- Thursday, November 6, 2003
  Learning parts

  CSE Many collections of data exhibit a common underlying structure: they consist of a number of parts or factors, each with a range of possible states. For example, in a collection of facial images,every image contains eyes, a nose, and a mouth, each of which has a number of appearances. We propose a new method, Multiple Cause

  Vector Quantization, for the unsupervised learning of parts-based representations of data. Our technique automates the segmentation of the data dimensions into parts, while simultaneously learning a discrete model of the range of appearances of each part. We pose MCVQ as a probabilistic graphical model, and derive an efficient and biologically plausible variational-EM algorithm for learning and inference. I will present applications of this model to problems in image decomposition, collaborative filtering, and document modeling

  Richard Zemel
  U of Toronto