CSE471: MINI PROJECT - FACE RECOGNITION USING EIGEN FACES

KIRAN RAJ S R
201307593
kiran.raj@research.iiit.ac.in

RAJESH KUMAR
201307556
rajesh.kumar@research.iiit.ac.in

***Abstract*—This mini project aims towards implementing face recognition systems using techniques like PCA (Principal Component Analysis), SVM (Support Vector Machines). We have implemented the above techniques on standard datasets maintained by Yale University, CMU University and also on a realtime dataset generated during CSE471(Statistical Methods in AI) class. We have performed validation, verification experiments on these dataset with the above mentioned techniques and results, inferences are recorded. Also in this paper, we have discussed the interesting observation made while carrying out these experiments.**

I. Introduction

Face recognition systems are used for automatically identifying or verifying a person from the digital image. Face recognition has wide variety of applications ranging from attendance system to security systems. As technology is evolving the quality of the digital image being captured is increasing by manifolds. Thus there is a real need to develop algorithms that can make use of high-resolution digital images to accurately recognize a person.

A. Related Work

Previous works in face recognition systems try to detect features like size, position and relationship among the features from the image and develop a model of face. Using the model developed recognition task is carried out. The major problem in these approaches is, it ignores the aspect of the face that are important for classification by machines. In simple words, all the features that human use to classify/recognize the individuals may not be highly useful for computers.

A good example for one of the previous work which uses extracting features from face was developed by Bledsoe[1]. It is a semi-automated recognition system in which, humans manually mark the features like nose, lips, ears, eyes on the image using hand. Features are now extracted on which standard pattern recognition techniques are applied.

Fischler & Elschlager[2] attemped to measure similar features automatically from images using template matching approach. Improvements on these template matching approach are proposed, also many geometric, statistical appraoches for face recognition have evolved in course of time but they all had the same drawback mentioned in the beginning of this section.

B. Eigen Faces for Recognition

M. Turk and A. Pentland[3] formalized the face recognition problem using Information theory approach [Enoding and decoding is the crux of their proposal]. Following are the steps involved in this approach.

1) Extracting relevant information from face image [*capture variation independent of judgement features*]
2) Encode as efficiently as possible
3) Finally compare encoded sample with similarly encoded samples in database

Encoded images are referred to as eigen faces in this approach. Eigen faces for face recognition is motivated by the PCA (Principal Component Analysis) technique that efficiently represents the images in encoded space. Implementation details of this method is discussed in detail in Section III A.

C. Organization of Report

The rest of the paper is organized as follows. In Section II, we briefly discuss the different datasets on which the experiments were conducted. Section III describes the implementation of various face recognition techniques and the results obtained. We end with some interesting observations, inferences in Section IV.

II. Datasets

A. YALE Dataset

The standard dataset maintained by YALE university has 50-60 images for each person. We have selected exactly 20 images for each person with azimuth and elevation angle in the range [-35, 35] to form the dataset. Other details are as follows,

- Number of Persons/Classes: 38
- Number of Images for each Person: 20
- Total Number of Images: 760
- Size of Each Image: 168 * 192 Pixels
- Training Set Size: 75% = 570
- Testing Set Size: 25% = 190
- Number of Folds = 4

Each partition i.e. training and testing set has images of all the persons in almost equal proportion. Experiments are conducted on 4 different testing set that are captured and average results are displayed.

B. CMU-PIE Dataset

CMU Dataset has images of different persons captured with different Pose, Illumination, and Expression unlike the YALE dataset that has images captured in a constrained environment.

- Number of Persons/Classes: 68
- Number of Images for each Person: 42
- Total Number of Images: 2856
- Size of Each Image: 32 * 32 Pixels
- Training Set Size: 75% = 2176
- Testing Set Size: 25% = 680
- Number of Folds = 4

Validation and verification experiments are done using 4-fold cross validation methods on the 4 different partitions obtained, average accuracies are reported.

C. S13SD Dataset

S13SD is a realtime dataset obtained in *CSE471 class proceedings at International Institute of Information Technology, Hyderabad.* The dataset has non-uniform number of images of different persons captured in different pose, illumination, expression, environment. Persons who have images less than 5 were ignored.

- Number of Persons/Classes: 76
- Number of Images for each Person: Atleast 5
- Total Number of Images: 483
- Size of Each Image: 80 * 80 Pixels
- Training Set Size: 482
- Testing Set Size: 1
- Number of Folds = 483

Hold-one validation technique has been used to report accuracy, where one image is used as test set while the rest of the data is used as training set.

III. Experiments and Results

A. Implementation Details

PCA exploits the fact that images of faces, being similar in overall configuration will not be randomly distributed in this huge image space, and thus can be described by a relatively low dimensional space.This low dimension space can be referred to as eigen space is obtained using the principal components.

Inorder to estimate the principal components, we need to compute the eigen vectors of scatter matrix of the training samples. Top k eigen vectors represent the image efficiently in the lower dimensional space where classification/recognition can be done accurately. Following are the steps involved in transformation of images in original space to low dimensional space.

1) Compute Eigen Faces: Consider a training set with n images each of size m*m pixels. Each image can be represented using a vector of size m^2*1 i.e. each pixel is used as feature and this is a very huge dimension usually for images. The vectors are merged horizontally to form the matrix of size m^2*n. Each column in the matrix now represents an image. Subtract mean image (average of all the training images) from each column to obtain the mean subtracted matrix(**A**) of the training samples.

Scatter matrix(**B**) is given by, AA^T. Computing eigen values of Scatter matrix(**B**) is a hard task as the size of the B matrix is n^2 * n^2 and this is very huge.

Consider the eigen vectors v_i of A^TA such that,

$$A^T A v_i = \mu v_i \tag{1}$$

Premultiply (1) with A on both sides,

$$A A^T A v_i = \mu A v_i \tag{2}$$

We can clearly see that Av_i is the eigen vectors of AA^T. Compute the top k eigen vectors u_1, u_2, u_3, ...u_k which are the eigen faces or principal components of the training samples.

2) Transformation to Eigen Space: Images both from training set and testing set can be now transformed by projecting the original image into eigen space formed by the top k eigen vectors. Multiplying a mean subtracted image with eigen vector results in a scalar value. Therefore multiplying the image with k eigen vectors results in a vector of dimension k*1. k is very low compared to n^2. Hence traditional kNN(k-Nearest Neighbours), SVM or any standard PR technique can be applied for classification and recognition tasks.

B. Identification

Identification experiment refers to the task of identifying the class label of unknown test image. This experiment has been carried out for YALE dataset using kNN & SVM technique while only kNN technique was used on CMU PIE and S13SD dataset. Also all these experiment were carried out using 4-fold cross validation technique i.e. 4 different partitions of same dataset is created and average accuracies are reported.

For all these experiments, dataset is partitioned and transformed into eigen space of k dimension. Observation of accuracy with respect to different eigen space dimension k is also presented.

1) YALE: KNN vs. SVM: Both kNN & SVM, PR techniques were applied on the YALE dataset. kNN uses k = 1 and SVM uses linear kernel function ”-t 0”. The accuracies didn't differ by much percentage. For low eigen space dimension i.e k ¡ 10, SVM outperformed kNN, however for higher eigen space dimensions both performed almost same.

- It is clearly seen from the Fig(1), that the accuracies are almost 100% using both the techniques kNN and SVM. High accuracies are achieved as the dataset doesn't introduce noise by varying illumination, pose or expression. Hence the dataset is highly efficient.
- SVM is fast compared to kNN and hence SVM is preferrable.

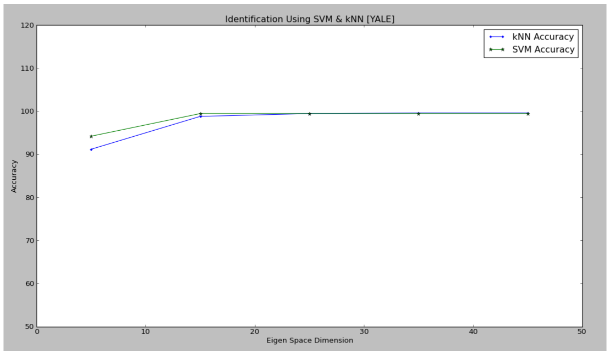

Figure 1: Comparision of accuracies of Identification experiments on Yale Dataset using kNN and SVM by 4-fold cross validation technique.

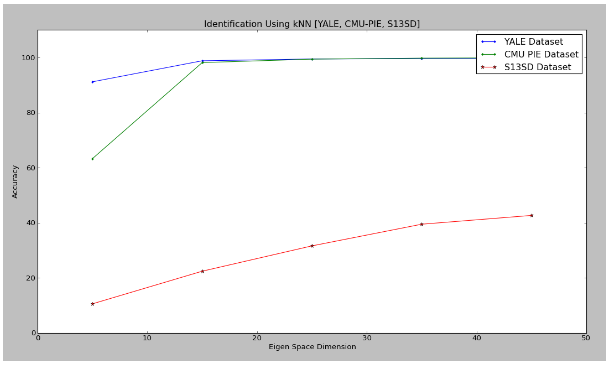

Figure 2: Comparision of accuracies of Identification experiments on YALE, CMU Dataset & S13SD using kNN by 4-fold cross validation technique.

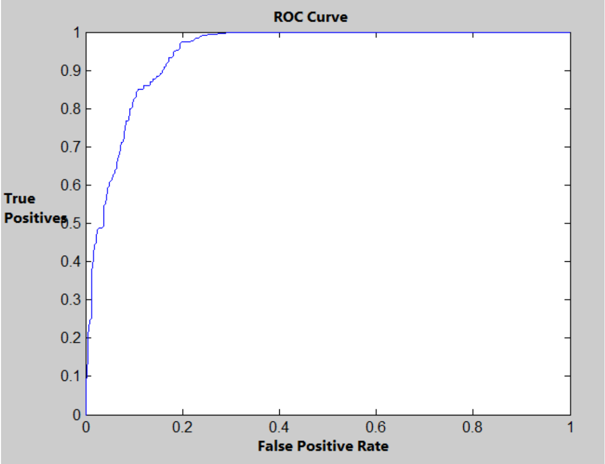

Figure 3: One of the ROC Curve obtained on Yale Dataset inorder to compute threshold at which the module operates for verification. k = 45

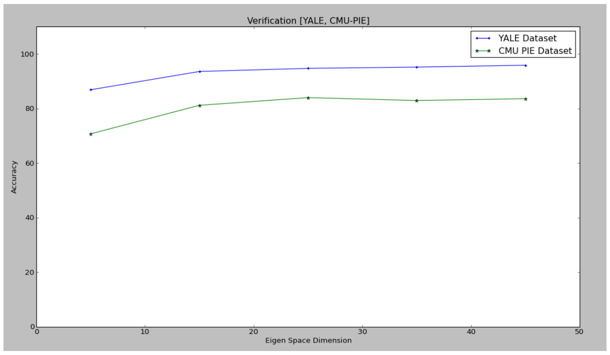

Figure 4: Comparision of accuracies of Verification experiments on YALE, CMU Dataset using kNN by 4-fold cross validation technique.

2) YALE vs. CMU-PIE vs. S13SD: Identification experiments on all three datasets clearly show how accuracy vary according to the dataset (See Fig(2)).

- Higher accuracies are achieved using the dataset that contains training set with not much changes in pose, illumination, expression. While real-time datasets like CMU-PIE and S13SD couldn't achieve high accuracy with less eigen faces.
- In all the datasets the accuracy increases with eigen space dimension.
- In all the cases, top 2 eigen vectors were not used as they were not beneficial for classification purpose. But they hold huge amount of information hence applications like compression can use the top 2 eigen vectors to the best.

C. Verification

Verfication refers to the task of verifying whether the claim of identity for a particular image is true or false. For example, if X's image is given as input and user claims it to be label ”A”, verification module returns whether the claim is TRUE(if x is A) or FALSE(if x is not A).

Verification takes the image, claim label as input. Now the samples from the claimed label are fetched and average distance between the claimed class and the input image is computed. If this distance is less than the threshold set by us the claim is valid else it is not. ROC curve (refer Fig(3)) is drawn inorder to identify the threshold at which the verfication module operates.

With the help of threshold value obtained from above ROC curve the verification experiments are run on possible pairs of test images and different claims. The accuracy of the verification process are reported.

1) YALE vs. CMU: Following are the results, inferences obtained from the verification experiments on YALE, CMU-PIE dataset.

- The Fig(4) clearly indicates that accuracy of verification in case of YALE dataset is high due to the trivial fact that the dataset doesn't contain any noise in the training samples. Noise here refers to change in various parameters of image for the same person.
- Also verification accuracy is lower than classification accuracy as the threshold is not ideal. This is caused due to the low interclass distance between few classes.

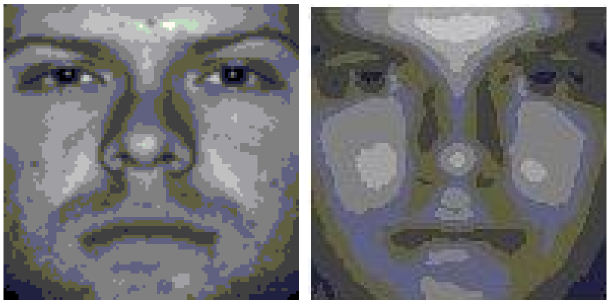

Figure 5: Left hand side image shows the original image and the right hand side image is reconstructed after mapping to a lower dimension.

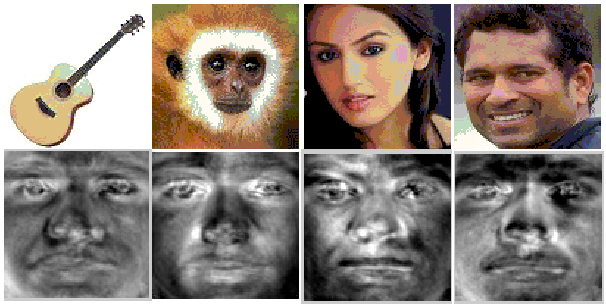

Figure 6: Top row has images in original dimension. First two are non-face images like guitar and a monkey. Next two are unknown images i.e. faces that are not available in training set. Bottom row shows the reconstruction of non-faces and unknown faces.

D. Reconstruction

Reconstruction refers to the task of mapping the image in eigen space dimension i.e lower dimension back to the original space. Consider an image in lower dimension space with k-features. Inorder to re-map it to original space multiply each feature with the corresponding eigen vector and sum it up. Also add back the mean vector subtracted from the image before mapping. Reshape this column vector to a square matrix to get back the image in original dimension.

Fig(5) shows an image in original space in left hand side. It was transformed to eigen space dimension with k = 45 and the image is reconstructed back to original dimension and is shown in right hand side of the same image.

1) Non-face & Unknown Faces: We conducted few experiments on non faces, unknown faces i.e faces not in training set. When mapped them to lower dimension space and reconstructed them back, refer Fig(6) which has top row that shows the original image and bottom row shows the reconstructed image from a lower dimensional space.

- Even non faces are reconstructed as face images due to the mean vector that contains lot of facial information of all the training images.
- Unknown images are reconstructed back but with lot less accuracy as the information pertaining to their images were never captured in the eigen space.

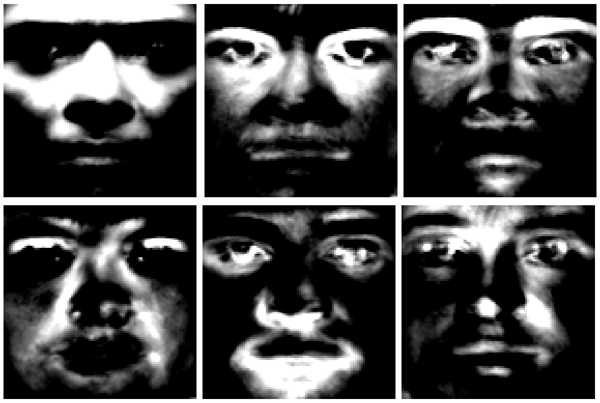

Figure 7: Top 6 eigen faces obtained from the eigen vectors of the scatter matrix of the YALE dataset's training set.

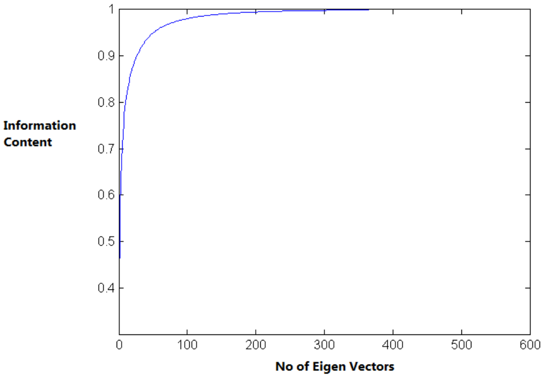

Figure 8: Curve that shows the information content of top k eigen vectors.

- Reproduction is unfaithful in case of non-faces and unknown faces. However reconstruction of known faces are highly accurate.

IV. Observations

A. Eigen Faces

Eigen Faces i.e. eigen vectors of the scatter matrix of the training samples also resemble like image. Top 6-Eigen faces are shown in Fig(7)

B. Information Content in Eigen Vectors

Very few eigen vectors are sufficient to represent the entire information content of an image in lower dimension. Fig(8) shows clearly that for k approximately equal to 10 around 95% of the information are already contained in these eigen vectors,

References

[1] Bledsoe, W.W, *Man-Machine facial recognition*, Panoramic Research Inc., Palo Alto, CA. *1966.*

[2] Fischler M.A, Elschlager R.A, *The representation and matching of pictorial structures*, IEEE Transaction os Computers, *1973.*

[3] M. Turk and A. Pentland, *Eigen Faces for Recognition*, Vision and Modeling group, The Media Laboratory, Massachusetts Institute of Technology. *Journal of Cognitive Neuroscience, Volume 3, Number 1, 1991.*