Image Processing Laboratory

Face Recognition using Eigen Face Decomposition

Group 15 and 16

Vineet Jain - 14EC35020
Oindrila Saha - 14EC35005
Arnab Mondal - 14EC35031
Sandeep Mishra - 14EC35033

Abstract

Given a dataset of faces, an approach to the recognition of human faces is presented. This approach treats face recognition as a two-dimensional recognition problem, taking advantage of the fact that faces are normally upright and thus may be described by a small set of 2-D characteristic views. Face images are projected onto a feature space that best encodes the variation among known face images. This feature space is defined by the 'eigenfaces', which are the eigenvectors of the set of faces; they do not necessarily correspond to isolated features such as eyes, ears, and noses. In the recognition process, an eigenface is formed for the given face image, and the Euclidean distances between this eigenface and the previously stored eigenfaces are calculated. The eigenface with the smallest Euclidean distance is the one the person resembles the most.

Introduction

The objective of this project is to build a face recognition system with eigen face decomposition given a dataset of faces.
Objectives:

1. Show the average 'ghost face' learned from the dataset
2. Show the eigen faces for test subjects
3. Report mean accuracy of recognition

Dataset used: Faces-152 dataset
The dataset has 152 persons with 20 images for each person. Randomly select 15 images from each person for training purpose. Remaining 5 images will be used for testing.

The basis of the eigenfaces method is the Principal Component Analysis (PCA). The Principal Component Analysis is a method of projection to a subspace and is widely used in pattern recognition. An objective of PCA is the replacement of correlated vectors of large dimensions with the uncorrelated vectors of smaller dimensions. Another objective is to calculate a basis for the data set. Main advantages of the PCA are its low sensitivity to noise, the reduction of the requirements of the memory and the capacity, and the increase in the efficiency due to the operation in a space of smaller dimensions. The strategy of the Eigenfaces method consists of extracting the characteristic features on the face and representing the face in question as a linear combination of the so called ‘eigenfaces’ obtained from the feature extraction process. The principal components of the faces in the training set are calculated. Recognition is achieved using the projection of the face into the space formed by the eigenfaces. A comparison on the basis of the Euclidean distance of the eigenvectors of the eigenfaces and the eigenface of the image under question is made. If this distance is small enough, the person is identified. On the other hand, if the distance is too large, the image is regarded as one that belongs to an individual for which the system has to be trained.

Algorithm

---

1. First, the training images of dimensions $N * N$ are read and they are converted to $N^2 * 1$ dimensions. A training set of $N^2 * M$ dimensions is thus created, where M is the number of sample images.

2. The average of the image set is calculated as:

   $$\psi = 1/M * \sum_{i=1}^{M} \Gamma_i$$

   where ψ : average image, M : number of images, Γ_i : image vector.

3. The eigenfaces corresponding to the highest eigenvalues are retained. Those eigenfaces define the face space. The eigenspace is created by projecting the image to the face space formed by the eigenfaces. Thus the weight vectors are calculated. Dimensions of the image are adjusted to meet the specifications and the image is enhanced in the preprocessing steps of recognition. The weight vector of the image and the weight vectors of the faces in the database are compared.

4. Average face is calculated and subtracted from each face in the training set. A matrix (A) is formed using the results of the subtraction operation. The difference between each image and the average image is calculated as :

   $$\phi_i = \Gamma_i - \psi \quad , \quad i = 1, 2, ..., M$$

   Where ϕ_i is the difference between the image and the average image.

5. The matrix obtained by the subtraction operation (A) is multiplied by its transpose and thus covariance matrix C is formed:

   $$C = A^T A$$

   where A is formed by the difference vectors, i.e.,

   $$A = [\, \phi_1, \phi_2, ..., \phi_M]$$

6. The dimensions of the matrix C is N*N. M images are used to form C. In practice, the dimensions of C is N*M. On the other hand, since the rank of A is M, only M out of N eigenvectors are nonzero.

7. The eigenvalues of the covariance matrix is calculated. The eigenvectors are computed from the eigenvalues of the covariance matrix. The selected set of eigenvectors are multiplied by the A matrix to create a reduced eigenface subspace.

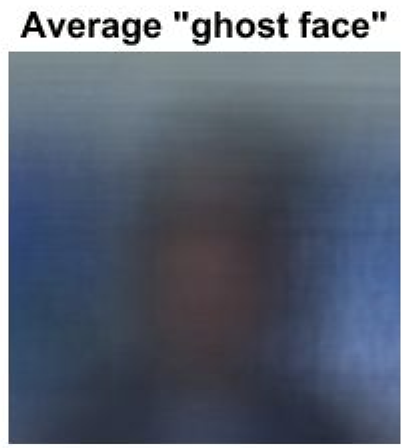
8. The eigenvectors of smaller eigenvalues correspond to smaller variations in the covariance matrix. The discriminating features of the face are retained. The number of eigenvectors depend on the accuracy with which the database is defined and it can be optimized. The group of selected eigenvectors are called the eigenfaces. Once the eigenfaces have been obtained, the images in the database are projected into the eigenface space and the weights of the image in that space are stored. To determine the identity of an image, the eigen coefficients are compared with the eigen coefficients in the database.

9. The eigenface of the image in question is formed. The Euclidean distances between the eigenface of the image and the eigenfaces stored previously are calculated.

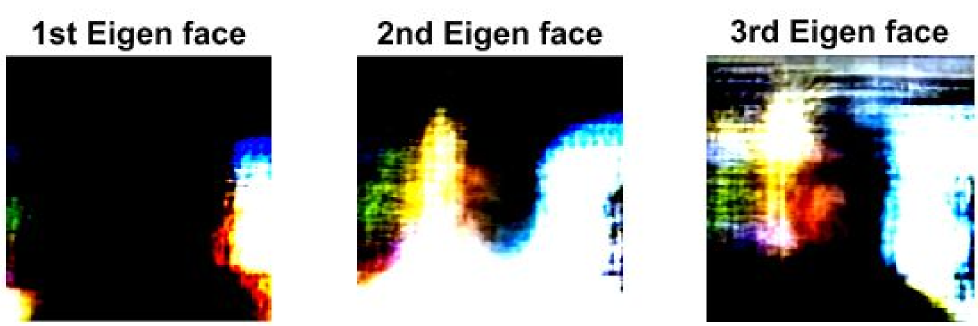
10. The person in question is identified as the one whose Euclidean distance is minimum below a threshold value in the eigenface database. If all the calculated Euclidean distances are larger than the threshold, then the image is unrecognizable.

Results

The average 'ghost face' learned by the algorithm based on the training data is shown below.



The eigen faces, which are the eigen vectors of the covariance matrix multiplied by the matrix A, are shown below. Since there are a large number of such eigen faces, only those corresponding to the highest three eigen values are shown.



For each person, 15 out of 20 images were randomly selected for training and the remaining were used for testing. After applying the algorithm on the test images, the accuracy was computed.

```
Testing Accuracy = 98.6684 %
```

Conclusion

The eigenfaces method, which is based on the concept of Principal Component Analysis (PCA) is applied on the Faces-152 dataset. Eigenfaces are obtained during training phase and recognition is achieved by projecting the faces onto this eigenface subspace. The average 'ghost face', a few of the eigenfaces and the mean accuracy of recognition were reported.