Geo-localization is the problem of discovering the location where an image or video was captured. Recently, large scale image geo-localization methods which 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 geo-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 uncontaminated 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.

Educational Career:
Bachelor's of Electrical Engineering, BS, 2008, Semnan University
Master's of Computer Engineering, MS, 2013, University of Central Florida

Committee in Charge:
Mubarak Shah, Chair, Electrical Engineering & Computer Science