**PES University, Bangalore**

(Established under Karnataka Act No. 16 of 2013)

**MAY 2020: IN SEMESTER ASSESSMENT (ISA) B.TECH. IV SEMESTER**

**UE18MA251- LINEAR ALGEBRA**

MINI PROJECT REPORT

ON

FACE RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS

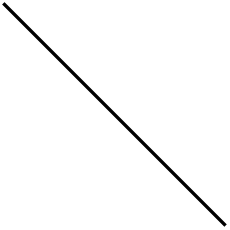
Submitted by

1. Name: CHETHAN M SRN PES2201800331

2. Name: ANIRUDH R SRN PES2201800068

3. Name: G SAI CHARAN REDDY SRN PES2201800483

Branch & Section : CSE (SECTION-A)



PROJECT EVALUATION

( For Official Use Only )

|  |  |  |  |
| --- | --- | --- | --- |
| Sl.No. | Parameter | Max Marks | Marks Awarded |
| 1 | Background & Framing of the problem | 4 |  |
| 2 | Approach and Solution | 4 |  |
| 3 | References | 4 |  |
| 4 | Clarity of the concepts & Creativity | 4 |  |
| 5 | Choice of examples and understanding of the topic | 4 |  |
| 6 | Presentation of the work | 5 |  |
|  | Total | 25 |  |

Name of the Course Instructor : Dr.Girish.V.R. Signature of the Course Instructor :

**INTRODUCTION**

**‘It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment’**

* **Carl Friedrich Gauss**

Security and authentication of a person is a crucial part of any industry. Face recognition is one of the most popular method used. Also it is a key factor in crime detection mainly to identify criminals.

Current system of face recognition will work precisely only if there is an exact match between the trained and tested image (face), which makes it difficult in authentication.

The system proposed in this literature survey is based on the fundamental concept of Linear Algebra, the **Principal Component Analysis (PCA).** Based on this concept, a new breakthrough concept of face recognition in the security industry is introduced called the **Eigenfaces.**

**REVIEW OF LITERATURE**

Enormous number of literature surveys have been published regarding the limitations of the current face recognition system used [1]. Some are listed below:

1. The existing system cannot tolerate variations in the new face. It requires new image to be almost exactly matching with one of the images in the database which will otherwise result in denial of access for the individual.
2. The performance level of the existing system is not appreciable.

Also there are many theories developed for face recognition, some being just mathematical formulas and some being just concept explanations but none were correlated and structured properly to give an accurate and fine system with a programming code for Face recognition and authentication.

A **Kernel principal component analysis (PCA)** [3] was recently proposed as a nonlinear extension of a PCA**.** This letter adopts the kernel PCA as a mechanism for extracting facial features. Through adopting a polynomial kernel, the principal components can be computed within the space spanned by high-order correlations of input pixels making up a facial image, thereby producing a good performance.

But the problem with Kernel PCA is the high amount of error rates that occur during data manipulation and computation of Eigenfaces.

A novel subspace method called **diagonal principal component analysis (DiaPCA) [2]** was proposed for face recognition. In contrast to standard PCA, DiaPCA directly seeks the optimal projective vectors from ***diagonal face images***without image-to-vector transformation.

The drawbacks with DiaPCA are high error rates with **Covariance matrices [4]** and less accuracy in the results if there is a change in facial expressions of the test image and train set.

**REPORT ON THE PRESENT INVESTIGATION**

**Concepts Used:**

**PRINCIPAL COMPONENT ANALYSIS (PCA) [6]**

PCA is a mathematical procedure that uses n orthogonal transformation to convert a set of values of possibly correlated M variables into a set of values of K uncorrelated variables (K<=M) called Principal Components (PC) or Eigenfaces.

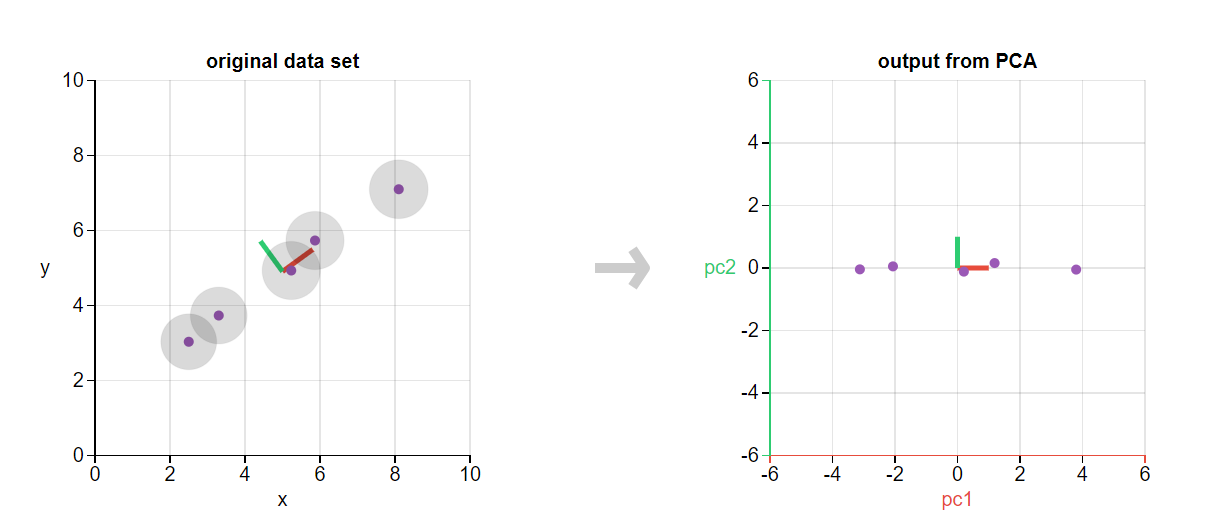
PCA can supply the user with a lower-dimensional picture, a shadow of this object when viewed from its most informative viewpoint.

PCA takes 4 or more variable (i.e. 4 or more dimensions of data) and make a 2D PCA plot.

PCA also tells us which variable is the most valuable for clustering the train dataset when plotted.

With reference to Linear Algebra point of view, PCA finds the best fitting line by maximizing the sum of the squared distances from the projected points to the origin by using Pythagoras theorem.

This transformation is defined in such a way that the first principal component (eigenface) shows the most dominant direction (features) of the dataset and each succeeding component in turn shows the next most possible dominant direction under the constraint that it can be uncorrelated to the preceding components.



**EIGENFACES [1]**

**Eigenfaces** refers to an appearance-based approach to face recognition that seeks to capture the variation in a collection of face images and use this information to encode and compare images of individual faces in a holistic (as opposed to a parts-based or feature-based) manner. Specifically, the eigenfaces are the principal components of a distribution of faces, or equivalently, the eigenvectors of the [covariance](http://www.scholarpedia.org/article/Covariance) matrix of the set of face images, where an image with *N* pixels is considered a point (or vector) in *N*-dimensional space.



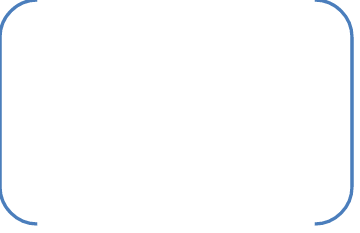
If one uses all the eigenfaces extracted from the original images, one can reconstruct the original images from the eigenfaces exactly. But one can use only a part of eigenfaces, so the reconstructed image is only an approximation of the original image.

These eigenvectors are derived from the covariance matrix of the probability distribution of the high dimensional vector space of possible faces of human beings and hence Eigenfaces are a set of Eigen vectors.

Eigen value of a square matrix is a scalar that is usually represented by λ and an Eigen vector (non-zero) denoted by x.

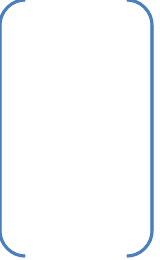
**Ax = λx**

Example:



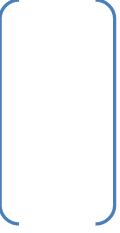
A = 1 2

1. 2



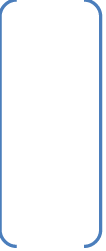
x = -2

3



Ax = 4

-6



Ax = (-2) -2

3

Ax = (-2) x

λ = -2

Hence, -2 is the Eigen value of A and vector X is called the Eigen vector of the matrix A.

**Properties of Eigen vectors**:

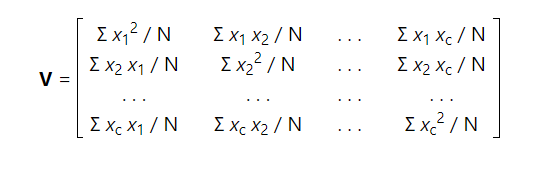
1. They can be determined for square matrices only
2. There are n eigen vectors in an nxn matrix
3. All eigen vectors are perpendicular to each other
4. x and λ are referred to as eigen pairs
5. Eigen space is a space consisting of all eigen vectors which have the same eigen value

**COVARIANCE MATRIX [5]**

**Variance** is a measure of the variability or spread in a set of data. Mathematically, it is the average squared deviation from the mean score.

**Covariance** is a measure of the extent to which corresponding elements from two sets of ordered data move in the same direction.

Variance and covariance are often displayed together in a variance-covariance [matrix](https://stattrek.com/help/glossary.aspx?Target=Matrix), (aka, a covariance matrix). The variances appear along the diagonal and covariance appear in the off-diagonal elements

****

**FACE RECOGNITION USING PRINCIPAL COMPONENT ANALYSIS [4]**

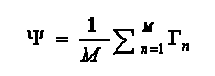
**STEP -1: PREPARE THE DATA**

The first step is to obtain a set S with M face images. Each image is transformed into a vector of size N and placed into the set.



**STEP-2: OBTAIN THE MEAN**

After obtaining the set, the mean image Ψ has to be obtained as,



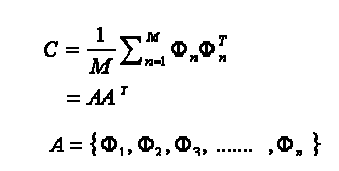
**STEP-3: SUBTRACT THE MEAN FROM ORIGINAL IMAGE**

The difference between the input image and the mean image has to be calculated and the result is stored in Φ



**STEP-4: CALCULATE THE COVARIANCE MATRIX**

The covariance matrix C is calculated in the following manner



**STEP-5: CALCULATE THE EIGENVECTORS**

In this step, the eigenvectors (eigenfaces) ui and the corresponding eigenvalues λi should be calculated. From M eigenvectors, u, only M' should be chosen, which have the highest eigenvalues. The higher the eigenvalue, the more characteristic features of a facedoes the particular eigenvector describe. Eigenfaces with low eigenvalues can be omitted, as they explain only a small part of the characteristic features of the faces. After M' eigenfaces are determined, the “training” phase of the algorithm is finished. Once the training set has been prepared the next phase is the classification of new input faces.

**STEP-6: RECOGNITION PHASE (Transform the New Face)**

The new face is transformed into its eigenface components and the resulting weights form the weight vectors.



**TESTING THE MODEL IN PYTHON**

face\_vector = []

for i in range(total\_images):  
 face\_image=cv2.cvtColor(cv2.imread(path),cv2.COLOR\_RGB2GRAY)  
 face\_image = face\_image.reshape(total\_pixels,)  
 face\_vector.append(face\_image)

face\_vector = np.asarray(face\_vector)  
face\_vector = face\_vector.transpose()

avg\_face\_vector = face\_vector.mean(axis=1)  
avg\_face\_vector = avg\_face\_vector.reshape(face\_vector.shape[0], 1)  
normalized\_face\_vector = face\_vector - avg\_face\_vector

covariance\_matrix = np.cov(np.transpose(normalized\_face\_vector))

eigen\_values, eigen\_vectors = np.linalg.eig(covariance\_matrix)

eigen\_vectors = sort(eigen\_vectors)  
k\_eigen\_vectors = eigen\_vectors[0:k, :]

eigen\_faces = k\_eigen\_vectors.dot(normalized\_face\_vector.T)

weights = (normalized\_face\_vector.T).dot(eigen\_faces.T)

test\_img = test\_img.reshape(total\_pixels, 1)

test\_normalized\_face\_vector = test\_img - avg\_face\_vector

test\_weight = (test\_normalized\_face\_vector.T).dot(eigen\_faces.T)

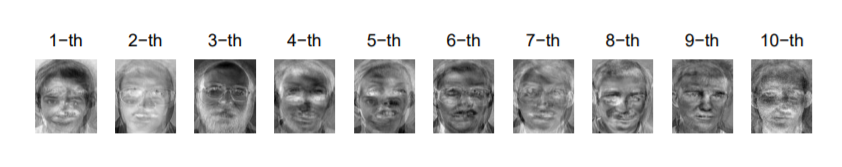
index = np.argmin(np.linalg.norm(test\_weight - weights, axis=1))

**RESULTS AND DISCUSSIONS**

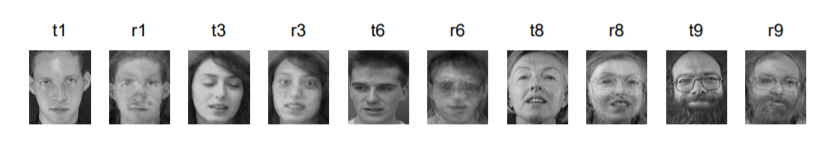
**Original training images**



**Eigenfaces**



**Test Images**



ALL THE TEST IMAGES WERE RECONSTRUCTED USING EIGENFACES AND WERE SUCCESSFULLY RECOGNIZED.

The eigenfaces produced during train model construction are the reshaped eigenvectors produced during the process of Principal Component Analysis.

Hence, the model demonstrated here gives the most accurate results and very less error rates with respect to the eigenface values.

The model also rejects the non-face test images passed since they can’t be constructed with the eigenfaces that were created using train image set.

**SUMMARY AND CONCLUSIONS**

Thus the proposed face recognition system based on PCA has been implemented. It accurately identifies input face images of an individual which differ from the set of images of that person already stored in the database thus serving as an effective method of recognizing new face images. The base code for training face images using Back Propagation Neural Network has also been completed. Hence when a new image is fed into the system for recognition the main features are extracted and computed to find the distance between the input image and the stored images.

Thus, some variations in the new face image to be recognized can be tolerated. When the new image of a person differs from the images of that person stored in the database, the system will be able to recognize the new face and identify who the person is. Recognition accuracy and better discriminatory power Computational cost because smaller images (main features) require less processing to train the PCA. Because of the use of dominant features and hence can be used as an effective means of authentication.

**SCOPE FOR FUTURE WORK**

Over the years, movies have fixed a futuristic fantasy in our minds that a time will come when software would be used to recognize people by their faces. A time when our faces will be our ID cards. With advent of facial recognition technology, that time is already here.

Today, along with drones, AI and IoT, facial recognition technology is also defining our millennium. Facial recognition is a biometric technology used for authentication and examination of individuals by correlating the facial features from an image with the stored facial database.

Face Recognition is one of the most popular applications of image analysis software and no more considered as a subject of science fiction. Earlier, this technology was only used for security and surveillance purposes, but it has safely transitioned to the real world in recent times. Today, companies are pitching facial recognition software as the future of everything from retail to policing.

**BIBLIOGRAPHY**

1. 3D Face recognition using 3D-eigenfaces [textbook]

Authors:

1. Brijesh Mehta

2. Divyarajsinh Parmar

3. Yagnish Parmar

Publication year: July 6, 2013

Publisher: LAMBERT

ISBN: 9783845407197

2. Diagonal Principal Component Analysis for Face Recognition [literature review]

Authors:

1. Daoqiang Zhang

2. Songcan Chen

3. Zhi-Hua Zhou

Nanjing University, Nanjing 210093, China

3. Face Recognition Using Kernel Principal Component Analysis [literature review]

Authors:

1. Hang Joon Kim
2. Keechul Jung
3. Kwang In Kim

Publication year: February 2002

4. Face recognition using principal component analysis and neural networks [literature review]

Authors:

1. A.S. Syed Navaz
2. Pratap Mazumder
3. T. Dhevi Sri

Publication year: Mar 2013

ISSN: 2250-1568

5. High dimensional covariance estimation [textbook]

Author: Mohsen Pourahmadi

Publication: 24 june 2013

ISBN: 9781118034293

Publisher: WILLEY

Reference section: Chapter 3

6. Principal Component Analysis [textbook]

Author: I. J. Jolliffe

Edition: Second

Publisher: SPRINGER

ISBN: 0387954422

Reference section: Chapter 1