**EE4408 Project Report: Face Recognition with Eigenfaces**

Zhou Yuanyuan

1. **Overview—Define Problem**

Using principal component analysis (PCA) features to get eigenfaces and implement the function of face recognition (MLE, KNN included).

1. **Data Description**

We obtain the dataset from the Yale face database which consists of a set of 15 grayscale face images. Each image is of dimension 195\*231 pixels and each pixel uses 8 bits for grayscale.

The training set consists of normal faces of 15 people and the test set consists of images of the same 15 people with different expressions or under different lighting.

1. **Project Process**

This project is based on the lecture of principal component analysis.

For training images: Use the training images to produce a set of eigenfaces. Use Euclidean distance as distance measure for computing 𝑑𝑖. Choose the threshold T that produce the proper results.

For each test image: the image after subtracting the mean face, its PCA coefficients, the reconstructed face image, distances 𝑑𝑖, and classification result.

1. **Steps and Results**
2. Read Images

Reads the image in grayscale mode and converts the image into a matrix, the length of the image is 450450. Combine them into an image matrix A450450\*15

1. Mean & Normalized Faces

The matrix values of the training images are added and divided to obtain the average value which is the mean face(fig.1).

The 15 images of the training set are subtracted from the average face, which is the individual feature of each image, i.e. the normalized face(fig.2).

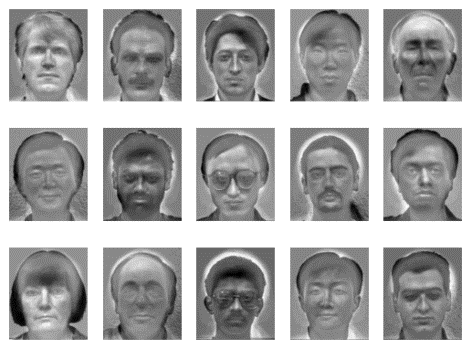
 

Fig Mean Face Fig Normalized Faces

1. Use Principal Component Analysis (PCA) to generate eigenmatrix

The original data minus the mean(m) of each feature (195\*231) to make it centered. Calculate the covariance matrix.

In order to simplify the calculation and increase running speed, calculate the eigenvalues and eigenvectors of theLow-dimensional matrix.

**X**450450\*450450 and **X** 15\*15 have the same eigenvalues, but the eigenvectors of **X**15\*15 equal to the eigenvectors of **X**450450\*450450 right multiplication **B**.

1. Extract Principal Components

Now we try to find the projected data, which will form the eigen space. The value of a certain eigenvalue divided by the sum of all eigenvalues is the variance contribution rate of the eigenvector. Then we get the cumulative proportion of variance explained vector, selecting the variance value greater than 0.8 as the principal component(fig.3).

At this point we have compressed the image dimension from 450450 to 8 and we plot the eigen faces(fig.4) from the principal components. The weights of the 8 principal components of each training image can also be calculated, on the basis of which the next recognition operation can be performed.

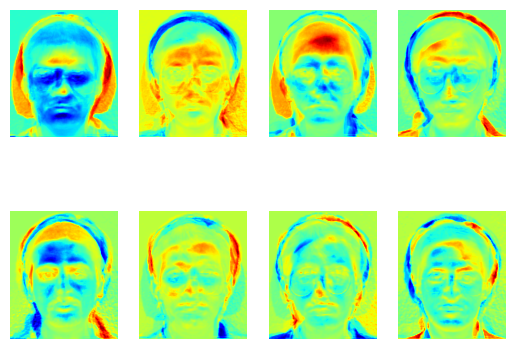
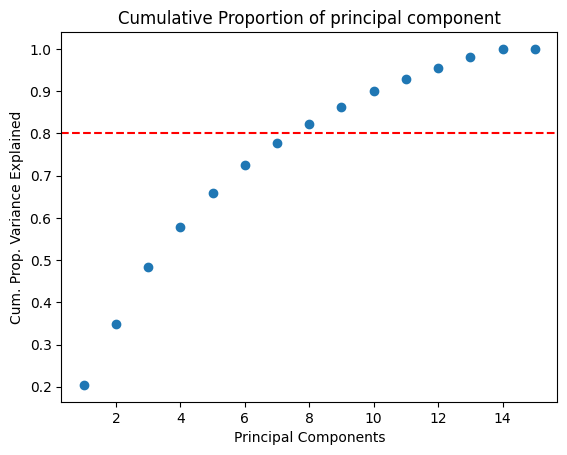


Fig Cum. Prop. Principal Component Fig Eigen Faces

1. Recognize test images

Project every test photo into the reduced dimensionality coordinate system, find out the weights, and then calculate their Euclidean distance, according to **MLE**, choose the minimum distance(di) one in the training set as the result. i.e. finding the min |W - W\_{unknown}|

For the recognition part, manually choose the threshold T, finding the nearest neighbor (**KNN**) among database images that are less than the threshold distance for matching.

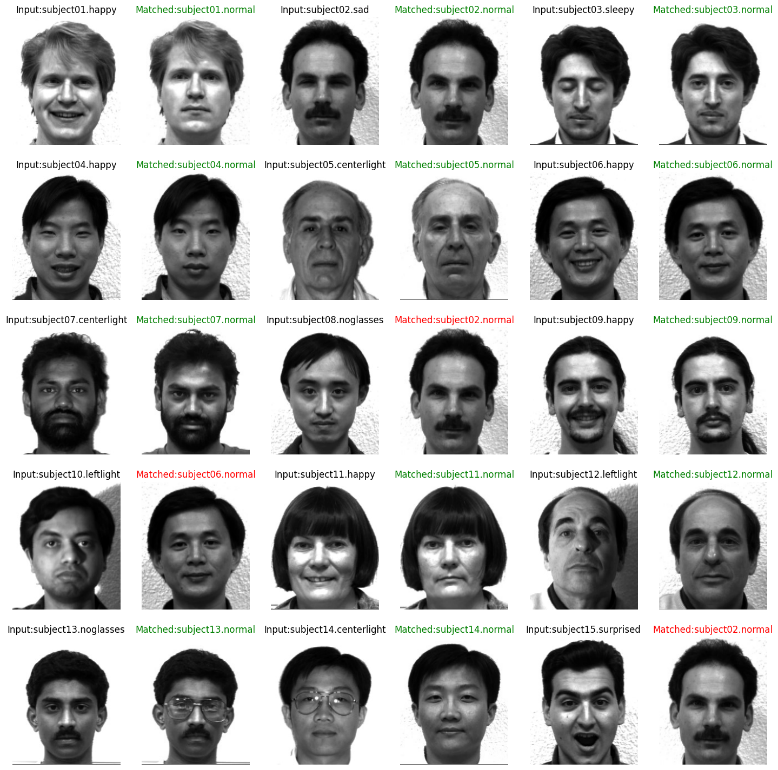


Fig Recognition Results

Correct predictions: 12/15 = 80.0%

1. **Problems**
2. The selection of the number of principal components and thresholds is somewhat subjective and affects the accuracy of the results. For example, for the eight principal components and threshold of 100,000,000 selected in this project, the correctness of the test set was only 80%.
3. Face recognition has certain requirements for the image, which cannot be scaled and rotated.
4. **Reference**

[1] The Yale Face Database. P. Belhumeur, J. Hespanha, D. Kriegman, eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection, IEEE Transactions on Pattern Analysis and Machine Intelligence, July 1997, pp. 711-720.