GENERALIZED PRINCIPAL COMPONENT ANALYSIS (GPCA):
AN ALGEBRAIC GEOMETRIC APPROACH
TO SUBSPACE CLUSTERING

ABSTRACT

Data segmentation is usually thought of as a “chicken-and-egg” problem. In order to estimate a mixture of models one needs first to segment the data, and in order to segment the data one needs to know the model parameters. Therefore, data segmentation is usually solved in two stages: (i) data clustering and (ii) model fitting, or else iteratively using, e.g., the Expectation Maximization (EM) algorithm.

This talk will show that for a wide class of segmentation problems (e.g., mixtures of subspaces, mixtures of fundamental matrices/trifocal tensors, mixtures of linear dynamical models) the “chicken-and-egg” dilemma can be tackled using algebraic geometric techniques. In fact, it is possible to eliminate the data segmentation step algebraically and then use all the data to recover all the models without previously segmenting the data as follows:

1. Fit a set of polynomials to all data points, without clustering the data.
2. Obtain the model parameters for each group from the derivatives of these polynomials.

Applications of GPCA to image/video segmentation, face clustering, 3-D motion segmentation, dynamic texture segmentation, and heart motion analysis will also be presented.