Face Recognition
Bayes Implementation of Eigenfaces

Ritu Ann Roy George
B170106EC
B-Batch
NIT Calicut

***Abstract*—This objective of this assignment is to implement Bayes theorem of conditional probability with the Eigenfaces method, the foundation of facial recognition systems. The Yale Face database is utilised in this experiment.**

I. Introduction

Face Recognition is one of the most popular applications of computer vision. Principal Component Analysis approach used for dimensionality reduction in face recognition systems, was first developed by Sirovich and Kirby in 1987 and implemented as EigenFaces by Turk and Alex Pentland in face classification in 1991 . The Bayesian Face Recognition paper by Baback Moghaddam, Tony Jebara and Alex Pentland discusses using an non eucliean distance based technique for facial recognition, i.e. using Bayes Theorem of conditional probability.

II. Bayes Theorem

Bayes's theorem is stated mathematically as the following equation:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \tag{1}$$

where A and B are events and $P(B) \neq 0$ and$P(B) \neq 0$.

$P(A \mid B)$ is a conditional probability: the likelihood of event A occurring given that B is true.
$P(B \mid A)$ is also a conditional probability: the likelihood of event B occurring given that A is true.
$P(A)$ and $P(B)$ are the probabilities of observing A and B respectively; they are known as the marginal probability.

We know that the Yale database contains 11 images of different facial expressions for each of the 15 persons (classes). Here 10 images from each class is used for training and 1 image as the testing image.

In our problem here, taking A as the class and B as the test image, the problem statement is to predict the class which the given test image belongs to.

III. Bayesian Similarity Score

The research paper that we implement in this assignment presents a probabilistic similarity measure based on the Bayesian belief that the image intensity differences, denoted by $\Delta = I_1 - I_2$ are characteristic of the typical variations in appearance of an individual. In particular, two classes of facial image variations are defined: Intrapersonal Ω_I and extrapersonal Ω_E. To simplify the Bayes Theorem, we consider only the Intrapersonal likelihood for different Δ into account (Utilising the ML (Maximum Likelihood Criteria). The similarity score is then expressed in terms of probability

$$S' = P(\Delta|\Omega_I) \tag{2}$$

Considering N images in the training dataset, each of which belongs to one of the k different classes (each class having multiple facial expressions of a particular individual), we have $\Delta = I_i - I_n$ where the new image is represented as I_n, and $i = 1, 2....N$. Assuming that the distribution of Ω_I is Gaussian in nature, the probability density can be expressed as:

$$P(\Delta|\Omega_I) = \frac{e^{\frac{-1}{2}\Delta^T \Sigma_I^{-1}\Delta}}{(2\pi)^{\frac{D}{2}}|\Sigma_I|^{\frac{1}{2}}} \tag{3}$$

Note that the density is Zero-Mean since for each $\Delta = I_j - I_k$ there exists a $\Delta = I_k - I_j$.

Now we apply Whitening transformation to improve the computation of $P(\Delta|\Omega_I)$ such that every image is stored as vectors of the whitened subspace coefficients i_j

$$i_j = \Lambda_I{}^{\frac{-1}{2}} V_I I_j \tag{4}$$

where Λ and V_I are the matrices of the largest eigenvalues and eigenvectors of the covariance matrix Σ_I. This now reduces the likelihood computation to calculating the Euclidean distances for the exponents. Hence

$$S' = P(\Delta|\Omega_I) = \frac{e^{\frac{-1}{2}||i_j - i_k||^2}}{(2\pi)^{\frac{D}{2}}|\Sigma_I|^{\frac{1}{2}}} \tag{5}$$

IV. Algorithm

A. Read and Vectorize the Input

Extract the images from the Yale Face Database and flatten each image into a row vector for easy computation. Here the last image from each of the 15 classes is set aside to act as the test image. The rest are utilised as the training dataset.

B. Principal Component Analysis

- Find the difference between all the pairs of images in the training data set.
- Compute Co-variance Matrix C
- Obtain eigenvalue and eigenvector
- Set the number of principal components for reducing dimensions
  Arrange the eigenvalue and the associated eigenvectors in the decreasing order of eigenvalues to obtain the components having the maximum information.

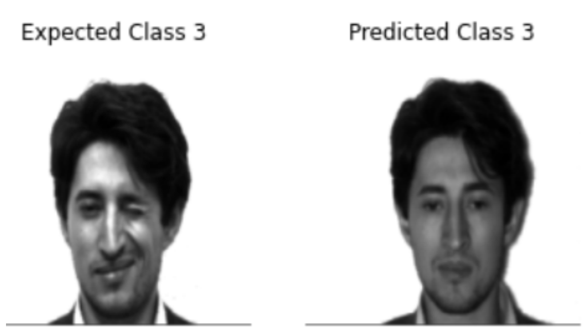
C. EigenFaces

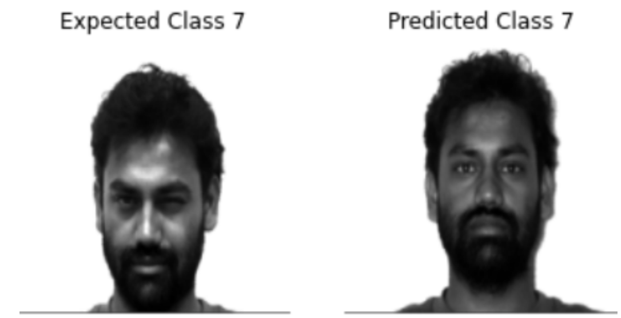
- Compute Eigenface for one class. Use the eigenvalue and eigenvectors to compute the coefficients for each image in the class and take the mean of all the coefficients to obtain the eigenface.
- Similarly find corresponding eigenfaces for all the classes and append them to the set of eigenfaces.

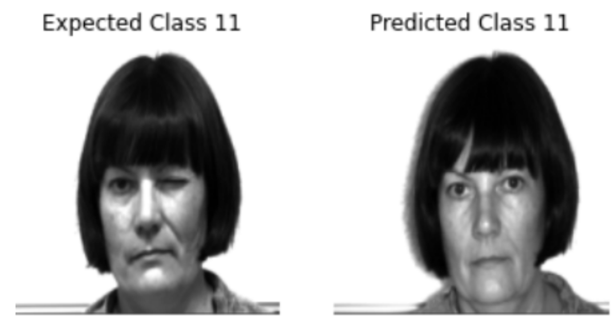
D. Recognition: Similarity Prediction Using Bayes Rule

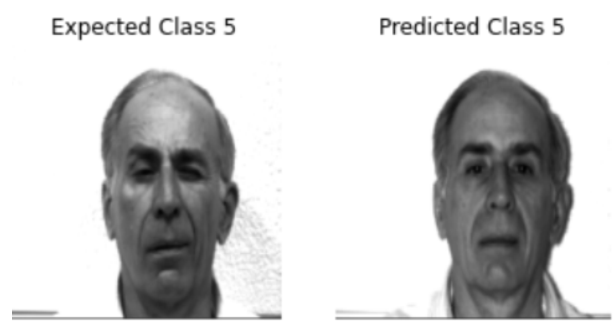
- Find the difference of coefficients of the test image at hand and each of the eigenfaces in the set.
- Find the probability of the test image being in class C_i by passing this difference in the equation for similarity prediction using Gaussian distribution. Repeat for all the classes.
- Find the class giving the maximum similarity value. This is the predicted class.
- Repeat the steps for N test images. The accuracy of the model for a fixed dimensionality can be expressed as $Acc = (C/N) * 100$, where C denotes the number of times the predicted class matched the expected class.

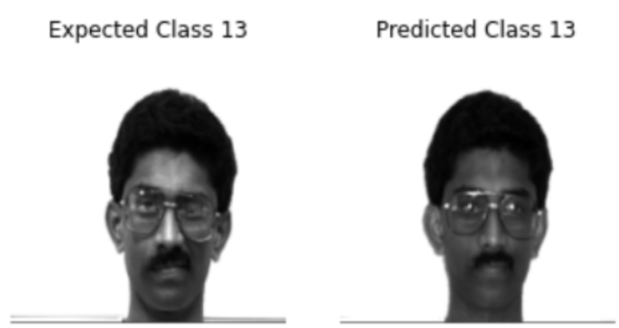
V. Observations

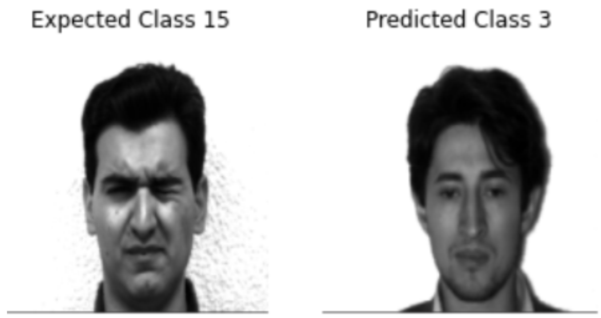


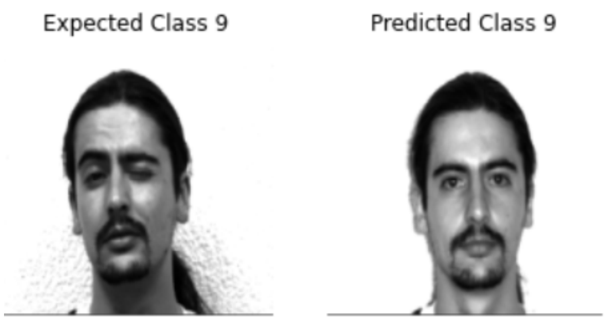












VI. Inference

From the Bayesian Face Recognition paper by Baback Moghaddam, Tony Jebara and Alex Pentland, we infer:

- Image intensity differences are characteristics of typical variations in the appearance of an individual
- Two kinds of facial variations are present: intra personal and extra personal.
- The problem can be approached with the maximum aposterior probability rule (MAP).
- An even simpler analysis can be performed using maximum likelihood (ML) estimation taking only intrapersonal likelihood into account.
- From the above experiment, an accuracy of 85.71% is observed.

VII. Result

Face Recognition is performed by Bayes Implementaion on Eigenfaces.