Assignment 3
Face Recognition using Eigenfaces

CMPE 58Z Introduction to Biometrics **Due: June 23, 2020, 23:59**

In this assignment, you are going to design a face recognition system using the Eigenface approach [1]. Eigenface technique is used to extract features from face images. The steps of the Eigenface feature extraction method are as follows:

1. Given a set of M face images: I_1, I_2, … I_M, vectorize pixel values to obtain pixel column vectors F_i. If the row and column size of an image I_i is P×Q, then the size of F_i is R×1 where R=P×Q. You can vectorize an image by simply concatenating rows. F_i will be referred to as a face image from now on.
2. Compute the average face S using a set of training faces.
3. Compute difference faces D_i by subtracting the mean face S from each face: $D_i = F_i - S$
4. Form a training matrix $A = [D_1, D_2, ..., D_M]$ containing difference faces in columns. The size of A is R×M.
5. Define matrix L as $L = A^TA$.
6. Compute the eigenvectors v_i of L. v_i represents the i^{th} eigenvector.
7. Pre-multiply eigenvectors with A to obtain Eigenfaces u_i: $u_i = Av_i$.
8. Select N Eigenfaces $u_1,u_2,…,u_N$ having the biggest corresponding eigenvalues and form the Eigenface transformation matrix $U = [u_1,u_2,…,u_N]$. It is useful to normalize eigenvectors so that they have unit length. Normalization is simply done by dividing each eigenvector to its norm.
9. Eigenface Projection: Project a given face image F_i into the Eigenface space by: $E_i = U^T(F_i-S)$.

Steps from 2 to 5 correspond to Eigenface training. The output of Eigenface training is the Eigenface transformation matrix U. Step 6 explains how to use U to extract Eigenface features from a given face image. The maximum feature dimensionality of reduced features is bounded by the number of images in the training set. During projection phase, Step 6, mean face is subtracted from the face image. For further details on the Eigenface technique, please refer to Calculating Eigenfaces section (pages 73-75) in [1].

It is possible to reconstruct the original face image, up to a certain quality, from the Eigenface features. To reconstruct a face image from Eigenface features E_i (i.e., Eigenface coefficients), you have to perform a different operation. Find out the reconstruction equation yourself. After obtaining the reconstructed face vector, you can form the reconstructed face image matrix by reversing the row concatenation operation given in Step 1.

In this assignment, you are a given a face database, the ORL face database, containing 40 subjects with 10 face images per subject. You will use the first five images of each person to form the training matrix A to compute U. Using U, you will compute Eigenface features for all the images in the ORL face database. The remaining five images of every subject will be used as a test set. You are going to find the identity of images in the test set. Each subject will have the first five images as their enrolled images. Therefore, in our system, each subject is represented by their five Eigenface feature vectors. In order to compute the dissimilarity of two feature vectors, Euclidean distance will be used.

In your report, provide the following results:

Eigenfaces
First 20 Eigenfaces, as shown in Figure 1.

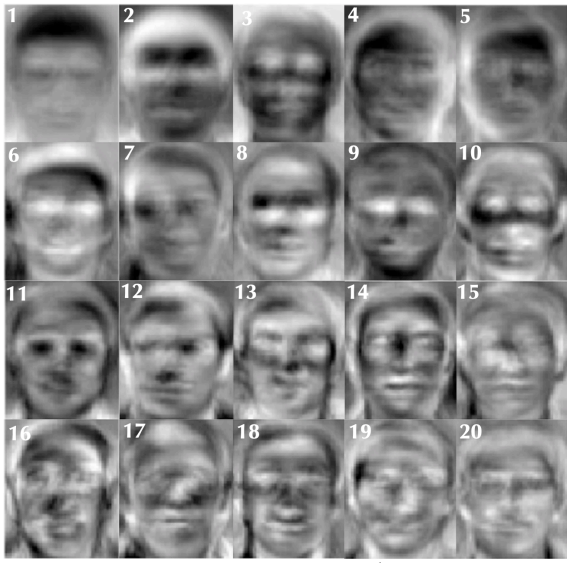

Figure 1. First 20 Eigenfaces.

Reconstructions

Reconstruct a sample face image from the test set using Eigenface features of size: [2, 5, 10, 20, 80, 100, 200] and plot the reconstructed images together with the original face image. See Figure 2a for an illustration which uses different Eigenface dimensionalities. Perform similar reconstruction for a non-face image, as shown in Figure 2b. Comment on the difference between these two cases in terms of the reconstruction ability of the Eigenface method.

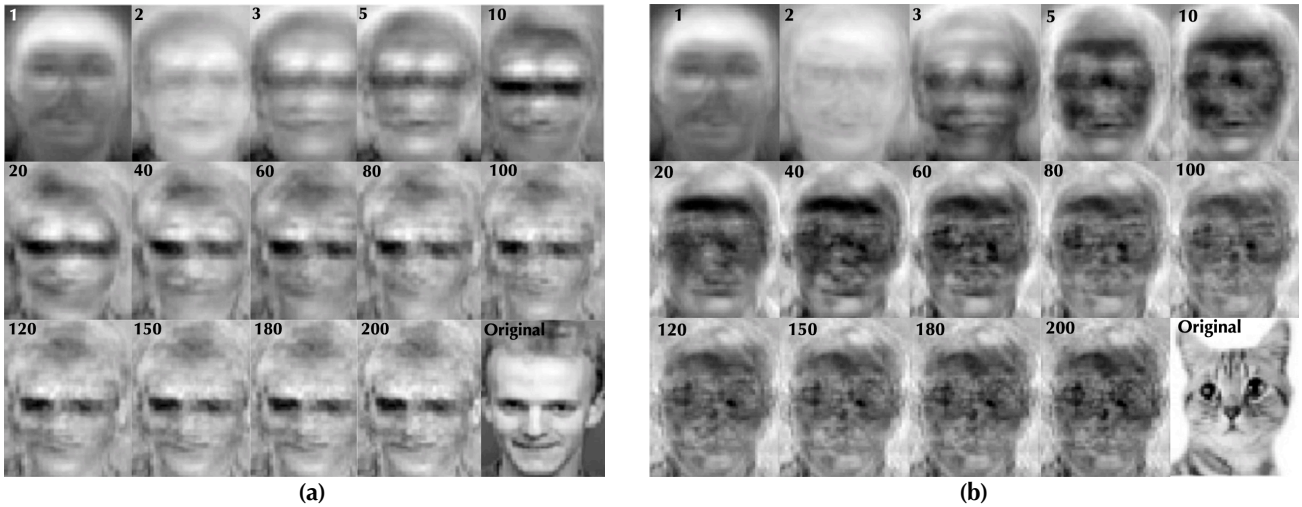

Figure 2. Reconstruction results for a face and a non-face image.

Classification Accuracy

Using your test set, plot the identification accuracy (y-axis) vs Eigenface dimensionality (x-axis) plot. See Figure 3 as an example. As Eigenface dimensionality (i.e., Eigenface feature vector size), you can choose [5, 10, 15, ..., 200]. Provide identification rates together with Eigenface feature vector size and the total variance explained as a table, as shown in Table 1. The total variance explained by selecting the biggest K eigenvectors is computed by $\frac{\sum_{i=1}^{K}\lambda_i}{\sum_{i=1}^{R}\lambda_i}$, where R is the maximum dimensionality. Note that eigenvalues are sorted in descending order: $\lambda_i > \lambda_{i+1}$

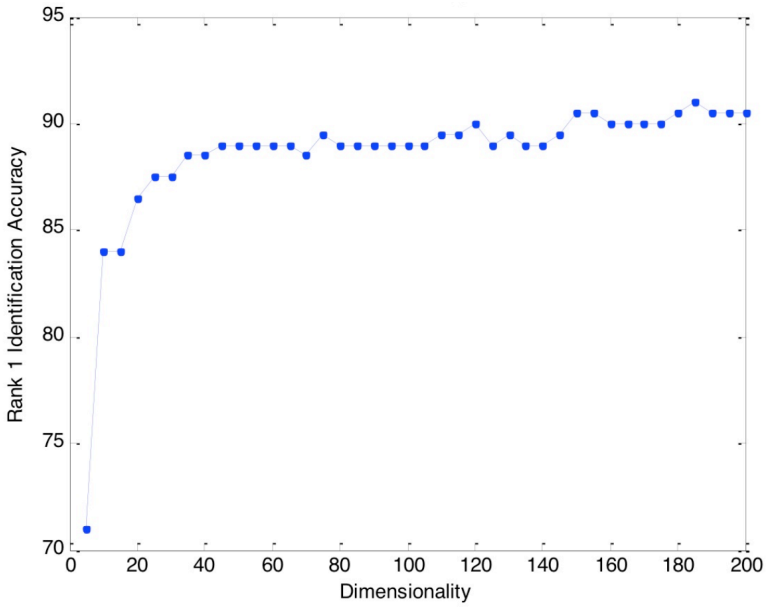

Figure 3. Identification accuracy vs Eigenface dimensionality.

Table 1. Identification performance and total variance explained with respect to Eigenface dimensionality.

#Eigenfaces	Variance	Identification Rate (%)
10	0.12	34.77
20	0.26	46.23
40	0.45	53.32
60	0.67	59.12
80	0.83	63.87
…	…	…
200	0.98	91.02

Submission Information
Deliver your source code and report in a compressed zip file. The main script of your assignment should be named as assignmentX.ext, where X is the assignment number and .ext is the file extension of the chosen programming language. See the submission guide section given at the end of this document for more details. When the script is executed, it should produce the 1) Eigenface images, 2) Reconstruction face example image (only one face image reconstruction is sufficient), 3) identification rate for a specific Eigenface dimensionality. Your code is not expected to produce Eigenface dimensionality vs identification rate plots.

Tip
For identification performance plots (accuracy vs Eigenface dimensionality), you do not need to re-compute eigenvectors/eigenvalues for each dimensionality. Compute them once and use the required part of the eigenvector matrix for feature extraction to save time.

References
1. Turk, M., and Pentland, A.: Eigenfaces for recognition. J. Cognitive Neuroscience 3(1), 71–86 (1991)

Evaluation Criteria

		Points
Correctness of the solution		60
Report (Contents, format, etc.)		30
Compliance to Submission Rules (Directory structure, file formats/naming, organization, etc.)		10
	TOTAL	**100**

Submission Guide

Submission Files
Submit a single compressed (.zip) file, named as name_surname.zip, to Moodle. It should contain all source codes (under the \code directory), report (in PDF format, under the \report directory) and all other files if needed (under \misc directory)

File Naming
Name your report as name_surname.pdf. Name the main code which is used to execute your assignment as assignmentX.py, where X is the assignment number and .py is the extension of for Python, given as an example.

Late Submission Policy
Maximum delay is two days. Late submission will be graded on a scale of 50% of the original grade.

Mandatory Submission
Submission of assignments is mandatory. If you do not submit an assignment, you will fail the course.

Plagiarism
Leads to grade F and YÖK regulations will be applied