Problem 2

Face modeling serves as one of the most fundamental problems in modern artificial intelligence and computer vision, and can be useful in various applications like face recognition, identification etc. However, face images are usually of high-dimensional (e.g., a small 100×100 gray-scaled face image has dimension 100 × 100 = 10, 000), therefore, find a suitable representation is utterly important. In this problem, we apply the linear model, principal component analysis (PCA), on face images to reduce the dimension and obtain eigenface representations.

Dataset: we use the dataset† which contains 177 face images. Each image contains 256 × 256 pixels and is gray-scaled (i.e., the value for each pixel is stored as unsigned integer between [0, 255], typically, 0 is taken to be black and 255 is taken to be white). You need to split the dataset to be train/test set, e.g., you could use the first 157 images for training, and the rest 20 faces for testing.

(1) **Write the PCA codes to compute K = 30 eigenfaces and visualize the top 10 eigen- faces.**

(2) **Use the learned K eigenfaces from (1) to reconstruct testing images, and record the reconstruction error. The reconstruction error can be defined as $|| \hat{Y} - Y||_2 / N$, where $\hat{Y}$ is the reconstructed face using the learned eigenfaces, Y is the testing faces and N is the total number of testing data. Please show 5 testing images and their reconstructed ones.**

(3) **Try different values of K , e.g., try K = 10, 30, 50, 100, 150..., and draw the curve to indicate the corresponding testing reconstruction errors. The x-axis of the curve can be different K values, and the y-axis can be testing reconstruction error defined in (2).**

Some useful hints:

(1) To construct the data matrix for PCA, you can reshape each image to a long vector. For example, the original 2D image of size [h, w] becomes 1D vector of size h × w. Each row of the data matrix represents one face image and data matrix has size [Ntrain,D] where Ntrain is the number of training samples and D = h ∗ w is the dimension of the reshaped image.

(2) To compute the PCA, you need to subtract the mean image from each training image. To get the mean image, just sum up all the training images and divide it by Ntrain.

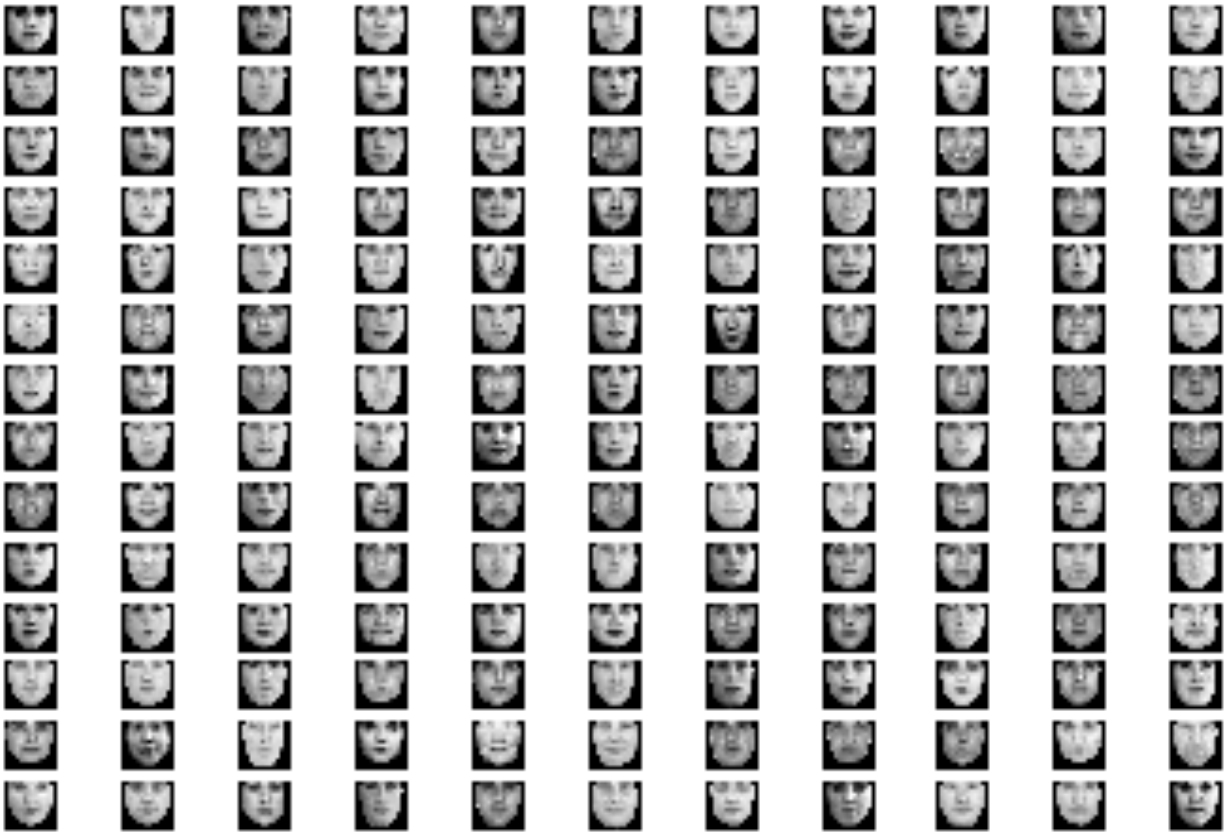
(3) You could use from PIL import Image, Image.open to read the images, and use mat-plotlib.pyplot to show the images. Feel free to use other functions to read and process the images as well.

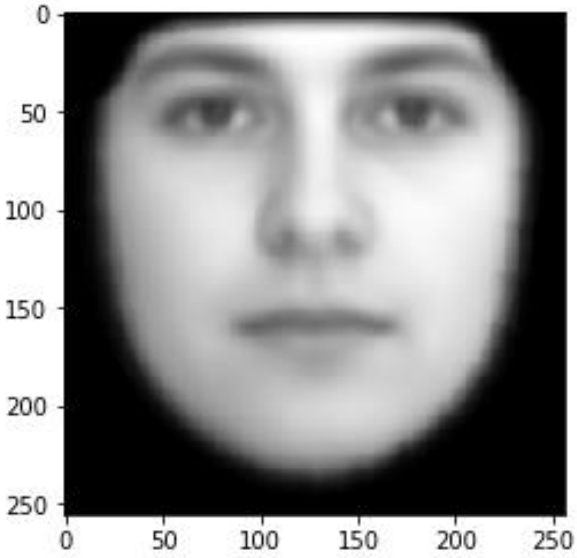
(4) For PCA, you could use linalg from numpy for eigen decomposition. Other eigenvalue decom- position/SVD functions can also be used. However, you can not directly call the pca functions in any languages for this problem

CS559-B HW2

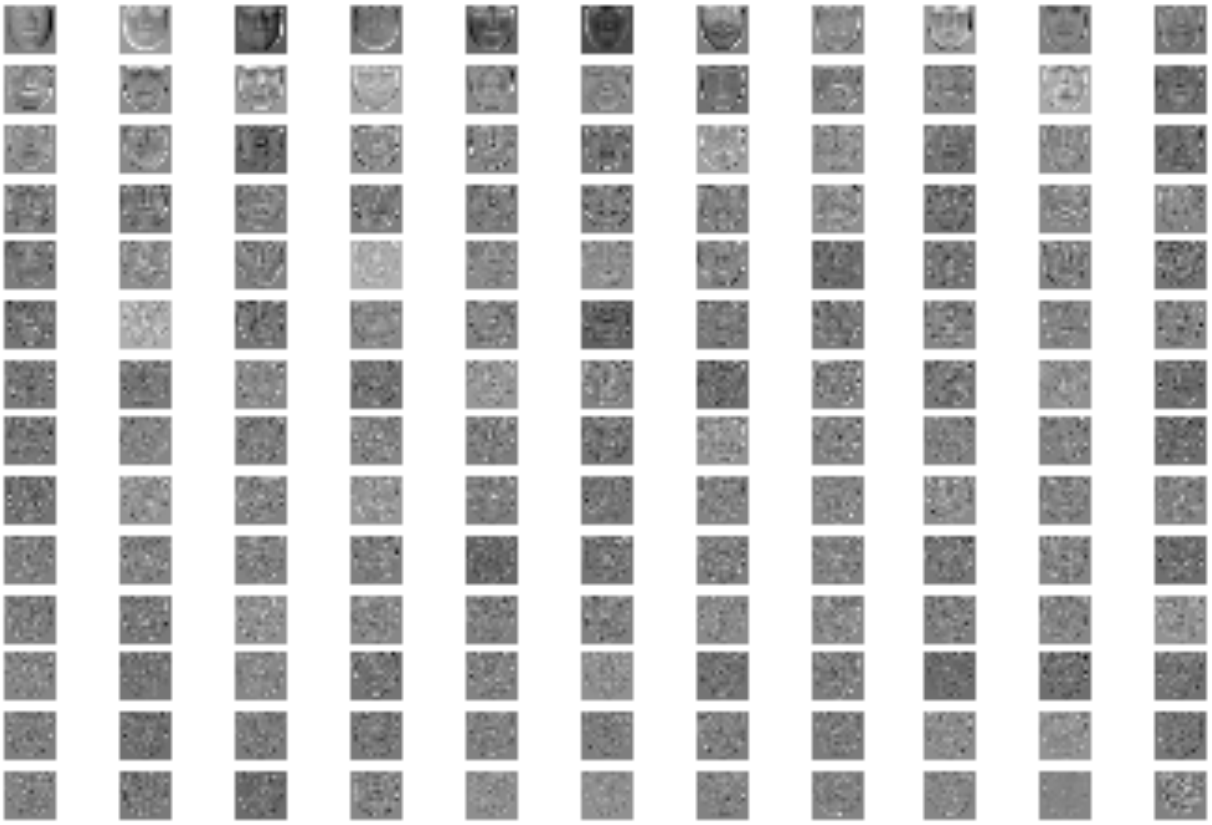
Output:

runfile('/Users/shreyavhadadi/Desktop/temp.py', wdir='/Users/shreyavhadadi/Desktop')

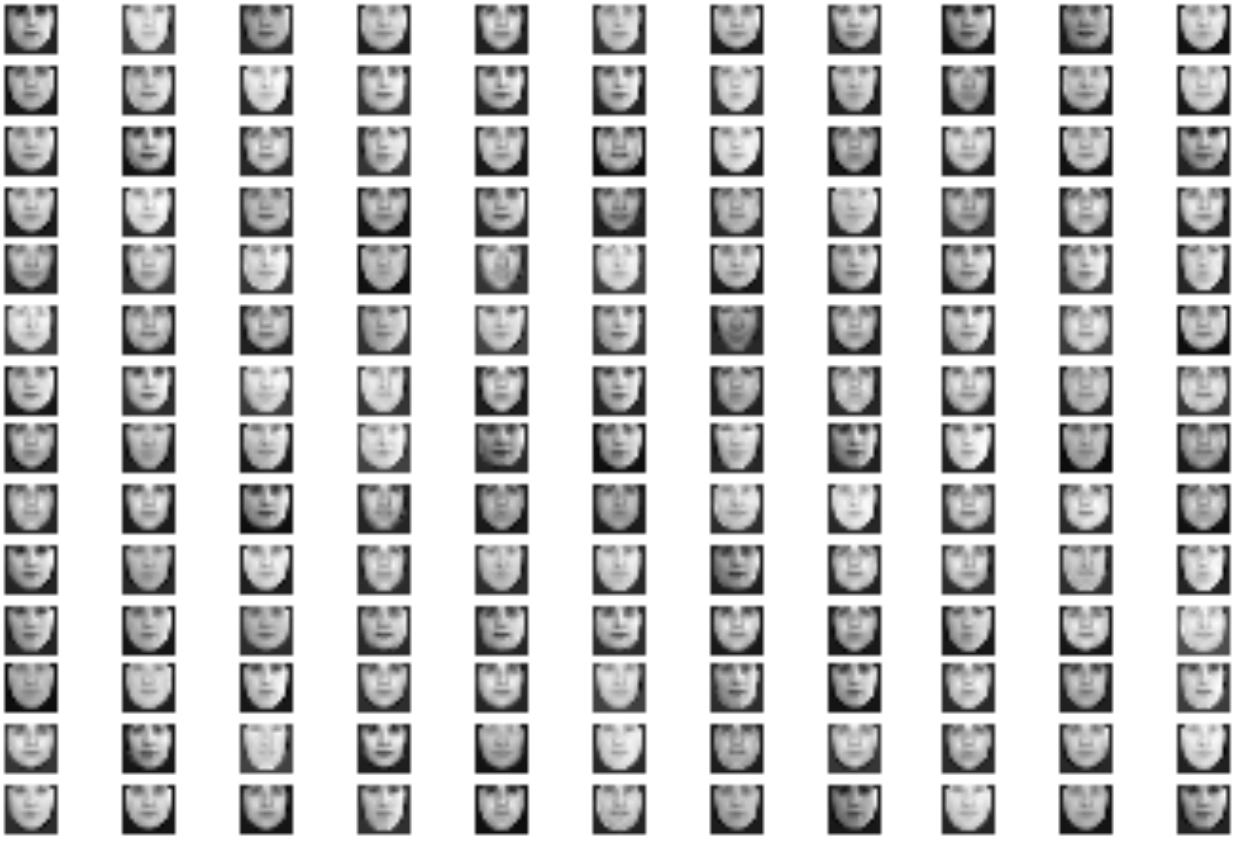




CS559-B HW2

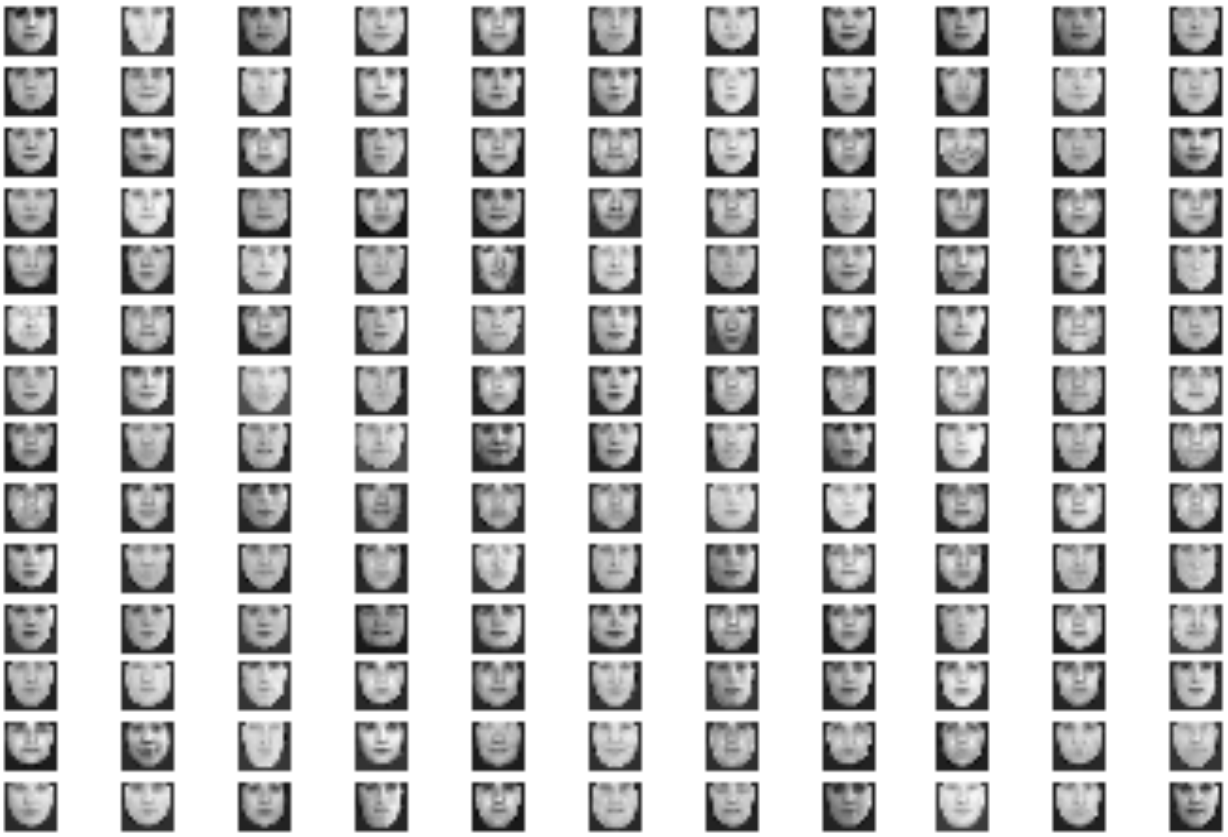


K=10



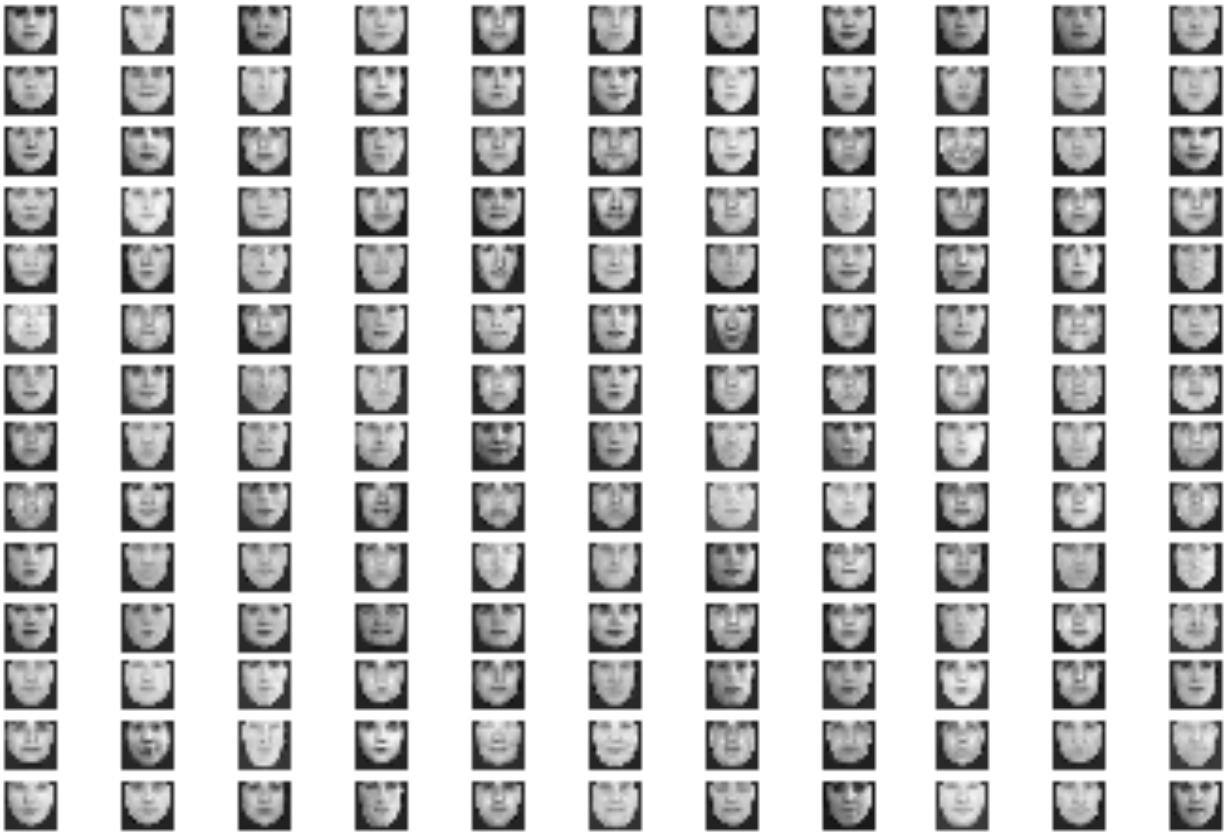
CS559-B HW2

K=30

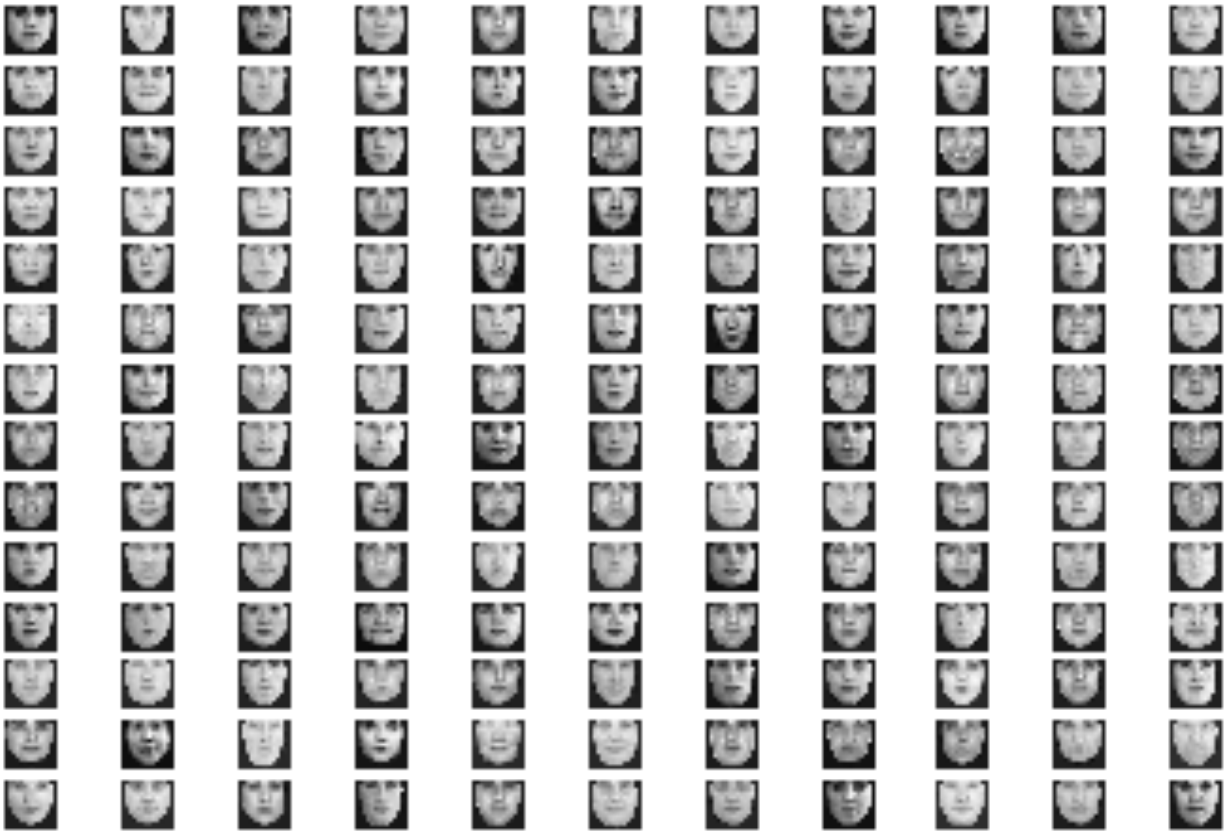


K=50

CS559-B HW2



K=100



K=10

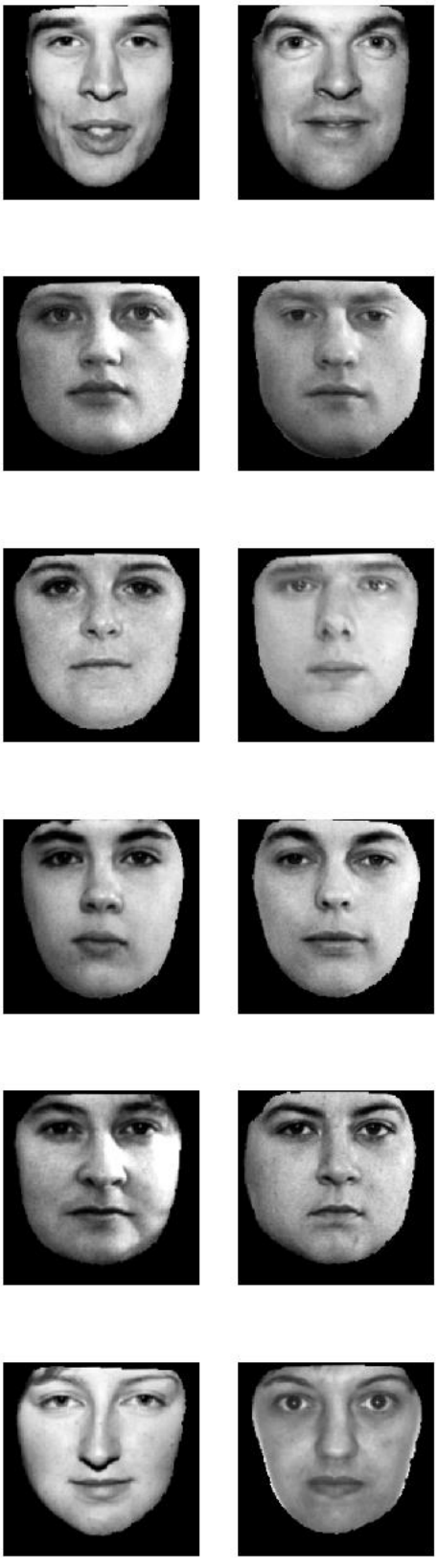
CS559-B HW2

/Users/shreyavhadadi/Desktop/temp.py:174: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).

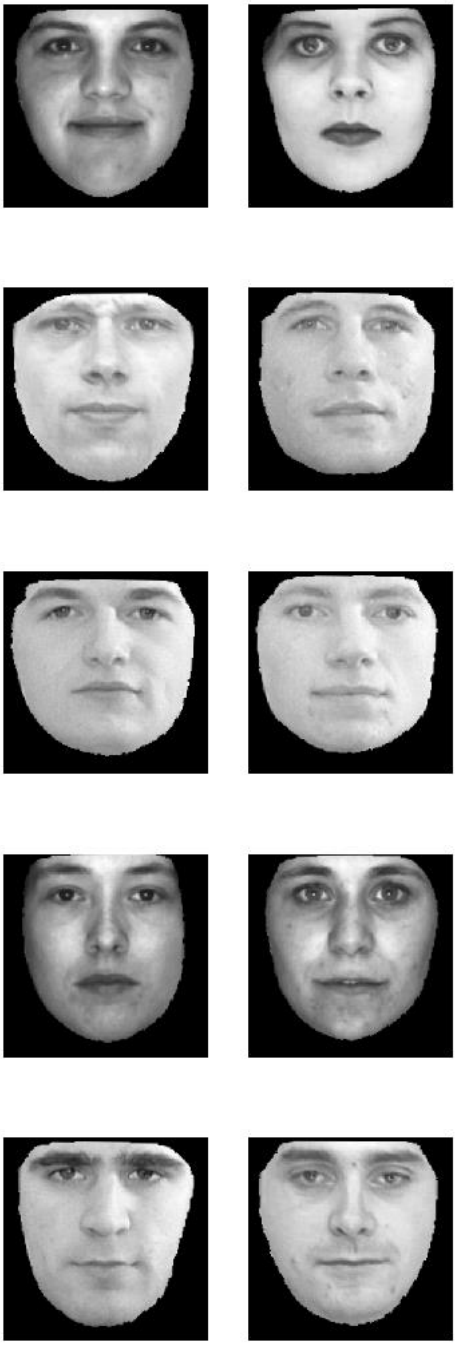
fig, axes = plt.subplots(1, 2)

CS559-B HW2

CS559-B HW2



CS559-B HW2

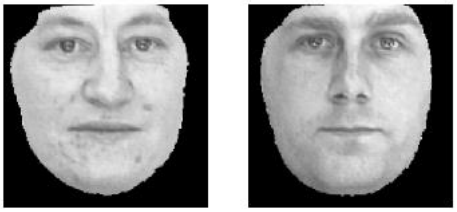


K=30

/Users/shreyavhadadi/Desktop/temp.py:192: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).

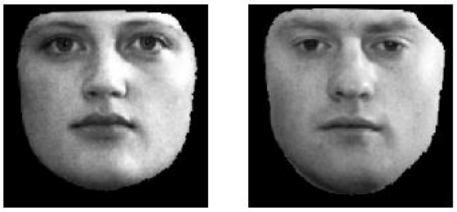
fig, axes = plt.subplots(1, 2)

CS559-B HW2

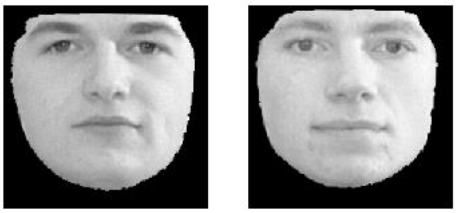


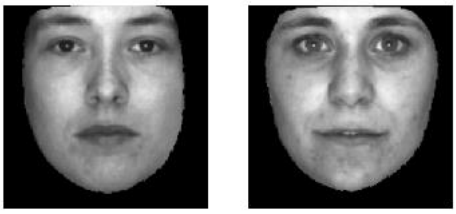
CS559-B HW2

CS559-B HW2



CS559-B HW2

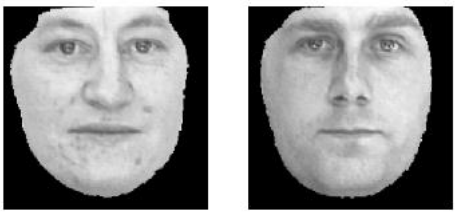


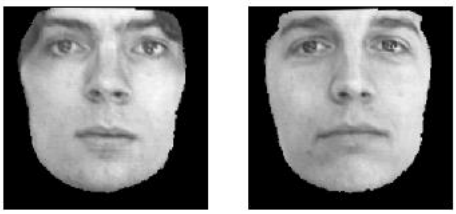


K=50

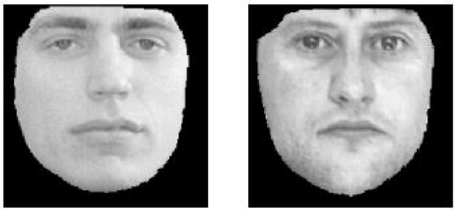
/Users/shreyavhadadi/Desktop/temp.py:210: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).

fig, axes = plt.subplots(1, 2)





CS559-B HW2



CS559-B HW2

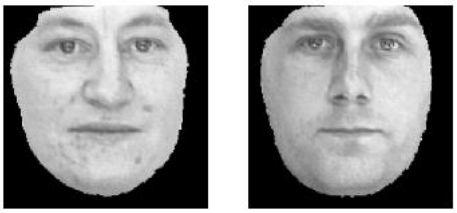
CS559-B HW2

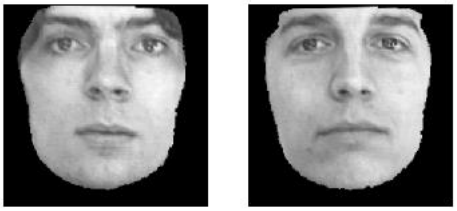
CS559-B HW2

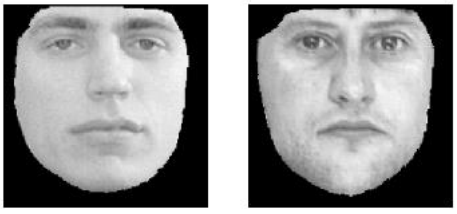
K=100

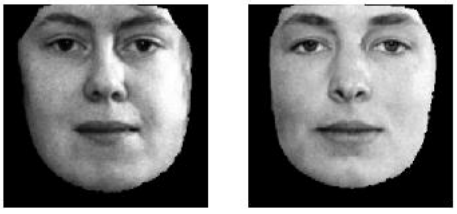
/Users/shreyavhadadi/Desktop/temp.py:228: RuntimeWarning: More than 20 figures have been opened. Figures created through the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly closed and may consume too much memory. (To control this warning, see the rcParam `figure.max_open_warning`).

fig, axes = plt.subplots(1, 2)





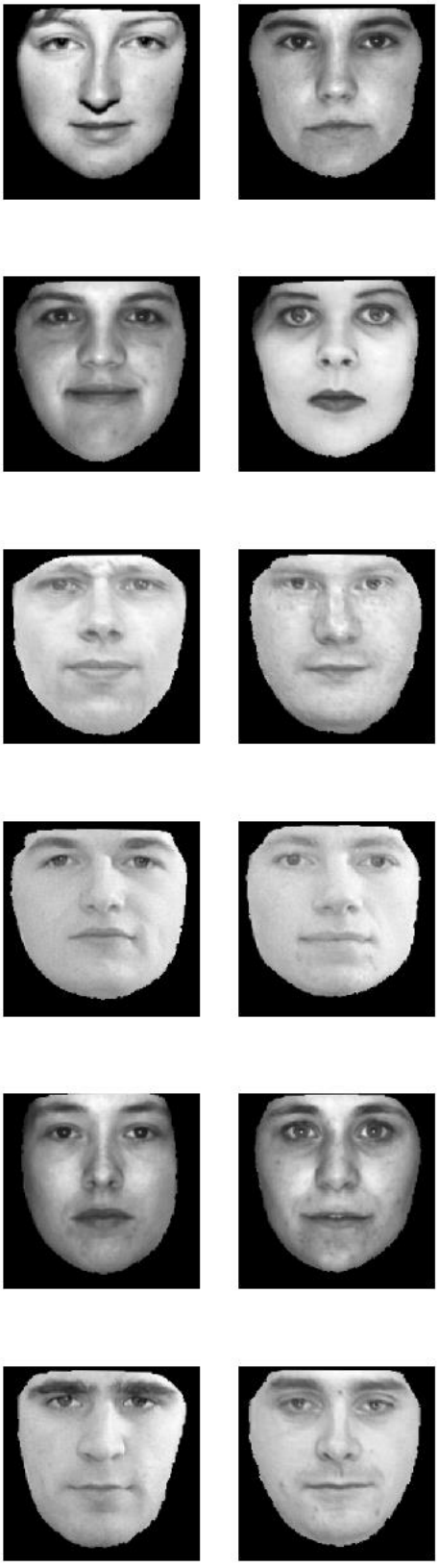




CS559-B HW2

CS559-B HW2

CS559-B HW2



CS559-B HW2