With the increasing pervasiveness of digital cameras, the Internet, and social networking, there is a growing need to catalog and analyze large collections of photos and videos. In this dissertation, we explore unconstrained still-image and video-based face recognition in real-world scenarios, e.g. social photo sharing and movie trailers, where people of interest are recognized and all others are ignored. In such a scenario, we must obtain high precision in recognizing the known identities, while accurately rejecting those of no interest.

Recent advancements in face recognition research has seen Sparse Representation-based Classification (SRC) advance to the forefront of competing methods. However, its drawbacks, slow speed and sensitivity to variations in pose, illumination, and occlusion, have hindered its wide-spread applicability. The contributions of this dissertation are three-fold:

1. For still-image data, we propose a novel Linearly Approximated Sparse Representation-based Classification (LASRC) algorithm that uses linear regression to perform sample selection for $l_1$-minimization, thus harnessing the speed of least-squares and the robustness of SRC. On our large dataset collected from Facebook, LASRC performs equally to standard SRC with a speedup of 100-250x.

2. For video, applying the popular $l_1$-minimization for face recognition on a frame-by-frame basis is prohibitively expensive computationally, so we propose a new algorithm Mean Sequence SRC (MSSRC) that performs video face recognition using a joint optimization leveraging all of the available video data and employing the knowledge that the face track frames belong to the same individual. Employing MSSRC results in a speedup of 5x on average over SRC on a frame-by-frame basis.

3. Finally, we make the observation that MSSRC sometimes assigns inconsistent identities to the same individual in a scene that could be corrected based on their visual similarity. Therefore, we construct a probabilistic affinity graph combining appearance and co-occurrence similarities to model the relationship between face tracks in a video. Using this relationship graph, we employ random walk analysis to propagate strong class predictions among similar face tracks, while dampening weak predictions. Our method results in a performance gain of 15.8% in average precision over using MSSRC alone.
The public is welcome to attend.