Announcing the Final Examination of Arslan Basharat for the degree of Doctor of Philosophy

Time & Location: November 4, 2009 at 11:00 AM in HEC 101
Title: Modeling Scenes and Human Activities in Videos

In this dissertation, we address the problem of understanding human activities in videos by developing a two-pronged approach: coarse level modeling of scene activities and fine level modeling of individual activities. At the coarse level, where the resolution of the video is low, we rely on person tracks. At the fine level, richer features are available to identify different parts of the human body, therefore we rely on the body joint tracks. There are three main goals of this dissertation: (1) identify unusual activities at the coarse level, (2) recognize different activities at the fine level, and (3) predict the behavior for synthesizing and tracking activities at the fine level.

The first goal is addressed by modeling activities at the coarse level through two novel and complementing approaches. The first approach learns the behavior of individuals by capturing the patterns of motion and size of objects in a compact model. Probability density function (pdf) at each pixel is modeled as a multivariate Gaussian Mixture Model (GMM), which is learnt using unsupervised expectation maximization (EM). In contrast, the second approach learns the interaction of object pairs concurrently present in the scene. This can be useful in detecting more complex activities than those modeled by the first approach. We use a 14-dimensional Kernel Density Estimation (KDE) that captures motion and size of concurrently tracked objects. The proposed models have been successfully used to automatically detect activities like unusual person drop-off and pickup, jaywalking, etc.

The second and third goals of modeling human activities at the fine level are addressed by employing concepts from theory of chaos and non-linear dynamical systems. We show that the proposed model is useful for recognition and prediction of the underlying dynamics of human activities. We treat the trajectories of human body joints as the observed time series generated from an underlying dynamical system. The observed data is used to reconstruct a phase (or state) space of appropriate dimension by employing the delay-embedding technique. This transformation is performed without assuming an exact model of the underlying dynamics and provides a characteristic representation that will prove to be vital for recognition and prediction tasks. For recognition, properties of phase space are captured in terms of dynamical and metric invariants, which include the Lyapunov exponent, correlation integral, and correlation dimension. A composite feature vector containing these invariants represents the action and will be used for classification. For prediction, kernel regression is used in the phase space to compute predictions with a specified initial condition. This approach has the advantage of modeling dynamics without making any assumptions about the exact form (polynomial, radial basis, etc.) of the mapping function. We demonstrate the utility of these predictions for human activity synthesis and tracking.

Major: Computer Science

Educational Career:
Bachelor's of Computer Systems Engineering, BS, 2001, GIK Institute, Pakistan
Master's of Computer Science, MS, 2006, University of Central Florida

Committee in Charge:
Mubarak A. Shah, Chair, Electrical Engineering & Computer Science
Niels da Vitoria Lobo, Electrical Engineering & Computer Science
Annie S. Wu, Electrical Engineering & Computer Science
Xin Li, Mathematics

Approved for distribution by Mubarak A. Shah, Committee Chair, on October 19, 2009.
The public is welcome to attend.