

Assignment PCA - IRM2016502

Problem:

To perform Face recognition using Principal Component Analysis (PCA).

Data:

I have used the 'ORL Database of Faces' from which I've chosen 15 different people each folder having 10 faces of a single person. Out of the 10 faces I chose 5 faces for training and remaining for testing purposes.

Procedure:

- As the first step I've loaded the data into a numpy array. We have a total of 15 classes and 5 images in each class for training, 5 images of each class for Testing.
- First I've found the mean of all train images so as to make the mean zero.

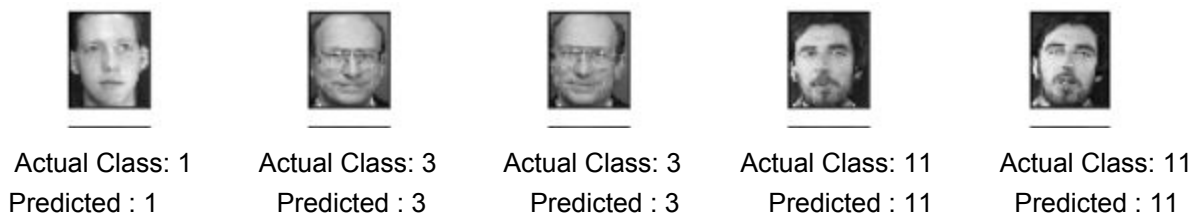
Mean Face



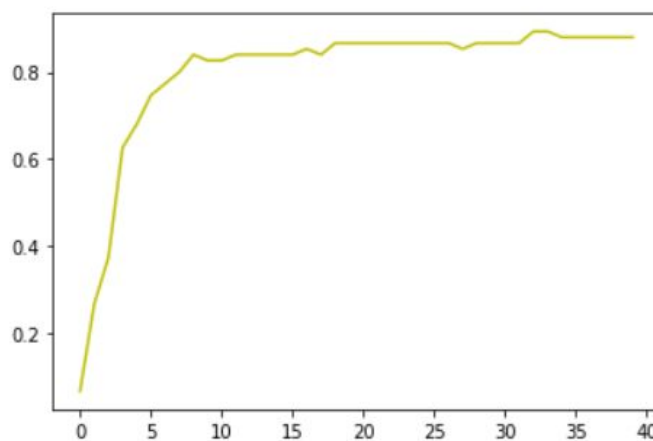
- Then after This mean face is subtracted from all the faces.
- Later I've found the covariance matrix and from which we could get its eigenvalues and eigenvectors.
- Now we need to choose a K to take top K eigenvectors corresponding to first K eigenvalues when they are sorted in descending order.
- Suppose K value is **k**. Then **k** eigenvectors are chosen according to their order of eigenvalues.

- Multiplied them with input to get eigen_faces.
- These are again multiplied to input faces to get signatures/ weights of each training image.
- During testing I've taken an image, normalised it, then multiplied it with eigen_faces to get it's weights.
- Then the image class with minimum weight difference is predicted as the class of test image.

The below is a sample of testing where actual class and predicted class along with the image are shown.



I've run a for loop from 1-40 to find out the best k value for the chosen dataset, it was found out to be 32. That is, considering 32 principal components is giving maximum accuracy.



The maximum accuracy of 0.8933333333333333 is observed at k = 32