Face Recognition using Eigenfaces SMAI Course Project

Satarupa Guha
IIIT Hyderabad
satarupa.guha@research.iiit.ac.in

Ayushi Dalmia
IIIT Hyderabad
ayushi.dalmia@research.iiit.ac.in

Abstract— **The goal of this project is to get familiarized with the state-of-the-art techniques for representation, identification and recognition of faces. We take the Information Theory approach, where the Principal Components of each of the face images are used to transform the faces into an Eigenface space. Once we get a new image, we project it on the Eigenface space, compare its position in this space with the position of known individuals, so as to perform face recognition, verification and reconstruction of faces as well as non faces.**

Keywords—Eigenface; Principle Component Analysis; Face Recognition; Face Reconstruction; Face Verification

INTRODUCTION

Human beings recognize hundreds of faces every day without the blink of an eye. With the advancement of the field of Artificial Intelligence, it is intriguing to model this facet of the human brain so as to simulate it by computers. Computers adept at face recognition can serve a multitude of purposes- security systems, criminal identification, Human Computer Interaction, etc. Face Detection, even without Face Recognition is very important nowadays, in the fields of Image Processing, Photography, Social Network Analysis, etc.

We implement the approach proposed in [1], i.e. the Information Theory approach where a face image is decomposed into a small set of characteristic feature images called "eigenfaces", which are most of the times much smaller than the entire image used as feature image. These "eigenfaces" are nothing but the principal components of the initial training set of images. They do not necessarily correspond to intuitive notion of face features such as eyes, ears or nose separately. The projection operation is done by representing a particular face by the weighted sum of the eigenface features. Recognition is done by projecting a new test image into a subspace spanned by the eigenfaces, i.e. the face space, and then determining which face it is by comparing its position in the face space with the positions of training set face images. This approach is simple, fast and insensitive to small or gradual changes in the face image, as compared to other approaches as proposed in [2] and [3]. The information theory approach is very useful because it gives insight into the information content of faces images. In order to extract the relevant information in a face image, it is essential to somehow capture the variation among the different face images, thereby making any intuitive judgment of features redundant. The idea is to represent each face in terms of the linear combination of those eigenfaces that have the largest eigen values, and hence account for the maximum variance. It is an extremely compact representation when compared to the images themselves.

However, the calculation of eigenfaces is far more challenging than it appears to be on the surface. Computing the eigen vectors of the images basically means computing the eigen vectors of the covariance matrix of the images. If each image is of size $N \times N$, then the covariance matrix would be of size $N^2 \times N^2$. Computing the eigen vectors for a matrix of such dimension is, needless to say, computationally very intensive. We tackle this by using basic linear algebra where the computation reduces significantly.

Apart from identifying a given image, we also try to develop a system which can verify it. Given an image along with a label, the system verifies whether the identity as claimed is true or not.

Imposing the eigenfaces on new faces as well as non faces leads us to interesting results. It is found that the non faces after reconstruction are found to have blurred facial features.

The rest of the paper is organized as follows: In Section II we briefly present some of the related works done so far. In Section III we describe the dataset used. In the next section details of the experiments performed is explained. In Section IV the Results for the experiments conducted are shown.

LITERATURE SURVEY

Researchers have focused in the area of face recognition since the middle nineties. During 1964 and 1965 Bledsoe along with Helen Chan and Charles Bisson worked on recognizing faces by using fiducial marks. This was a hybrid human computer system where the features were marked on the photograph by a human and these features were used by the computer to recognize the image. In order to automate this, Goldstein, Harmon and Lesk worked at Bell Labs to extract features chosen by the human. Now these features were extremely subjective in nature and required human evaluation. Hence a complete automated system was needed. After this in the 1970s and 1980s the template matching approach was taken. As the number of images and the complexity of problem increased people started taking other resorts. Connectionist approaches to face identification attempted to incorporate the overall shape of the face and not the individual features. The statistical based approach of eigenface came after this. Most practical face recognition systems use this method. Later, face recognition system using Linear/Fischer Discriminant Analysis have also been found to perform successfully. Neural Networks have also been used in face recognition. They address several problems of face recognition like the gender classification, face recognition and classification of facial expressions.

DATA SETS

The experiments are conducted on the following datasets:

Yale Dataset

The Yale Dataset consists of 760 gray scale images. There are 38 different subjects, with 20 images per subject. All of them are captured under controlled conditions (illumination, pose and expression) and cropped to contain the face only. We have further done some processing on this dataset so as to include only those images that have azimuthal and elevation angles within the range [-30,+30].

CMU-PIE Dataset

The CMU-PIE dataset consist of 2856 gray scale images. It has 68 different subjects, and for each subject there are 42 images. This dataset too include images take under various Poses, Illuminations, and Expressions (PIE).

S13SD Dataset

The S13SD Dataset consists of 680 gray scale images under various positions, illuminations and expressions with 136 subjects. The number of images per subject varies from subject to subject. As a preprocessing step, we remove the images of those subjects who have less than 5 samples. We are hence left with 483 images in total, and 76 subjects.

THE EIGENFACE APPROACH

We use the eigenface approach to conduct Face Identification, Face Verification and Face Reconstruction on the datasets described above.

Face Identification

Given a set of images, we first create the face space. Now when a new test image comes in we need to identify and return the closest match of the test image in the face space. The evaluation of the accuracy of the system is done using 4-fold

cross validation using 75% of training data and 25% of testing data. For the S13SD dataset, we perform hold-one-out technique of cross validation, where we take only one image from each class as test data and the remaining as training data. The method employed is as described next. The classification task has been done using both K-Nearest Neighbor as well Support Vector Machine.

Let A be the image matrix where each column corresponds to an image denoted as A_1 , A_2 , A_3 . . . A_n . Let there be M images and let the size of each image be NxN. Hence A is of dimensions $N^2 \times M$. Each of the training images are transformed into eigenfaces. In order to do so we first calculate the mean face for all the M images. Let $\varnothing$ denote the mean of all the images. It is given by $\frac{1}{M}\sum_{i=1}^{M} A_i$. The mean image for the Yale, CMUPIE and the S13SD Dataset is given in Figure 1.

Each face A_i differs from the mean image given by $B_i = A_i - \varnothing$ Matrix B is computed by subtracting each image from

Figure 1a Figure 1b Figure 1c

Figure a: Mean Image for Yale Dataset, Figure 1b: for CMUPIE Dataset, Figure 1c: for S13SD Dataset

the mean image. Now we need to compute the principle components of the matrix B which returns N^2 orthogonal vectors. The orthogonal vectors are basically the eigen vectors, μ_k for the covariance matrix of B. Let C be the covariance matrix of the matrix B. The vectors p_k and values μ_k are the eigen vectors and the eigen values respectively, of the covariance matrix,

$$\frac{1}{M}\sum_{i=1}^{M} B_i B_i^T$$

Of this large set of vectors we calculate the select the best M vectors which best describes the image on the basis of the eigen values, μ_k. We call these M orthogonal vectors as eigenfaces and the orthogonal space as face space.

The covariance matrix C is formed by $B*B^T$, the dimension of which is $N^2 \times N^2$. Therefore, the computation of eigen vectors of this matrix is very computationally intensive. As N^2 >> M eigen vectors beyond M are not significant. Hence, we rather compute the eigen vectors of B^T*B, which is of dimensions $M \times M$. If the number of data points in the image space is less than the dimension space, ie. If $M<N^2$, there will be M-1, instead of N^2, meaningful eigen vectors.

Let the eigen vectors v_i of B'*B such that

$$B^T B v_i = \text{¿} \quad \mu_i v_i$$

Pre-multiplying both sides by B, we have

$$B B^T B v_i = \text{¿} \quad \mu_i B v_i$$

After we get the eigen vector matrix V, we multiply it by the matrix B, so as to account for the changes we made to make the computations easy. B*V is of dimesions $N^2 \times M$. This reduced our computations by a huge extent from the order of the number of pixels in the images (N^2) to the order of the number of images in the training set (M). Now, among the M eigen vectors obtained we take only the top K eigen vectors based on the eigen values they correspond to. So, the eigen vector matrix is further reduced to dimension $N^2 \times K$ where K<M. Since accurate reconstruction is not an issue, and since the eigenfaces have been observed to be adequate for describing face images under controlled conditions, we can use the same for the purpose of face identification. Finally, we contruct the $M \times k$ matrix of top eigen vectors. The eigenfaces u_l is computed as follows

$$u_l = \sum_{i=1}^{k} v_i B_i$$

After obtaining the eigenface for the training data, we do the same for the test data. From the test data, we take one test image at a time from the test data, compute its eigenface and compare it with all the eigenfaces computed from the training data. The label of the eigenface, that is closest to the eigenface of the test image, is returned as the predicted label. This classification task has been done using both K-Nearest Neighbor as well Support Vector Machine.

Face Verification

For the purpose of face verification, we generate pairs of images. The pairs may contain images belonging to same class or to different classes. We randomly generate these pairs. A pair is called a genuine pair and labeled with '0' if the images in the pair belong to same class. It is called an imposter pair and labeled with '1' if the images belong to different classes. We also calculate the distance between the images in each pair. Next we plot the Receiver Operating Characteristic curve using this distance metric and the labels generated in the previous step. From the ROC curve, we also get a threshold which is the optimal operating point of the ROC curve.

After this, we take a test image and randomly assign a label to it. We call this label as claim label, and this can be the true label of the image or any other label. We need to check whether our system can identify the claim label as true or false. For this purpose, we compute the distance of this test image from all the training images having the claim label as their label. If the distance is less than the threshold, computed in the previous step, our system returns true. Otherwise, it returns false. In order to check the accuracy of face verification performed by our system, at each step we compare the claim label and true label. If they are same and our system returns True, or if they are different and our system returns False, we say, our system correctly verified the test sample. Otherwise, we conclude our system failed.

Face Reconstruction

Given a new image, whether a face or a non-face, we transform the image into eigen space. Then we add the mean face (obtained from the training data) to this eigenface representation of the test image. The image we get after this operation is the reconstructed image for the test image.

RESULTS

Face Identification

INSERT

We have compared the results of KNN between the Yale dataset, CMU-PIE and S13SD datasets. The following table shows the results

TABLE I.

Dataset	K-Eigen					
	5	***10***	***20***	***30***	***40***	***50***
Yale	91.97	98.16	99.61	99.87	100	100
CMU-PIE	92.44	95	99.85	100	100	100
S13SD	12.26	15.79	26.32	28.95	25	25

As is evident from the table, as the number of eigen vectors increases, the accuracy increases. But it is not true always. As we can see, the accuracies for the S13SD dataset increases at first as the number of eigen vectors increases, but after a point it starts dipping. This can be attributed to the fact that the eigen vectors towards the end represents the very minute details of an image, which is not very useful for discriminating between different classes. Similarly, the first few eigen vectors represent the broad overview of an image, which again is not too useful for the purpose of discrimination. For this reason, in all our experiments, we have left out the top 4 eigen vectors.

We have also compared the accuracies of the Yale dataset when the classification task is done by K-Nearest Neighbour or by Support Vector Machine respectively. The best accuracy that Yale dataset gives is 100% while for SVM it is 96.28%.

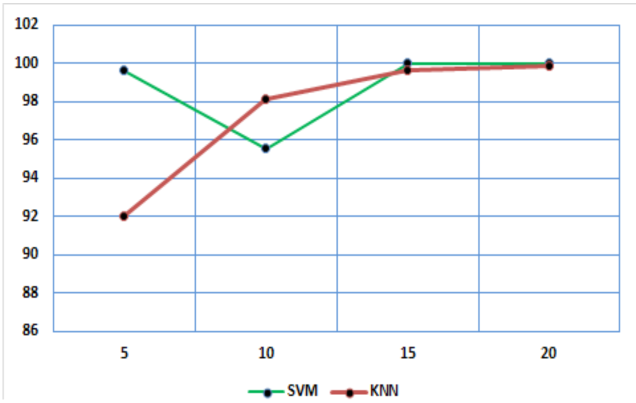
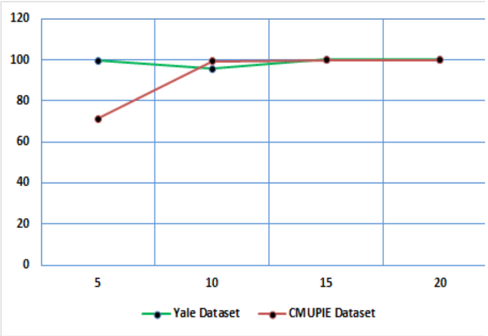


Figure : Comparison of Performance of KNN and SVM with the variation of K on Yale dataset

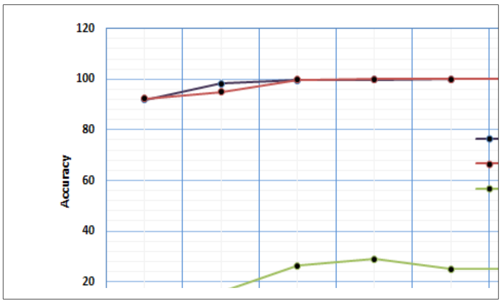
As we can see from the above statistics, KNN seems to perform better overall, and the increase in accuracy for KNN is very steep. Whereas, in case of SVM, the increase in accuracy is not very as fast. One reason for this may be that the SVM's hyperplane selection is a slow adjusting course.

The

Figure : Variation of Accuracy with K for Yale and CMUPIE Datasets on SVM



performance for Yale and CMUPIE datasets have been compared as we vary K. The classifier used here is Support Vector Machine.



Next, we have plotted the variation of the accuracy with the number of top K eigen vectors for all the datasets- Yale, CMU-PIE and S13SD datasets.

We have conducted the experimented under different conditions. We have added salt and pepper noise to some of the images and added them to the training data. Results show that this small modification boosts the overall accuracy by a small margin. We then rotate some of the images in the training data, and add the modified images to the training data. The results obtained from this step is slightly better than that obtained in the previous step. Finally, we incorporate both noise and rotation to a small fraction of the training data. This shows slightly better performance than the previous two approaches. The following table shows the variation of accuracies with the number of top eigen vectors under the three conditions.

Dataset	K-Eigen				
	10	***20***	***30***	***40***	***50***
Rotation	17.15	24.58	26.41	27.48	27.06
					87
					47

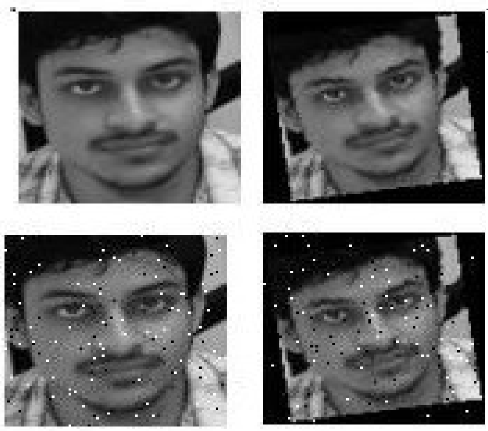


Figure 3: Top Left: Original Image, Top Right: Rotated Image, Bottom Left: Image with Noise, Bottom Right: Rotated Image with Noise *Face*

Verification

For all the datasets, we have plotted the Receiver Operating Curves. Figure 1 shows the ROC for Yale dataset. Figure II shows the ROC for CMU-PIE dataset, while Figure III shows the same for the S13SD dataset.

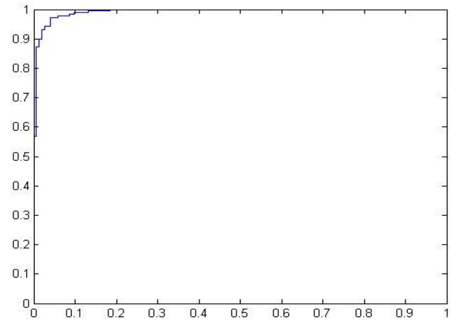


Figure : ROC Curve for Yale dataset

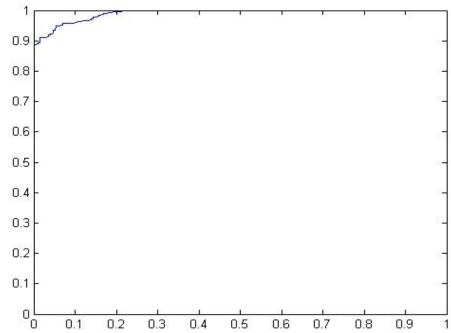


Figure II: ROC Curve for CMU-PIE dataset

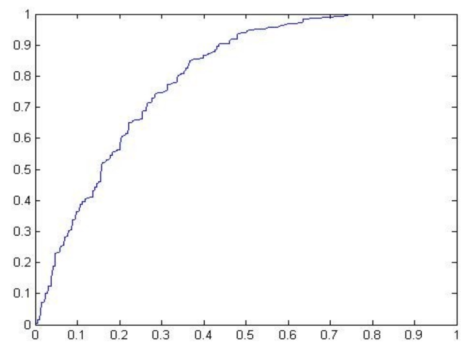


Figure III: ROC Curve for S13SD dataset

Further, we have recorded the variation of accuracy of verification with the number of top eigen vectors for all the datasets- Yale, CMU-PIE and S13SD. We plotted the first two on two different graphs because the number of top eigen vectors are different for the different datasets. This has been done to incorporate the fact that the number of classes in each of the datasets is different. Since, we implement class-wise verification, our selection of top eigen vector changes from class to class. However, for S13SD, the number of images in all the classes range roughly between 5 to 10. So in this case, we cannot vary the eigen vectors much and the accuracy after taking top 3 eigen vectors is 3.95%. For top 4 eigen vectors the accuracy has been found to be 17.12%. Some of the classes may not have more than 5 images and thus may not have more than 5 eigen vectors. Therefore we cannot vary the eigen vectors beyond 4.

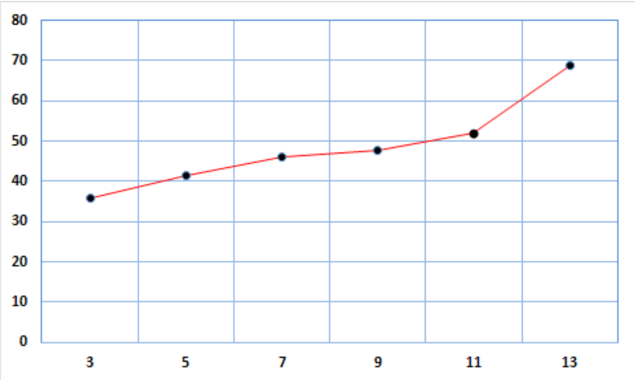


Figure : Variation of Accuracy with K for Yale

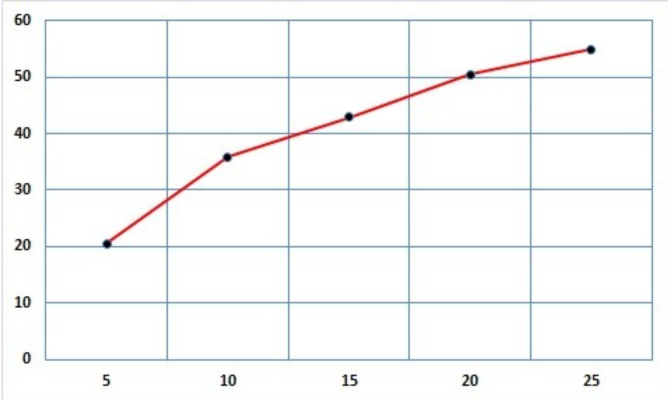


Figure : Variation of Accuracy with K for Yale

Face Reconstruction

We get some very interesting results from face reconstruction of test images. We have taken 5 test images from different sources, that include both faces and non-faces. Their corresponding reconstructed images have been shown below.

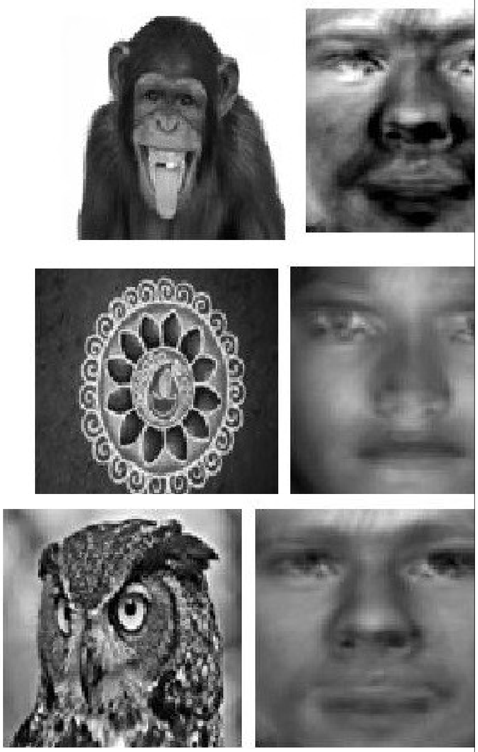


Figure : Reconstructed Non Faces

CONCLUSION

Face Recognition has widespread applications in security systems, criminal identification, Human Computer Interaction. Recognizing faces was the main motivation behind the development of eigenfaces. Deploying the concept of PCA and eigenface one can achieve significant improvement in the complexity of the space and time requirements. Using this, one can represent many subjects by a small set of data. One might argue that it may lead to loss of information but the benefits achieved far supersedes. Also, systems which uses face recognition need not achieve perfect accuracy. So the eigenface approach works decently on a majority of real life applications under controlled conditions. On top of that, this approach is relatively simple, fast and reliable compared to other approaches proposed in the literature. Again, this approach faces challenges when the dataset is not well illuminated, have non uniform pose or does not have enough number of samples for every subject. The system performs poorly in these cases.

REFERENCES

M. Turk, A. Pentland, “Face Recognition using eigenfaces,” Journal of Cognitive Neuroscience, 1991, pp. 71-86.

Bruce, “Recogning faces”, Journal of Neurophysiology, 1981, pp.46, 369-384.

S. Carey, R. Diamond, “From piecewise to configurational representation of faces”, Science, 1977, pp 195, 312-313.