Statistical Machine Learning
Problem set 7

Alan Joseph Bekker

Eigenfaces for recognition

The goal of the assignment is to implement a nearest-neighbor face recognition algorithm based on Eigenfaces as discussed in class. We will be using a database consisting of 11 images each of 15 different people. The image resolution is 32x32, corresponding to the extrinsic dimension of 1024. The task is to correctly identify the person from a test image using a low-dimensional eigenface representation.

Directions:
Load the face database from the moodle site and partition it into training and test sets. We will be using 8 images per person for training and 3 images for testing. Find the mean face image. Perform PCA on the training faces and extract top K components. Use the MATLAB princomp function & compare it to your PCA code. Compute its K-dimensional projection of the test images onto the face space. For each test image, find the training image that is "closest" (in the sense of Euclidean distance) to the test image in the face space, and assign the label (person index) of the training image to the test image.

What to hand in:

The picture of the mean face and top 20 eigenfaces computed by PCA. A plot of the nearest-neighbor classification rate as a function of K. You can choose any sampling of values K from 1 to 1024, as long as it captures the trend of how classification performance changes as a function of K (i.e., we expect performance to be poor for extremely low K, but then to rise very rapidly and level off at some point).