

CSC320: Assignment #3

Due on Saturday, March 21, 2015

Ying Qi He

March 21, 2015

Part 1

Dataset description

The dataset consists of 800 grayscale 32×32 -pixel of images of the faces of actors: 'Aaron Eckhart', 'Adam Sandler', 'Adrien Brody', 'Andrea Anders', 'Ashley Benson', 'Christina Applegate', 'Dianna Agron', 'Gillian Anderson'. The face appears in different angle of views and lighting, and there is considerable variation in the appearance of the face, most of cropped-out faces cannot be aligned with each others. However, examining the dataset, it appears that there are more front faces than profile faces, and most of the faces has good lighting that we can see the features on the face clearly. Also, some of the bounding boxes is not accurately cropping the face out. There are some faces in the dataset do not include part of the face, and some cropped-out pictures are even not the face of the person. A random sample of 25 faces is shown in Figure 1.



Figure 1: A random selection of 25 faces from the dataset. Generated using `display_25_rand_images()`

Part 2

Seperate datasets and compute the average face and the eigenfaces

Dataset is seperated into training set, validation set, and test set by their rendering order. The first 100 faces rendered are in training set, next 10 faces are in validation set, and last 10 faces are in test set. Using the training set of faces, Figure 2 displays the first 25 eigenfaces and Figure 3 displays a the average face in the training set. These images were generated using the function `display_save_25_comps()`.



Figure 2: The first 25 principal components of the faces. Generated using `display_save_25_comps()`.



Figure 3: The mean image. Generated using line 211 - 213.

Advice: make observations about the output, and try to explain them.

Part 3

Compute and report the accuracy

With $k = 2, 5, 10, 20, 50, 80, 100, 150, 200$, code is written to recognize every face in a set of faces using the top k eigenfaces.

In the Validation Set, the recognition performance is as following:

k	Accuracy
2	0.275
5	0.475
10	0.55
20	0.6375
50	0.6875
80	0.7375
100	0.725
150	0.7375
200	0.725

Then I applied $k = 80$ which had a highest accuracy to the test set, and got a resulting accuracy rate of 0.55. Some failure cases are shown in 4.



Figure 4: some of the failure cases

Failure could be caused to lighting in background, angle of view, or accuracy of bounding boxes.

Part 4

Recognizing gender of person

With $k = 2, 5, 10, 20, 50, 80, 100, 150, 200$, code is written to recognize the gender of every face in a set of faces using the top k eigenfaces.

In the Validation Set, the gender recognition performance is as following:

k	Accuracy
2	0.725
5	0.85
10	0.875
20	0.875
50	0.9375
80	0.9375
100	0.925
150	0.9125
200	0.925

Then I applied $k = 50$ which had a highest accuracy to the test set, and got a resulting accuracy rate of 0.85.