



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 4, April 2023

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# Comparison of Linear and Kernel-Based Methods for Face Recognition

Poonam V. Kadam

Assistant Professor, Department of Computer Engineering, Metropolitan Institute of Technology & Management (MITM), Sindhudurg, Maharashtra, India

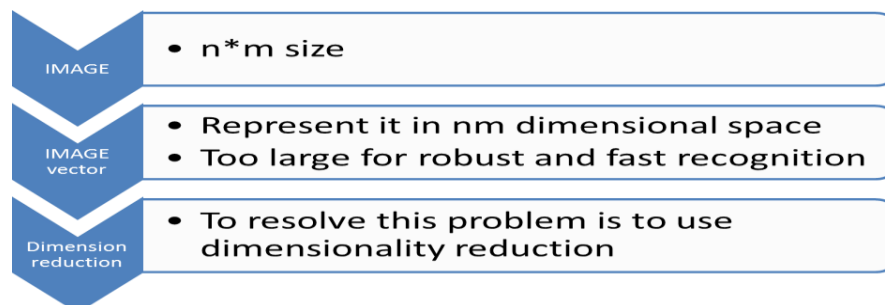
**ABSTRACT:** This paper presents the results of a comparative study of linear and kernel-based methods for face recognition. The methods used for dimensionality reduction follow linear or Kernel approach. Linear methods are Principal Component Analysis (PCA) & Linear Discriminant Analysis (LDA). Kernel based methods used for face recognition are Kernel Principal Component Analysis (KPCA) and Kernel Discriminant Analysis (KDA). The method used for classification is Euclidean Distance. In addition, these classification methods are applied on raw images to measure the performance of these dimensionality reduction techniques. All experiments have been performed on images from ORL Face Database.

**KEYWORDS:** Face recognition, Principal component analysis, Linear discriminant analysis, Kernel Methods, KPCA, KDA.

## I. INTRODUCTION

Face recognition is a formal method which is first proposed by Francis Galton in 1888. As one of the most successful applications of image analysis and understanding, face recognition has significant demand especially during the past two years, Face recognition can be performed in both a still image and video, which has its origin in still image face recognition. Different approaches of face recognition for still images can be categorized into two main groups such as

- **Local feature based methods:** In the local feature-based approach, typically a set of features is extracted from the image (e.g. locations of facial features such as nose, eyes etc.) and used to classify the face. The main advantage of this approach is its relative robustness to variations in illumination, contrast, and small amounts of out-of-plane rotation. Its main disadvantage is that there is generally no reliable method to extract an optimal set of features; hence useful information may be lost in the feature extraction step.
- **Global approach:** This method uses the whole face as the raw input to the recognition system. It is also known as HOLISTIC approach. Global approaches use the entire image as the pattern to be classified, so no information is lost. In this project we will concentrate on the global approach.



When using global methods, we usually represent an image of size  $n \times m$  pixels as a vector in  $nm$  dimensional space. In practice, however, the  $nm$ -dimensional space is often too large to allow robust and fast recognition. A common way to

resolve this problem is to use dimensionality reduction techniques. Two of the most effective techniques for this purpose are PCA and LDA.

## II. LITERATURE SURVEY

### PCA (Principle component analysis)

A principal component analysis (PCA) is a powerful technique for extracting a structure from potentially high-dimensional data sets, which corresponds to extracting the eigenvectors that are associated with the largest Eigen values from the input distribution. This eigenvector analysis has already been widely used in face processing. PCA of a high temporal resolution image series can attenuate temporal autocorrelation, thereby increasing the suitability of the data for image segmentation and classification procedures.

The eigenface approach applies PCA transform for feature extraction. One of the goals that the feature extraction routine wishes to achieve is to increase the efficiency. One simple way to achieve this goal is using alternative orthonormal bases other than the natural bases. One such basis is the Karhonen-Loeve (KL). KL bases are formed by the eigenvectors of covariance matrix of the face vector. In the high dimensional “face” space, only the first few eigenvalues have large values. in other words , energy mainly locates in the subspace constituted by the first few eigenvectors. Therefore a great compression can be achieved by letting those eigenvectors with large eigenvalues to represent face vector.

### PCA Algorithm

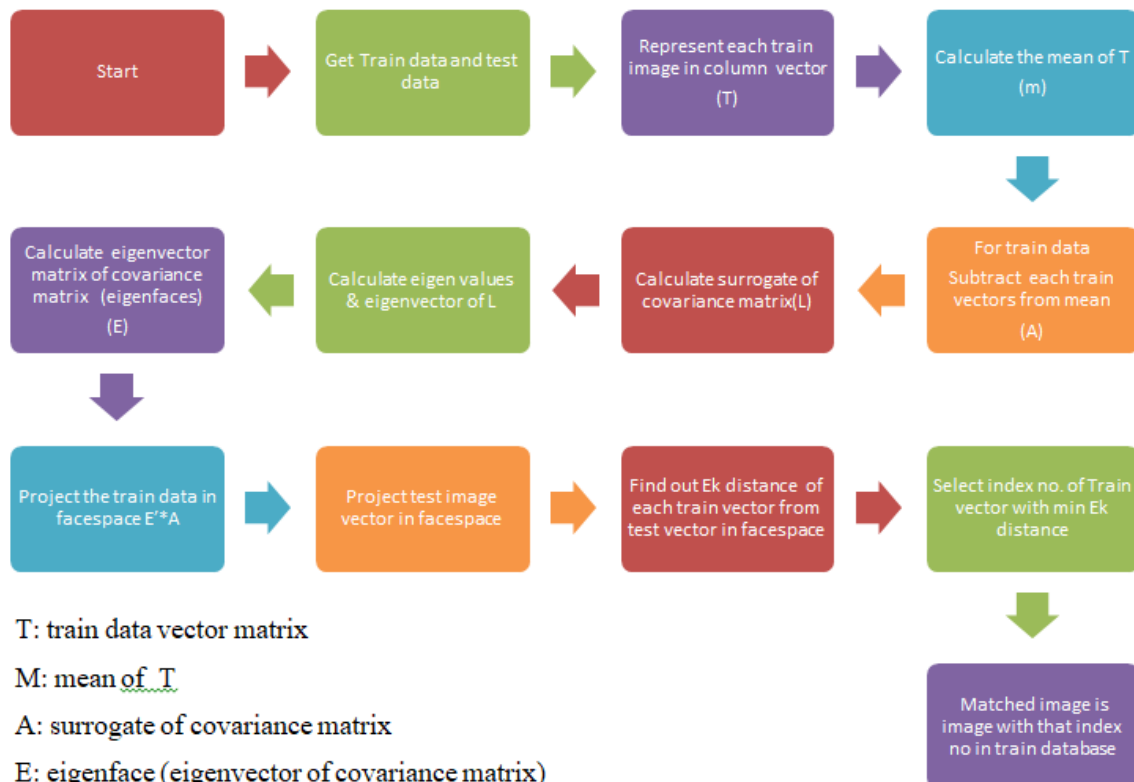


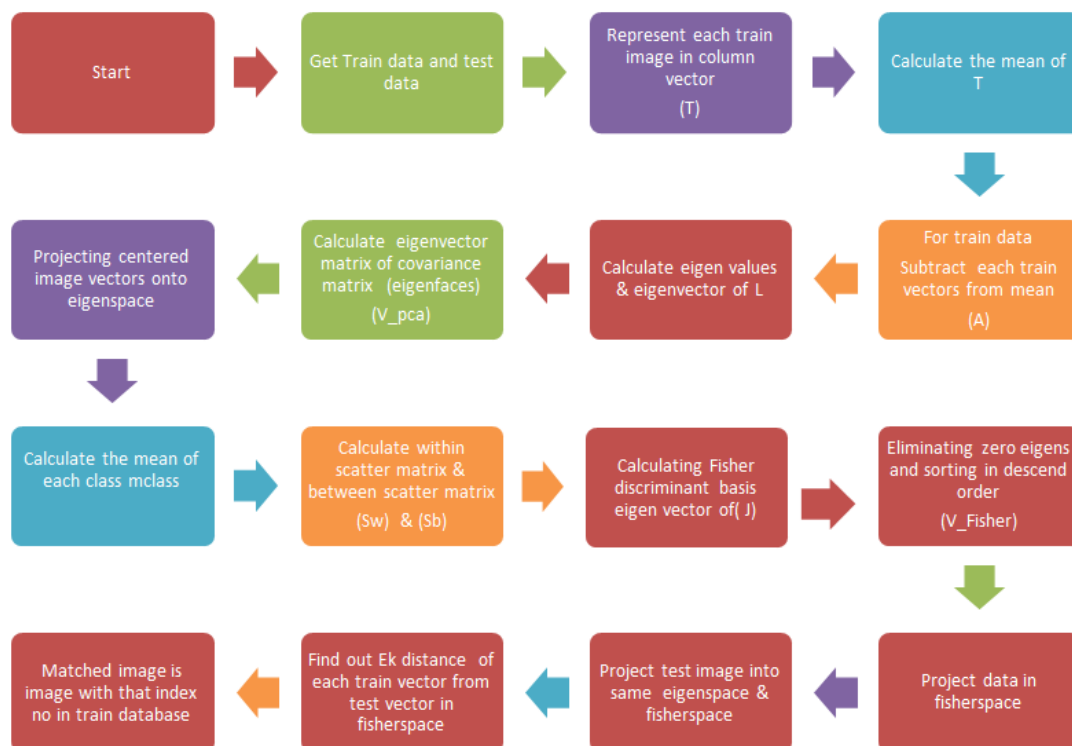
Fig.PCA algorithm flowchart

### LDA (Linear Discriminant Analysis)

The LDA transform for feature extraction reduces the dimensionality of a high dimensional image and project it. Its

peculiarity is that it even takes into consideration scatters within and between classes. LDA based algorithms optimizes the low dimensional representation of the object with the focus on most discriminant feature extraction. This method combines the strength of D-LDA and F-LDA approaches while at the same time overcome their shortcoming and limitations. From the method described above we can say that, these methods makes FRT more advanced. Also significant advances have been made in the design of classifier for successful Face Recognition. Hence, we can say that among Appearance based holistic approaches Eigen faces and Fisher faces have proved to be effective in experiments with large database. Because as compared to holistic approaches, feature based methods are less sensitive to variations in elimination and to inaccuracy.

### LDA Algorithm



- T: train data vector matrix
  - m: mean of T
  - A: surrogate of covariance matrix
  - V\_pca: eigenface (eigenvector of covariance matrix)
  - Sw: within scatter matrix
  - Sb: Between scatter matrix
  - J: Fisher discriminant basis's
  - V\_fisher: fisherspace
  - Ek :Euclidian distance
- Fig. LDA algorithm flowchart

### KERNEL Methods

The recognition technique in well control environment gives better performance. However, Face recognition task cannot be assumed to be linear. face recognition in an uncontrolled environment is still very challenging. These challenges are: The Illumination variation problem, Pose variation problem and occlusion, image background difference these issues can cause serious performance degradation in face recognition. Such problems are known as non-linearities in face recognition



**Illumination:** In Illumination problems face appears different due to a change in lighting. Hence performance of face recognition becomes poor.

**Pose Variance:** The performance of Face Recognition System drops significantly when large pose variations are present in the input images. Reducing this difficulty is not difficult, but accurate pose estimation is hard.

**Occlusion:** Sometimes the face images are partially occluded that is any object comes over the face in the images. And due to this it becomes difficult to recognize a person from an image in the database.

For such non-linear input conventional linear methods such as LDA and PCA perform poorly  
Proposal: Use Kernel approach

**What is KERNEL trick?**

Consider a Dataset  $X$  in high dimensional space  $R^n$

$$[X_1, \dots, X_n] \in R^n,$$

via the nonlinear mapping

$$\Phi : R^n \rightarrow F$$

$$X \rightarrow \Phi(x)$$

$X$  is mapped into a very high (possibly infinite) feature space  $F$ .

we use nonlinear kernel function  $\Phi(x)$  to project input data  $X$  into feature space  $F$  which is nonlinearly related to the input space and performs PCA/LDA in feature space  $F$ .

**Types of kernel function**

Following are type of KERNEL function

- Polynomial  $((x,y)+\theta)^n$   $\theta = \text{constant}$   $n = \text{degree of polynomial}$
- Gaussian RBF  $\text{Exp}(-|x-y|^2/2\sigma)$   $\sigma = \text{variance}$
- Sigmoidal  $\tanh(k(x,y)+\theta)$   $\theta = \text{constant}$
- Inverse multiquadric  $(1/\sqrt{|x-y|^2+c^2})$

According to reference Gaussian RBF and polynomial KERNEL is commonly used KERNEL functions.

Polynomial Kernel perform good for 2<sup>nd</sup> degree whereas Gaussian kernel gives good performance at  $\sigma = 0.26^*$  dimension.

In our project we are using Gaussian RBF as KERNEL function

**KPCA**

Kernel based methods can map the input space to a high (possibly infinite) dimensional feature space in a Non-linear way. The projection into higher space is done using inner products without explicit computation of the mapping, which makes it computationally feasible essentially, The computations are done in a subspace of dimension  $N$  of a higher dimensional space. In contrast to PCA, the dimension of the projected subspace can reach using KPCA. The basic concept of kernel PCA is to first map the input data  $\chi$  (training set) into a feature space  $F$  via a non-linear mapping  $\Phi(\bullet)$  and then perform a linear PCA in  $F$ . The motivation is that a training set, which may not be linearly separable in input space, may be linearly separable in the mapped space.



Again, let  $X = [x_1 | x_2 | \dots | x_N]$  denote the matrix containing the training vectors, and let  $\Phi = [\varphi(x_1) | \dots | \varphi(x_N)]$  be its image in the feature space. Assuming that the mapped data are centered, i.e.,  $\sum_{i=1}^N \varphi(x_i) = 0$  (the centering method in  $F$  is explained at the end of this sub-section). Let  $K = \Phi^T \Phi$  with  $K_{ij} = \varphi(x_i) \cdot \varphi(x_j)$ . Then the principal directions satisfy equation (2) and are given by  $e = \Phi \alpha$ . (After normalization of  $\alpha$ ). For a test vector  $x$ , the principal component  $y$  corresponding to eigenvector  $e$  is given by:

$$y = e^T \varphi(x) = \alpha^T \Phi^T \varphi(x) = \sum_{i=1}^N \alpha \varphi(x_i) \cdot \varphi(x_j).$$

The dot-product matrix  $K$  can be computed by choosing a kernel  $k(x, y)$  such that  $k(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) = K_{ij}$  and thus avoiding any computation in the high dimensional feature space. This is referred to as the “kernel trick”. Then, the first  $q$  principal components (assuming that the eigenvectors are sorted in a descending order of eigenvalues) constitute the  $q$ -dimensional feature vector for a face pattern. In general, the assumption of centered data in feature space made above is not reasonable. A method to center the mapped data is described here: Let

$\tilde{\varphi}(x_j) = \varphi(x_j) - 1/N \sum_i \varphi(x_i), 1 \leq j \leq N$  be the centered mapped data in the feature space.

$\tilde{\Phi} = [\tilde{\varphi}(x_1) | \tilde{\varphi}(x_2) | \dots | \tilde{\varphi}(x_N)] = \Phi (I - 1_{N \times N}/N)$  be the matrix containing the centered mapped training vectors.

Using Gaussian Kernel to construct Kernel  $K$

$$K(x, y) = -e^{-\frac{(x-y)^2}{\sigma^2}}$$

Where  $K$  is symmetric Kernel Then, the inner product matrix  $\tilde{K}$  for the centered mapped data can be obtained from inner product matrix  $K$  of non-centered data by  $\tilde{K} = \tilde{\Phi}^T \tilde{\Phi} = (I - 1_{N \times N}/N)^T K (I - 1_{N \times N}/N)$  Where  $I_{N \times N}$  is an  $N \times N$  identity matrix and  $1_{N \times N}$  is a  $N \times N$  matrix of all ones.

### KDA

As in the case of PCA, we can use the kernel idea for LDA to find non-linear directions by first mapping the data non-linearly into some feature space  $F$  and computing Fisher’s linear discriminant there, thus implicitly yielding a non-linear discriminant in input space. Let  $\Phi(\bullet)$  be a non-linear mapping to some feature space  $F$ .

Let  $\Phi = [\Phi(x_1) | \dots | \Phi(x_N)]$  be the matrix which contains the non-linear mappings of all the training samples. To find the linear discriminant in  $F$  we need to maximize

$$J(w) = \frac{w \cdot w^T S_B^\Phi}{w \cdot w^T S_W^\Phi}$$

Again, this problem can be reduced to an eigenvalue problem of the same form as in LDA. Instead of mapping the data explicitly into the feature space, we seek formulation of the algorithm which uses only the dot products  $\varphi(x_i) \cdot \varphi(x_j)$  of the images of training patterns in feature space. We are then able to compute these dot products efficiently without mapping explicitly to  $F$ . As shown in the LDA case, the cost function in this can be reduced to the form in Equation 7. In this case, the dot-product matrix  $K$  is  $\Phi^T \Phi$ , and can be computed by choosing a kernel of the form  $k(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) = K_{ij}$ . the projection  $y_j$  of a test vector  $x$  corresponding to

$j^{\text{th}}$  Eigenvector is given by

$$y_j = \sum_{i=1}^N k(x, x_i) \alpha_j^i$$

### III. PROPOSED SYSTEM

#### System Architecture

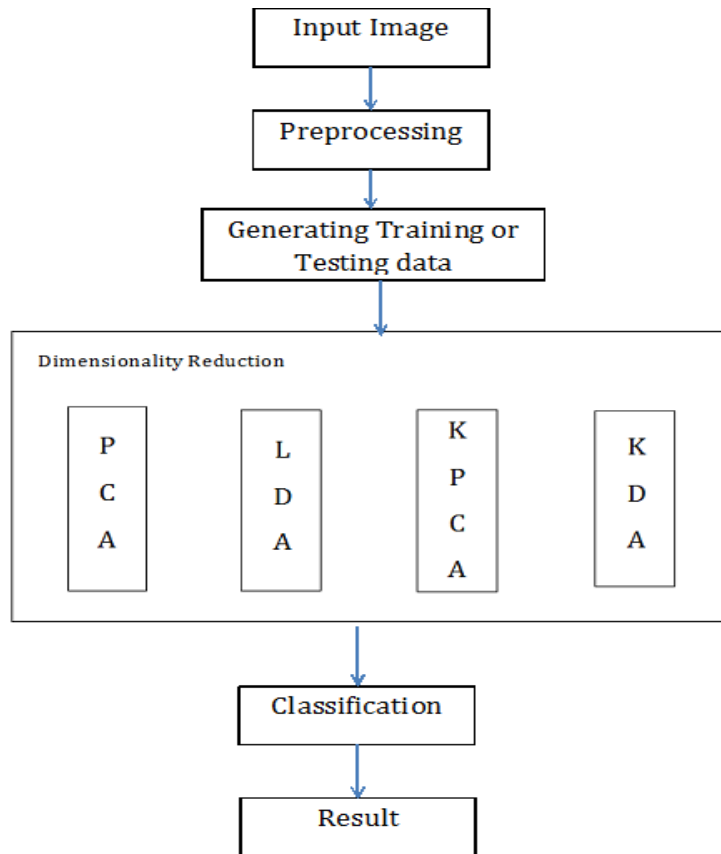


Fig. System Architecture

#### Pre-processing:

The original images were downsampled by a factor of 4 in each dimension to 28x23. The downsampled images were histogram equalized, and normalized to zero mean and unit variance.

#### Generating Training and Test Data

The database has variable number of images for different subjects, but the same number of training and test images for each subject was chosen so that the results are not affected by number of images of any particular subject. Ten runs of each experiment were performed and the results were averaged. In each experiment, nine randomly chosen images for each subject were used in the training set, and nine randomly chosen images for each subject were used in the test set. The test and training set did not have any images in common.

#### Dimensionality Reduction

Dimensionality reduction is achieved by linear and kernel-based methods. Linear methods used are PCA and LDA, and kernel-based methods used are KPCA and KDA.

#### Classification

After global features are extracted using PCA/KPCA or LDA/KDA, classification is performed using two approaches: Nearest Neighbor using the standard L2-norm for Euclidean distance, and Support Vector Machine which is widely used for classification because of certain desirable properties such as provable optimality and low computational cost. Since the classic SVM was formulated for the two-class problem, its basic scheme is extended to multi-class face recognition by adopting one-versus-rest decomposition. This works by constructing c SVMs for a c-class problem.



Each SVM is trained to separate its class from all the other classes and then an arbitrator is used between the SVM outputs to produce the final result.

#### IV. EXPERIMENTALRESULT

##### ERROR RATE USING 2 IMAGES PER INDIVIDUAL

EXP/ ALGORITHM	PCA	KPCA	LDA	KLDA
EXPERIMENT NO.1	6.66%	0%	66.67%	40%
EXPERIMENT NO.2	20%	13.33%	86.66%	46.66%

##### ERROR RATE USING 3 IMAGES PER INDIVIDUAL

EXP/ ALGORITHM	PCA	KPCA	LDA	KLDA
EXPERIMENT NO.3	6.66%	6.66%	60%	33.34%
EXPERIMENT NO.4	0%	0%	60%	13.34%

#### V. CONCLUSION

According to experimental result, Both kernel based methods perform better than standard eigenface (PCA), fisherface (LDA) method. For 2 images per individual the KPCA improves its performance reducing the error rate, compare to 3 images; so we can say that for PCA method kernel approach give better result for less no image per individual. In case of LDA it is different case in 2 images and 3 images per individual kernel LDA reduces the error rate resulting in improved performance of LDA. We conclude that the KPCA is the best method for face recognition.

#### REFERENCES

- [1] B.S. Manjunath, R. Chellappa, C. von der Malsburg, "A Feature Based Approach to Face Recognition", in Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, pp.373-378, 1992.
- [2] Kwang In Kim, Keechul Jung and Hang Joon Kim, "Face Recognition Using Kernel Principal Component Analysis", IEEE Signal Processing letters, Vol 9, No. 2, Feb 2002.
- [3] Sebastian Mika, Gunnar Rätsch, Jason Weston, Bernhard Schölkopf, Klaus-Robert Müller, "Fisher Discriminant Analysis With Kernels", Neural Networks for Signal Processing IX, 1999
- [4] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller, "Nonlinear Component Analysis as a kernel Eigenvalue Problem", Tübingen, Germany, Tech. Report No 44, Dec 1996.
- [5] Christopher J C Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", Bell Laboratories, Lucent Technologies, 1998.
- [6] T. Joachims, "Making Large-Scale SVM Learning Practical. Advances in Kernel Methods - Support Vector Learning", B. Schölkopf and C. Burges and A. Smola (ed.), MIT-Press, 1999.
- [7] W. Y. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face recognition: A literature survey", UMD CfAR Technical Report CAR-TR-948, 2000.
- [8] M. Turk and A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, 3(1):71-86, 1991.
- [9] K. Etamad and R. Chellappa, "Discriminant Analysis for recognition of human face images", Journal of the Optical Society of America A, 14(8):1724-1733, 1997





Impact Factor: 8.423



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details