamir and M. Shah, in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
7. **UCF101: A Dataset of 101 Human Actions Classes From Videos in The Wild**, K. Soomro, A. Roshan Zamir and M. Shah, *CRCV-TR-12-01*, 2012.
8. **Automatic Detection and Tracking of Pedestrians in Videos with Various Crowd Densities**, A. Dehghan, H. Idrees, A. Roshan Zamir, and M. Shah, in *International Conference on Pedestrian and Evacuation Dynamics (PED)*, 2012.
9. **Identification of Commercial Entities in Street View Imagery**, A. Roshan Zamir, A. Darino, R. Patrick and M. Shah, in *International Conference on Machine Learning and Applications (ICMLA)*, 2011.
10. **Accurate Image Localization Based on Google Maps Street View**, A. Roshan Zamir and M. Shah, in *European Conference on Computer Vision (ECCV)*, 2010, (ECCV Travel Award).

DISSERTATION

VISUAL GEO-LOCALIZATION AND LOCATION-AWARE IMAGE UNDERSTANDING

Geo-localization is the problem of discovering the location where an image or video was captured. Recently, large scale image geo-localization methods that employ techniques similar to image matching have attracted much interest. In these methods, given a reference dataset consisting of geo-tagged images, the problem is to estimate the location of a query image by finding its matching reference images.

In this thesis, we present a new framework for geo-locating an image utilizing a novel multiple nearest neighbor feature matching method using Generalized Minimum Clique Graphs (GMCP). First, we extract local features (e.g., SIFT) from the query image and retrieve a number of nearest neighbors for each query feature from the reference data set. Next, we apply our GMCP-based feature matching to select a single nearest neighbor for each query feature such that all matches are globally consistent. Our approach to feature matching is based on the proposition that the first nearest neighbors are not necessarily the best choices for finding correspondences in image matching. Therefore, the proposed method considers multiple reference nearest neighbors as potential matches and selects the correct ones by enforcing the consistency among their global features (e.g., GIST) using GMCP. Our evaluations using a new data set of 102k Street View images shows the proposed method outperforms the state-of-the-art by 10 percent.

Geo-localization of images can be extended to geo-localization of a video. We have developed a novel method for estimating the geospatial trajectory of a moving camera with unknown intrinsic parameters in a city-scale. The proposed method is based on a three step process: 1) individual geo-localization of video frames using Street View images to obtain the likelihood of the location (latitude and longitude) given the current observation, 2) Bayesian tracking to estimate the frame location and video's temporal evolution using previous state probabilities and current likelihood, and 3) applying a novel Minimum Spanning Trees based trajectory reconstruction to eliminate trajectory loops or noisy estimations.

So far, we have assumed reliable geo-tags for reference imagery are available through crowdsourcing. However, crowdsourced images are well known to suffer from the acute shortcoming of having inaccurate geo-tags. We have developed the first method for refinement of GPS-tags which automatically discovers the subset of corrupted geo-tags and refines them. We employ Random Walks to discover the accurate subset of location estimations and robustify Random Walks with a novel adaptive damping factor that conforms to the level of noise in the input.

In location-aware image understanding, we are interested in improving the image analysis by putting it in the right geo-spatial context. This approach is of particular importance as the majority of cameras and mobile devices are now being equipped with GPS chips. Therefore, developing techniques which can leverage the geo-tags of images for improving the performance of traditional computer vision tasks is of particular interest. We have developed a location-aware multimodal approach which incorporates business directories, textual information, and web images to identify businesses in a geo-tagged query image.



AMIR ROSHAN ZAMIR

- 1986 Born in Tehran, Iran.
2008 B.S., Tehran Polytechnic, Iran.
2007 Undergrad Research Assistant, PRL Lab, Tehran, Iran
2009-14 Ph.D., University of Central Florida, Orlando, FL.
2013 Program Chair, ICCV THUMOS Challenge Workshop, Sydney, Australia.
2014 Research Associate, Center for Research in Computer Vision at UCF, Orlando, FL.

SELECTED AWARDS

- 2010 ECCV Travel Award
2013 UCF Research Excellence Award
2013 National Geospatial-Intelligence Agency (NGA) Best Research Poster Award.

INVITED TALKS

- 2013 *Large Scale Image and Video Geo-localization Using Street View Imagery*, Department of Electrical and Computer Engineering, **University of Maryland, College Park**.
2013 *Global Multi-object Tracking Using Generalized Graphs*, College of Computing, **Georgia Institute of Technology**.
2013 *Large Scale Image and Video Geo-localization and Location-aware Image Understanding*, Department of Computer Science, **Cornell University**.