Lydia LaSeur
October 8, 2014
ISC 4223

Lab #1 – Face Recognition using PCA

Principal component analysis (PCA) is a method used to reduce the dimensionality of high-dimensional data while preserving the overall variance of the data. Dimensionality reduction is extremely important for computation in order to conserve memory, minimize error, and achieve high performance. With almost any type of data, the most important characteristic is the variance. It numerically describes the overall characteristics of the data set and what makes one particular instance distinct from another. PCA looks at the covariance of a set of vectors, in the form of a covariance matrix. A basis for a lower dimensional subspace of the data set is formed by the "best" eigenvectors (ones with the largest magnitude) of the covariance matrix. Projecting the high-dimensional vector onto this lower-dimensional subspace, or eigenspace, will reduce the original dimensionality. Now these new, much smaller vectors can be used in place of the original vectors for a desired method.

PCA is especially useful in image processing since images are represented with very high dimensionality on computers, i.e. a matrix or vector where each element represents each pixel. Therefore, even relatively "small" images are represented by impractically large matrices that are very expensive and difficult to work with.

One particular application in image processing is face recognition. Given a training set of faces the basis of the eigenspace is found. The image of each training face can reconstructed as a linear combination of "eigenfaces" i.e. the best eigenvectors of the covariance matrix. The coefficients of this linear combination form a vector the represents the training face in a lower dimensional space. Using the same basis, an unknown face can be represented as a linear combination of the eigenfaces as well. By comparing the unknown face to the training faces, using a distance measure, the test face can be recognized as a specific face/person from the training set. In this assignment, the training set will consist 16 faces from unique subjects. In the training set, all subjects are facing the camera straight on and the lighting is consistent throughout. The testing set will contain 4 test images. All are subjects that were included in the training set but each image will have some characteristic that differs from the corresponding image in the training set, i.e. subject photographed from an angle, change in facial expression, addition or removal of glasses, change in lighting. For each test face, the closest match from the training set will identify the test face as that specific person.

Test #1:

Test face sqrtest1.pgm

Test face recognized as sqrs03.pgm

Test #2:

Test face sqrtest12.pgm

Test face recognized as sqrs05.pgm

Test #3:

Test face sqrtest4.pgm

Test face recognized as sqrs04.pgm

Test #4:

Test face sqrtest7.pgm

Test face recognized as sqrs07.pgm

The first test shown above illustrates that facial recognition using PCA is sensitive to changes in the perspective of the face. The training face for subject 1 shows him photographed head-on, but the test face views him from a slight angle. This change in perspective has a significant effect on the variance of the image data. Therefore, the test face is matched to the wrong training face since it has a similar orientation. If additional faces showing different perspectives were included in the training set, the program would have most likely performed better.

The second test fails as well. This is due to the change in shadows/lighting. The training face for subject 12 shows her in brighter light, which diminishes the shadows on her face. As opposed to the test face for subject 12 in which shadows appear under her eyes and smile lines. Again, the addition of the shadows will alter the variance of the face and PCA is sensitive to significant changes in variance between training and testing.

Of course with all these test faces, there is some change in the variance between the training faces. The last two tests shown above illustrate the types of changes that PCA is not as sensitive to. In test #3 the removal of the subjects glasses is not enough change for PCA to fail. Similarly in test #4 the addition of glasses does not cause PCA to fail, as well as the big change in facial expression. However, the orientation and lighting between the training and testing faces are virtually exactly the same for tests #3 and #4.

From this assignment it clear to me the importance of orientation and lighting of the faces when using PCA for recognition. This makes sense since changes in those characteristics will affect the variance over the whole image. However, features like glasses or facial expression only affect the variance in part of the image so PCA is not as sensitive to these alterations.