A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques

Jamal Hussain Shah, Muhammad Sharif, Mudassar Raza, and Aisha Azeem
Department of Computer Sciences, COMSATS Institute of Information Technology, Pakistan

Abstract: *Face recognition is considered to be one of the most reliable biometric, when security issues are taken into concern. For this, feature extraction becomes a critical problem. Different methods are used for extraction of facial feature which are broadly classified into linear and nonlinear subspaces. Among the linear methods are Linear Discriminant Analysis (LDA), Bayesian Methods (MAP and ML), Discriminative Common Vectors (DCV), Independent Component Analysis (ICA), Tensor faces Multi-Linear Singular Value Decomposition (SVD), Two Dimensional PCA (2DPCA), Two Dimensional LDA (2D-LDA) etc., but Principal Component Analysis (PCA) is considered to be one the classic method in this field. Based on this a brief comparison of PCA family is drawn, of which PCA, Kernel PCA (KPCA), 2DPCA and Two Dimensional Kernel (2DKPCA) are of major concern. Based on literature review recognition performance of PCA family is analyzed using the databases named YALE, YALE-B, ORL and CMU. Concluding remarks about testing criteria set by different authors as listed in literature reveals that K series of PCA produced better results as compared to simple PCA and 2DPCA on the aforementioned datasets.*

Keywords: *Linear, non-linear, PCA, 2DPCA, 2DKPCA, facial features extraction, face recognition, survey.*

Received November 21, 2010; accepted May 24, 2011; published online August 5, 2012

1. Introduction

In pattern recognition feature extraction is the most difficult problem; its reason is that we want to obtain the best features with minimum classification error and low running time. In some image processing techniques, feature extraction [38] is very complicated in some critical biometrics like finger prints [14, 36] where features are so, minute that cannot be seen with the necked eye. But face is one of the most important corporals biometric that shows reliable results as compare to other biometrics [41]. One of the classical dimensional reduction methods namely Principle Component Analysis (PCA) is commonly used in image processing as well as in signal processing methods in particular feature extraction via PCA is used by many authors in the recognition of facial images. Sirovich *et al*. [15, 43], is the one who exploit PCA based classifier in order to represent the images of human faces. Similarly Turk *et al*. [49], used eigenfaces to represent face images. Up till now the aforementioned methods are used for face recognitions systems. Pentland *et al*. [33], on the other hand utilize Modular eigenspaces for pose variation [26]. Sharif *et al.* [39], introduced appearance-based methods for recognition purpose. An additional problem handled by various researchers is the illumination [39] factor to enhance the performance of face recognition. For this various authors i.e., [6, 11, 42, 61, 62] contribute to eliminate the unwanted illumination factor via PCA. Among the subspace reduction methods Linear/Fisher Discriminate Analysis (LDA/FDA) [1, 2, 27, 46, 53] is used for facial feature classification under varying light and pose. Although, LDA extract better discriminate features but its performance is degraded under small sample size problem (3S) [8]. Recently, further enhancements are depicted in classical PCA algorithm resulting in Kernel version of PCA (KPCA) [37], similarly for LDA the advancement have been seen under the name of Kernel LDA (KLDA) [23, 60]. By using higher dimensional space i.e., kernel hilbert space provides more dependency among the input image. One most interesting property of PCA is that it allows the conversion of 2Dimensional face images into 1D feature vector. Each row of the 2D image corresponds to a row in the feature vector. But the problem might arise in the sense that the spatial information of the face image may lose during the conversion, and dimensionality may arise. Moreover, small amount of data is available in images like faces or textual information; in such cases small sample size problem is expected. In order to avoid these problems Two Dimensional PCA (2DPCA) is used. An advantage of 2DPCA over 1DPCA is that the feature vector is now two-dimensional. So, the problem of dimensionality is greatly reduced [29, 58]. Technically speaking the co-variance matrix may be directly constructed via 2D image vector. But the drawback of using 2DPCA is that more coefficients are required in the representation phase of image in the 2DPCA subspace. Also, higher order vector of the image are ignored due to the linearity of 2DPCA.

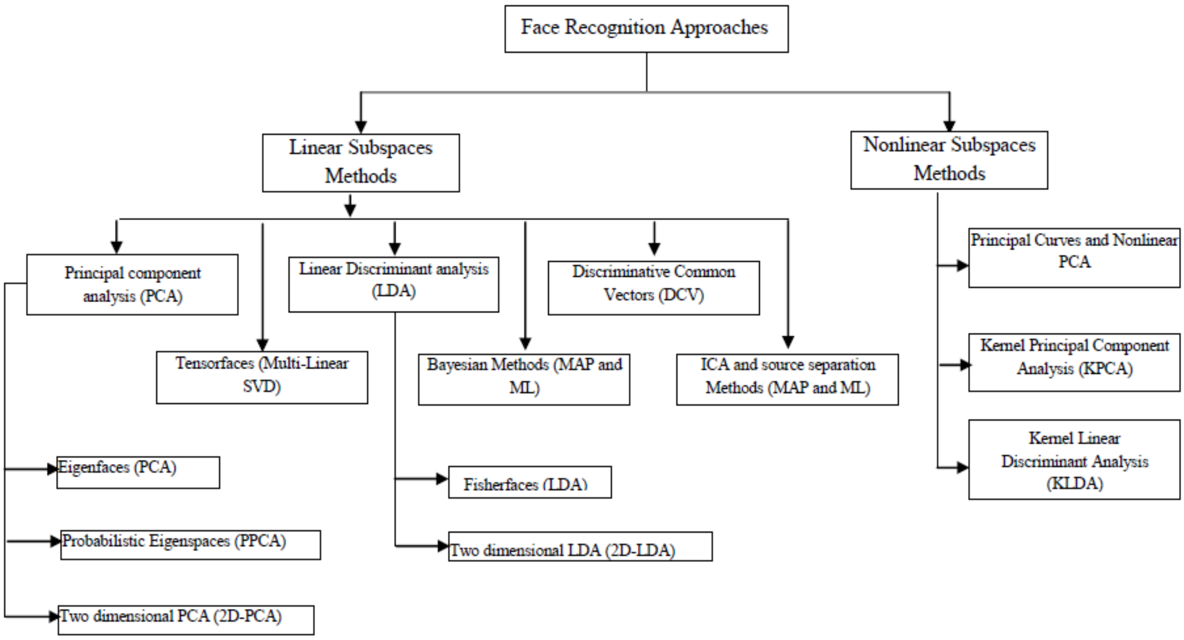

Figure 1. Face recognition methods regarding linear and nonlinear.

However, face images have nonlinear nature, when the face images are projected under uncontrolled laboratory condition like lighting factors [26]. This reduces the recognition rates when using the 2DPCA for recognition. Most commonly used methods regarding linear and nonlinear subspaces and there is represented in Figure 1.

This paper is organized as follows: in section 2 a brief description of selected methods is made and highlighting their pros and cons draws evaluation strategy. They are also, evaluated based on their functionality in controlled and uncontrolled laboratory conditions. In section 3 a comparison among PCA, KPCA, 2DPCA and 2DKPCA, is conducted in which different datasets were used under varying lighting conditions, facial expressions and under complex backgrounds. Their results and their corresponding graph are also, shown for clear evaluation. Finally, conclusion is drawn in section 4.

2. Overview of Linear and Nonlinear

2.1. Subspace Methods

In this section comparison is drawn among various methods related to PCA namely KPCA, 2DPCA, 2DKPCA and last but not the least binary two-dimensional PCA. These algorithms are evaluated on the basis of their performance by comparing their results. After comparing the family of PCA algorithms each method proposed by different researchers are evaluated at end.

2.1.1. Principle Component Analysis (PCA)

PCA, also known as Karhunen-Loeve (KL) transformation or eigenspace is basically a statistical technique used in image recognition and classification. It is also, used for image compression. First time Kirby and Sirovich developed low-dimensional characteristic for face images at Brown University [15, 43]. Which was further elaborated to eigenspace projection by [7, 13, 26, 28]. Recently, Nayar, Nene and Murase used eigenspace projection to view objects at various angles [10, 43]. Finlayson *et al.* [7], used angles and eigenvectors for color images (eigenvectors are also, called eigenfaces).

Main emphasis of PCA is to transform the 2D image into 1D feature vector in subspace. This subspace is also, called eigenspace in which the covariance matrix is obtained as a result of facial features. The subspace formed as a result of PCA conversion makes use of facial feature to characterize different reference images or eigenfaces from the sample dataset. The orthonormal vector obtained as a result of Singular Value Decomposition (SVD) is projected on to a database, which results in a major reduction in a size of the coefficient used to symbolize the image. This property also, helps in reducing the overall analysis time by obtaining the best match among the probe and reference image.

1. *Evaluation:*
   a. PCA converts a high dimensional data into low dimensional image in a linear fashion, in which the principle component are not correlated.
   b. The recognition rate of PCA based face recognition outperforms when the number of test images increases, but the rate of recognition decreases of the certain number.
   c. Size of the image is not superior issue in PCA based system but the important thing is that the number of probes images before PCA projection is large as compared to the number of reference images.

2. *Discussion:* Those authors that have used PCA for face recognition such as: Jose and Gottumukkal *et al.* [9, 34] proposed the local and global feature extraction method by exploiting Modular Principal Component Analysis (MPCA) that incorporates Modular PCA and 2D PCA method. The algorithm deals with N-smaller subdivision of faces in order to increase its recognition rate by taking the benefit of un-harmful regions. Consequently average image is computed from N-sub images. On the whole the results are accomplished on Yale, ORL and UMIST database under the un-controllable conditions of face recognition like pose, illumination and expression variation and attain the results as shown in Table 1.

Table 1. Best recognition rates [34].

	UMIST	ORL	Yale
MPCA	63.00(10/09)	94.44(25/04)	94.44(25/09)

Nicholl and Amira [31], focused on automatically estimating the discriminative coefficients in a DWT/PCA that deals with inter-class and intra class standard deviations. Consequently, eigen-faces are elected on the basis of eigen-values with discrepancy due to the illumination factors between trained images. Recognition rate of such system is shown in Table 2.

Table 2. Comparative results [31].

	Accuracy (%)
DWT/PCA with Coefficient Selection	96.5

Li *et al.* [18], gained an attention to feature extraction through PCA for feature extraction and LDA for classifying features. Similarly, Nearest Neighbor Classifier (NNC) is used for face recognition. Table 3 demonstrates the outcome produced by the approach in ORL database.

Table 3. Recognition rate [18].

	5	10	15	20	25	30	35	40
PCA+ LDA	0.635	0.78	0.845	0.835	0.89	0.895	0.9	0.93

Moon and Philips [25], access and analyze different PCA based algorithms on the basis of comparison with illumination normalization; upshots of compressed images, varied eigen-vectors illustration and alter the similarity measure in feature classification methods.

Chen *et al.* [3], on the other hand proposed Adaptive Principal Component Analysis (APCA) in order to progress the PCA results by functioning PCA on faces, rotate face space and then warping is applied between class and within class covariance.

Zhang *et al.* [59], presents subspace method so called Diagonal Principal Component Analysis (DiaPCA) that deal with superlative projective vectors and conclude the results that proposed algorithm provides more accuracy in contrast to PCA and 2DPCA.

Hongta [44], proposed an idea in a new direction by adapting multi feature extraction approach that incorporates PCA and LDA. On the other hand exploit Radius Basis Function Network (RBFN) for feature classification.

In [35], PCA based face recognition is outlined which is varied in terms of testing criteria. The recognition rates vary depending on the number of training and testing sets used size of the image and even presence of noise in the face images. In [40], PCA based reconstruction procedure is applied with its novelty in using the median vector rather than the average of the class samples. It has the inherit advantage of preserving the spatial information of image samples as well as its robust nature towards outliers. In [61], a Particle Swarm Optimization (PSO) is adopted as a new method for selecting important face features. This variation of selecting the discriminatory feature instead of PCA showed great performance in terms of recognition rates. A combination of wavelet transform and SVM methodology is discussed in [54]. The former is used as a preprocessing technique while later is for classifying the facial features. A modular PCA based face recognition system is illustrated in [55], in combination with within-class median. For all the training and testing images normalization of sub-images samples is performed before projection. This result in increased recognition performance in contrast to other modular based PCA based methods.

2.1.2. Kernel Principal Component Analysis

PCA is one of the renowned methods for feature extraction. PCA performs computation on original inputs of the covariance matrix and calculate eigenvectors to transform high dimensionality into low dimensional input vector whose components are not overlapped with each other. Nonlinear PCA is also, derived from different algorithms. Different kernel methods are generalized in order to form one of the major types of nonlinear PCA called Kernel Principal Component Analysis (KPCA) [37]. Primarily KPCA calculates PCA by mapping original input into high dimensional feature space by means of different kernel methods. The functional form of the mapping function $\phi(x)$ is absolutely defined with choice of kernel, $k(xi, xj)=(\phi\ (xi),\ \phi\ (xj\ ))$, or inner product in the feature space. The suitable selection of kernel encourages the data to become separable in feature space rather than reliant in the original input space. Consequently preceding algorithms, which are incapable to grip linearly, separable data sets, provide a direction to acquire non-linear algorithms using kernel substitution methods. Three major steps of KPCA algorithm are as under:

- *Step 1:* Determine dot product of the matrix *K* using kernel function:

$$Kij=k(xi, xj) \tag{1}$$

- *Step 2:* Now calculate the Eigen vectors from the resultant *K* matrix and normalize with function:

$$\lambda k\ (\alpha k\ \alpha k)=1 \tag{2}$$

- *Step 3:* Calculate test point projection on to Eigen vectors V *k* using kernel function.

$$kPCk(x) = (Vk\Phi(x)) = \sum_{i}^{m} \alpha ki\ \ k\ \left(xi\ ,\ x\right) \tag{3}$$

1. *Evaluation:* Basically KPCA provides nonlinear geometry of face images and instructs higher order statistics [20]. Prior related algorithms do not consider the structure of manifold on which face image possibly reside so, kernel based algorithms handles the limitation of linear problem [12], by introducing the idea of nonlinearity. Additionally, kernel functions provides a way to analyze that where this type of functions are suitable and how to attain a nonlinear structure for discrimination.
2. *Discussion:* Those authors that have used KPCA for face recognition such as: Liu *et al.* [21] employ Weighted Kernel Principal Component Analysis (WKPCA) for feature extraction that integrates kernel matrix with weighted inter-class association. Consequently, results are taken from weighted kernel matrix as exposed in Table 4.

Table 4. Recognition rate and comparison [21].

Dimension	KPCA
70	0.9165
65	0.9165
60	0.9130
55	0.9009
50	0.8061
45	0.8991
40	0.9009
35	0.8991
30	0.9009
25	0.9026
20	0.8748
15	0.8643
10	0.8157
05	0.6313

Cong-De *et al.* [5], gave an idea of Kernel based 2D Symmetrical Principal Component Analysis (K2DPCA) that derives from kernel method and 2D PCA. It plots the even and odd images to high dimensional feature space in Reproducing Kernel Hilbert Space (RKHS). Finally, eigenvectors for even and odd images are calculated in RKHS. Finally utilizing high variance eigenvectors depicted in Table 5 draws the results.

Table 5. Recognition rate (%) k2dspca on cbcl database.

	1	2	3	4	5
K2DPCA	82.2 (19×3)	89.5 (19×4)	93.6 (19×5)	95.8 (19×5)	97.6 (19×4)

Nat *et al.* [30], facilitate with an idea of K2DPCA to extract nonlinear features as in KPCA. On the whole the results are accomplished on Yale and ORL databases as shown in Tables 6 and 7 respectively.

Table 6. Comparison of the top recognition accuracy (%) on yale database.

	1	2	3	4	5
K2DPCA	76.30	85.12	89.00	89.00	89.95

Table 7. Comparison of the top recognition accuracy (%) on orl database.

	2	3	4	5
K2DPCA	74.18	78.71	88.83	91.13

Lu *et al.* [22], proposed Kernel based SPCA (KSPCA) that integrates SPCA with kernel method and extract detailed information from symmetric facial image and nonlinear PCA. Finally a comparison is drawn with different kernel based method and considered KSPCA to be more accurate method. Experimental results are exposed in Tables 8 and 9.

Table 8. Recognition rate (%) of kspca on cbcl database.

	20	40	100	200	361
KSPCA	94	95.25	96.5	97	98.75

Table 9. Recognition rate (%) of kspca on orl and yale databases.

	ORL	Yale
KSPCA	97.7	95

In [19], the authors have used kernel PCA for face hallucination. Main concept in using KPCA is that it has the property of classifying linear and non-linear data. It has also, been observed that different resolution images contain similar features in the kernel subspace. Another approach for handling facial expressions and extracting suitable feature is outlined in [52]. In this regard polynomial kernel is effectively employed to deal with nonlinear structure due to expression changes. In addition nearest neighbor and Euclidean distance is used to provide compact representation. Experimentation results reveal outstanding results as compared to traditional PCA based approaches. An automatic face detection system is proposed in [48] making use of Ada Boost algorithm and followed by wavelet based feature extraction method. Fast face detection is also, proposed in [40]. In order to apply the classification mechanism support vector machines is used to effectively provide recognition results.

2.1.3. Two-Dimensional PCA

Image representation and feature extraction are pervasive techniques that are commonly used for face recognition process. In order to deal with such techniques PCA is one of the notorious methods for feature extraction [49]. In PCA the matrix of an image is transformed into high dimensional vector matrix, which is helpful for calculating the covariance matrix in high dimensional vector space. But the core constraint of such covariance matrix is that the matrix is of very large size, which creates large number of small training samples that produce hindrance to evaluate it accurately [57]. Additionally it takes much time to calculate the subsequent eigenvectors. In order to resolve these difficulties 2DPCA provides a way to handle these limitations [25]. Based on 2D images it directly calculates the covariance matrix by eradicating the image in to vector conversion phase. In contrast to PCA the size of covariance matrix is small which reduce its training samples as well as the time of eigenvectors calculation is reduced near to real time. On the other hand the results become more accurate than PCA [57]. According to Wang *et al.* [50], 2DPCA works like block based approach where each block represents one row in the input image. Even though Yang *et al.* [57], illustrates superior results of matrix based representation as compare to traditional one-dimensional vector based approach, but this method also, contains some weaknesses. First it requires extra coefficients for representing an image as compare to one-dimensional method [57]. In order to resolve such problem Kong *et al.* [17], introduces two-projection sets direction called generalized two-dimensional PCA. He proposed an iterative algorithm to determine the projection direction that provides local optimal solution, which totally based on prior conditions. Secondly, unlike PCA, it drops the facts and data of covariance matrix of different local structures in the image, which is mandatory for recognition [39].

1. *Evaluation:* When comparing 2DPCA with PCA two major advantages of 2DPCA are extracted firstly the covariance matrix is evaluated accurately and secondly small amount of time is required to compute the eigenvector [30].
2. *Discussion:* Those authors that have used 2DPCA for face recognition such as: Xu *et al.* [56] proposed a method named complete 2DPCA for face recognition. The main advantage of this method is that it gives higher recognition rate and another property of it, is that small number of feature coefficient is required. Experiments were performed on ORL database, the recognition results and their comparison with other methods is listed in Table 10.

Table 10. Comparison of the top recognition rate (%) using the ORL database.

	1	2	3	4	5
Complete 2DPCA	78.5 (12x4)	87.5 (32x4)	90.7 (15x5)	91.7 (16x4)	91.5 (15x7)

Kong *et al*. [16], present a variation of 2DPCA named Generalized 2DPCA, it is basically an extension of 2DPCA which comprises of Bilateral-projection-based 2DPCA (B2DPCA) and a Kernel based 2DPCA (K2DPCA). Main purpose of this method was to reduce the coefficient required to represent an image and gain higher-order statistics for the image. Experiments were performed on ORL and UMIST database, the recognition results and their comparison with other methods is listed in Tables 11 and 12.

Table 11. Experiment on ORL database.

	1	2	3	4	5
K2DPCA	74.5	86.9	92.0	94.6	96.2

Table 12. Experiment on umist database.

	#5, #14	#1, #7, #14	#3, #9, #15	#4, #10, #16	#5, #11, #17	#6, #12, #18
B2DPCA	90.7	91.7	95.3	95.8	94.0	92.8
K2DPCA	92.7	94.0	95.3	97.0	95.7	94.0

In [56], a similarity measure is used for classifying patterns using boosting learning. In this the authors have used 2DPCA for extracting facial features. Various experiments on ORL and YALE database showed enhanced performance. In [51], a comparison is drawn among 2DLDA and 2DPCA with its application in face recognition tasks. In [4], a color model representation for face images is outlined using quaternion matrix. This matrix uses information from R, G, B colors and also, preserves the spatial structure of face images. Experiments on FERET dataset showed great performance of the algorithm. Presented in [45] a Dual-Tree Complex Wavelet Transform (DTCWT) algorithm for representing face images is illustrated. This representation is robust to illuminations, shift invariant and showed great performance on AR and ORL face databases.

2.1.4. Kernel Based 2DPCA

The prior PCA based face recognition methods first convert the 2D face image into its corresponding 1D feature vectors. The obtained image vectors have high dimensional vector space and it complex to accurately assess the covariance matrix. The reason being is the large size of the vector and few number of the reference image samples. In order the overcome this difficulty and to calculate the eigenvector efficiently SVD based methods are used to extract the covariance matrix. Even though SVD can compute accurate results for obtaining the methods but unfortunately this does not mean the obtain eigenvector would be accurately calculated, the reason being that eigenvector are statistically obtained.

In order to overcome this new approach in [59], named 2DPCA is used for such purpose. In this case the image is directly transformed into its feature vectors using the original 2D image. As mention before every method has pros and cons so, to further improve the performance K2DPCA method came into the scene. This method makes use of the nonlinearity of principle components. In a similar way to KPCA, K2DPCA can also, extract the nonlinear features effectively instead of projecting the image on to the subspace [30]. K2PCA is an unsupervised feature extraction technique that closely resembles 2DPCA. Basic idea of kernel-based 2DPCA is discussed as under:

1. The standard 2DPCA requires the dot product when computing the nonlinear mapping or while obtaining the high dimensional features matrix $\Re^{k\times f}\,\Psi:\tau\rightarrow\mathcal{F}$ this can be shown by the following symbols. τ, $\Re^{k\times s}$.
2. Using the dot product of the high dimensional feature vector, allows the kernel based method to efficiently work on the dataset $\{\Psi(X_i)\}^N_{i=1}$.

2.1.5. Binary 2DPCA

In the public security the requirement for biometrics data with high resolution is of great concern. But an efficient and high-speed subspace is a massive problem in security related applications. In this regard PCA is an effective technique regarding the subspace selection [46]. But its computational cost in terms of time is comparatively high when in testing phase. This increased cost is due to the floating-point multiplication with each and every pixel of the given image.

Another reason being is that floating point multiplication requires more CPU cycles so, the testing phase becomes complex in the real time application where the datasets used are of high dimensions. In this regard binary B2DPCA [33, 44] came into the scene, which avoids using floating-point multiplication, instead uses addition-using integers. This greatly reduces the cost of time during the testing procedure. Another plus point of binary 2DPCA is that it's 50 times faster as compared to simple PCA. Following are the major feature listed about B2DPCA:

1. The size of the dictionary used in B2DPCA is relatively small as compared to binary PCA and selection of base is faster in contrast to BPCA.
2. Second main feature about B2DPCA is that the size of covariance matrix is small with dimension $w\times w$, while that of binary PCA is $hw\times hw$.
3. Lastly the Harr like function in B2DPCA is not effected by divided-by-zero ambiguity [7, 13].

3. Experimental Results

The databases required for evaluating the recognition accuracies are Yale, Yale-B, ORL and CMU. Performance was measured on the basis of number of correct probe image identified by the algorithms. Secondly the criterion for recognition performance was based on the number of training samples used. Different authors whom worked on these algorithms under set condition obtain varying results. The reason behind this is their varying methodologies and preprocessing conditions. Detailed in the text below, gives precise evaluation criteria for different dataset. Their average recognition rates of generalized version of algorithms are also, outlined in tabular format under the required databases.

3.1. The Yale Database

This dataset contains 165 grayscale images in GIF format of 15 individuals is included in the dataset. 11 images per subject exist, with one different facial expression, also lighting condition vary in images. Various expressions and accessories that are used are w/glasses, happy, w/no glasses, normal, right light, sad, and sleepy, surprised, and wink.

Experimsents were executed on Yale database using the above-mentioned methods. These experiments differ in the number of training samples. Recognition rates of these methods are shown in Table 13 and Figure 2. In this dataset 91.58% recognition rate is to be seen in 2DKPCA with training sample of 5 images. However, PCA could not perform well in this case due to the reason of large variation in the light and expressions.

Table 13. Recognition rate at yale database, using different training sets.

Training Samples	PCA	KPCA	2DPCA	2DKPCA
2	65.29%	74.21%	68%	70.27%
3	68.81%	81.78%	79.72%	82.28%
4	70.25%	85.33%	86.75%	87%
5	78.00%	89%	90.08%	91.58%

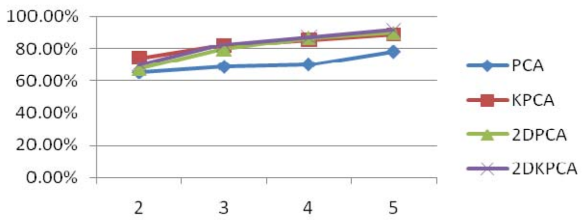

Figure 2. Graphical representation of recognition rate at yale database, using different training samples.

3.2. The Yale Face Database B

This dataset contains 5760 images of 10 subjects, each having 576 different viewing conditions (9poses×64 illumination conditions). In case of particular pose variations their background information is also, extracted.

Experiments were performed on Yale-B database using the aforementioned techniques. These experiments vary in the number of training samples i.e., from 2 training images to maximum 5 samples. Recognition rates of these methods are shown in Table 14 and graphical representation in Figure 3. Best results were obtained in this dataset depends upon the number of training samples used and it can be concluded that *K* series of PCA showed outstanding results as compared to classic PCA and 2DPCA methods.

Table 14. Recognition rate at yale database, using different training sets.

Training Samples	PCA	KPCA	2DPCA	2DKPCA
2	82.00%	85.21%	75%	87.50%
3	85.00%	86.78%	84.20%	87.75%
4	87.00%	87.90%	85.75%	91.50%
5	88.50%	90.50%	87.08%	91.58%

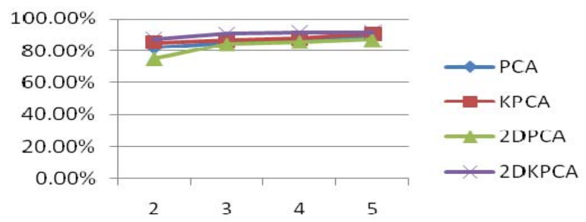

Figure 3. Graphical representation of recognition rate at yale database, using different training samples.

3.3. ORL Database

This database contains 10 various images samples of every 40 subjects. For some subjects, the captures images are taken under different timings. Some images contain varying expression, or illumination variation. The expressions are not that varying just smiling or not smiling faces or use of certain accessory like glasses or without glasses. Background of all the images is black where most of the images are taken in frontal position with slight variation of pose.

When ORL database was used for experimental purpose, which is known to be one of the classic datasets used in face recognition systems. This datasets does not contain large variation of expression, light but only slight head orientation is seen. So PCA gave effective results just like the K-series of PCA apart from 2DPCA which showed 86.08% recognition results. Whole list of RR with varying training sample is illustrated in Table 15 and Figure 4.

Table 15. Recognition rate at yale database, using different training sets.

Training Samples	PCA	KPCA	2DPCA	2DKPCA
2	83.00%	85.21%	76%	84.50%
3	85.00%	86.78%	84.20%	88.50%
4	88.00%	87.90%	84.75%	91.50%
5	90.50%	90.50%	86.08%	93.91%

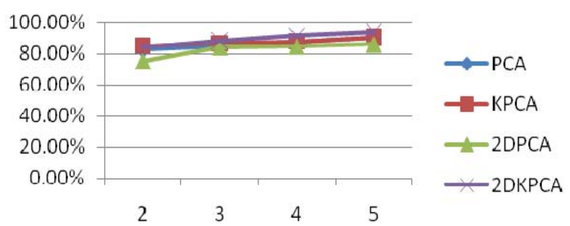

Figure 4. Graphical representation of recognition rate at yale database, using different training samples.

3.4. CMU Database

This database contains 41,368 images of 68 individuals, each having different pose, varying expressions and lighting conditions. Experiments conducted on CMU database shows that only 2DKPCA and KPCA was able to outperform with recognition rates of 90.58% and 87% respectively. The poor results shown by other methods is due to the fact that it's large variation of pose, with complex background hindered in obtaining effective results. The results are shown in Table 16 and Figure 5.

Table 16. Recognition rate at yale database, using different training sets.

Training Samples	PCA	KPCA	2DPCA	2DKPCA
2	64.29%	74.21%	68%	70.27%
3	66.81%	81.78%	73.72%	82.28%
4	70.25%	85.33%	76.75%	87%
5	75.83%	87%	80.08%	90.58%

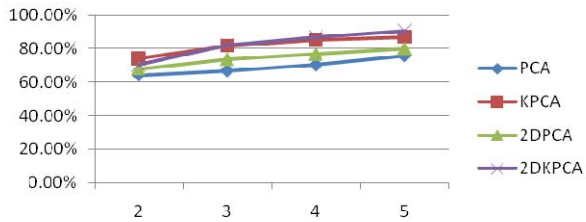

Figure 5. Graphical representation of recognition rate at yale database, using different training samples.

4. Conclusions

Main purpose of conducting this survey is to evaluate the different variations of PCA in face recognition application. Experiments conducted on different datasets by the respective authors work discussed in literature. Different authors test their results on controlled and uncontrolled conditions like glasses, scarf, facial expression, facial hair, and cosmetic etc. It is concluded that K version of PCA showed better results as compared to simple PCA and 2DPCA. The reason is that PCA could not perform well on above-mentioned datasets because of its linear nature. While KPCA and 2DKPCA performed well in uncontrolled situations of varying illumination and slight change in expression and pose.

References

[1] Belhumeur N., Hespanha P., and Kriegman J., "Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.

[2] Cevikalp H., Neamtu M., Wilkes M., and Barkana A., "Discriminative Common Vectors for Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 4-13, 2005.

[3] Chen S., Shan T., and Lovell C., "Robust Face Recognition in Rotated Eigenspaces," *in Proceedings of Image and Vision Computing New Zealand*, New Zealand, pp. 1-6, 2007.

[4] Chen S., Sun Y., and Yin B., "A Novel Hybrid Approach Based on Sub-Pattern Technique and Extended 2DPCA for Color Face Recognition," *in Proceedings of the 11th IEEE International Symposium on Multimedia*, China, pp. 630-634, 2009.

[5] Cong-De L., Yu-Lei C., and Bin-Bin H., "Kernel Based 2D Symmetrical Principal Component Analysis for Face Classification," *in Proceedings of the 7th International Conference on Machine Learning and Cybernetics*, Kunming, vol. 1, pp. 442-447, 2008.

[6] Epstein R., Hallinan P., and Yuille A., "5±2 Eigenimages Suffice: An Empirical Investigation of Low-Dimensional Lighting Models," *in Proceedings of the Workshop on Physics Based Modeling in Computer Vision*, Cambridge, pp. 108-116, 1995.

[7] Graham F., Dueck J., Funt B., and Drew M., "Color Eigenfaces," *in Proceedings of IEEE 3rd International Workshop on Image and Signal Processing Advances in Computational Intelligence*, UK, pp. 607-610, 1996.

[8] Fukunaga K., *Introduction to Statistical Pattern Recognition*, Academic Press Professional, USA, 1990.

[9] Gottumukkal R. and Asari K., "An Improved Face Recognition Technique Based on Modular PCA Approach," *Pattern Recognition Letters*, vol. 25, no. 4, pp. 429-436, 2004.

[10] Gross R., Matthews I., and Baker S., "Eigen Light-Fields and Face Recognition Across Pose," *in Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition*, Washington, pp. 1-7, 2002.

[11] Hallinan W., "A Low-Dimensional Representation of Human Faces for Arbitrary Lighting Conditions," *in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, pp. 995-999, 1994.

[12] He X., Yan S., Hu Y., Niyogi P., and Zhang H., "Face Recognition using Laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 328-340, 2005.

[13] Hjelmas E. and Wroldsen J., "Recognizing Faces from the Eyes Only," *in Proceedings of the 11th Scandinavian Conference on Image Analysis*, pp. 1-5, 1999.

[14] Jaam J., Rebaiaia M., and Hasnah A., "A Fingerprint Minutiae Recognition System Based on Genetic Algorithms," *International Arab Journal Information Technology*, vol. 3, no. 3, pp. 242-248, 2006.

[15] Kirby M. and Sirovich L., "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108, 1990.

[16] Kong H., Li X., Wang L., Teoh E., Wang J., and Venkateswarlu R., "Generalized 2D Principal Component Analysis," *in Proceedings of International Joint Conference on Neural Networks*, Canada, vol. 1, pp. 108-113, 2005.

[17] Kong H., Wang L., Teoh K., Li X., Wang J., and Venkateswarlu R., "Generalized 2D Principal Component Analysis for Face Image Representation and Recognition," *Neural Networks*, vol. 18, no. 5-6, pp. 585-594, 2005.

[18] Li J., Zhao B., and Zhang H., "Face Recognition Based on PCA and LDA Combination Feature Extraction," *in Proceedings of the 1st International Conference on Information Science and Engineering*, Nanjing, pp. 1240-1243, 2009.

[19] Liang Y., Lai J., Zou Y., Zheng W., and Yuen P., "Face Hallucination Through Kpca," *in Proceedings of the 2nd International Congress on Image and Signal Processing*, Tianjin, pp. 1-5, 2009.

[20] Liu C., "Gabor-Based Kernel PCA with Fractional Power Polynomial Models for Face Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 572-581, 2004.

[21] Liu N., Wang H., and Yau W., "Face Recognition with Weighted Kernel Principal Component Analysis," *in Proceedings of the 9th International Conference on Control, Automation, Robotics and Vision*, Singapore, pp. 1-5, 2006.

[22] Lu C., Zhang C., Zhang T., and Zhang W., "Kernel Based Symmetrical Principal Component Analysis for Face Classification," *Neurocomputing*, vol. 70, no. 4-6, pp. 904-911, 2007.

[23] Lu J., Plataniotis N., and Venetsanopoulos N., "Face Recognition using Kernel Direct Discriminant Analysis Algorithms," *IEEE Transactions on Neural Networks*, vol. 14, no. 1, pp. 117-126, 2003.

[24] Matthew T., "A Random Walk Through Eigenspace," *IEICE Transactions on Information and Systems*, vol. 84, no. 12, pp. 1586-1595, 2001.

[25] Moon H. and Phillips J., "Computational and Performance Aspects of PCA-Based Face-Recognition Algorithms," *Perception*, vol. 30, no. 3, pp. 303-322, 2001.

[26] Murase H. and Nayar K., "Visual Learning and Recognition of 3-D Objects from Appearance," *International Journal of Computer Vision*, vol. 14, no. 1, pp. 5-24, 1995.

[27] Murtaza M., Sharif M., Raza M., and Shah J., "Face Recognition using Adaptive Margin Fisher's Criterion and Linear Discriminant Analysis," *International Arab Journal of Information Technology*, vol. 11, no. 2, pp. 1-11, 2014.

[28] Nayar K., Nene A., and Murase H., "Subspace Methods for Robot Vision," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 5, pp. 750-758, 1996.

[29] Nhat V. and Lee S., "Improvement on PCA and 2DPCA Algorithms for Face Recognition," *in Proceedings of Lecture Notes in Computer Science*, Berlin , pp. 568-577, 2005.

[30] Nhat V. and Lee S., "Kernel-Based 2DPCA for Face Recognition," *in Proceedings of IEEE International Symposium on Signal Processing and Information Technology*, Gaza, pp. 35-39, 2007.

[31] Nicholl P. and Amira A., "DWT/PCA Face Recognition using Automatic Coefficient Selection," *in Proceedings of the 4th IEEE International Symposium on Electronic Design, Test and Applications*, Hong Kong, pp. 390-393, 2008.

[32] Pang Y., Tao D., Yuan Y., and Li X., "Binary Two-Dimensional PCA," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 4, pp. 1176-1180, 2008.

[33] Pentland A., Moghaddam B., and Starner T., "View-Based and Modular Eigenspaces for Face Recognition," *in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, pp. 84-91, 1994.

[34] Pereira J., Cavalcanti G., and Ren T., "Modular Image Principal Component Analysis for Face Recognition," *in Proceedings of International Joint Conference on Neural Networks*, USA, pp. 2481-2486, 2009.

[35] Poon B., Amin M., and Yan H., "PCA Based Face Recognition and Testing Criteria," *in Proceedings of the 8th International Conference on Machine Learning and Cybernetics*, Baoding, vol. 5, pp. 2945-2949, 2009.

[36] Qader H., Ramli R., and Al-Haddad S., "Fingerprint Recognition using Zernike Moments," *International Arab Journal of Information Technology*, vol. 4, no. 4, pp. 372-376, 2007.

[37] Schölkopf B., Smola A., and Müller K., "Nonlinear Component Analysis as a Kernel Eigenvalue Problem," *Neural Computation*, vol. 10, no. 5, pp. 1299-1319, 1998.

[38] Sharif M., Javed M., and Mohsin S., "Face Recognition Based on Facial Features," *Research Journal of Applied Sciences*, *Engineering and Technology*, vol. 4, no. 17, pp. 2879-2886, 2012.

[39] Sharif M., Mohsin S., Jamal M., and Raza M., "Illumination Normalization Preprocessing for Face Recognition," *in Proceedings of the 2nd Conference on Environmental Science and Information Application Technology*, Wuhan, vol. 2, pp. 44-47, 2010.

[40] Sharif M., Mohsin S., and Javed M., "Real Time Face Detection using Skin Detection (Block Approach)," *Journal of Applied Computer Science and Mathematics*, vol. 5, no. 10, pp. 75-81, 2011.

[41] Sharif M., Mohsin S., Javed M., and Ali M., "Single Image Face Recognition using Laplacian of Gaussian and Discrete Cosine Transforms," *International Arab Journal of Information Technology*, vol. 9, no. 6, pp. 562-570, 2012.

[42] Shashua A., "Geometry and Photometry in 3D Visual Recognition," *PhD Thesis*, Massachusetts Institute of Technology Cambridge, USA, 1992.

[43] Sirovich L. and Kirby M., "Low-Dimensional Procedure for the Characterization of Human Faces," *Jorurnal of Optical Society of America*, vol. 4, no. 3, pp. 519-524, 1987.

[44] Hongta S., Feng D., and Zhao R., "Face Recognition using Multi-Feature and Radial Basis Function Network," *in Proceedings of the Pan-Sydney Area Workshop on Visual Information Processing*, Sydney, pp. 183-189, 2002.

[45] Sun Z., Yang W., Sun C., and Shen J., "Face Recognition using DT-CWT Feature-Based 2DPCA," *in Proceedings of Chinese Conference on Pattern Recognition*, Chongqing, pp. 1-5, 2010.

[46] Swets L. and Weng J., "Using Discriminant Eigenfeatures for Image Retrieval," *IEEE Transactions on Pattern Analysi s and Machine Intelligence*, vol. 18, no. 8, pp. 831-836, 1996.

[47] Tang F., Crabb R., and Tao H., "Representing Images using Nonorthogonal Haar-Like Bases," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2120-2134, 2007.

[48] Tao H., Rui L., and Mei-Juan Z., "Face Recognition under Complex Conditions," *in Proceedings of International Conference on Electrical and Control Engineering*, Wuhan, pp. 960-963, 2010.

[49] Turk M. and Pentland A., "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[50] Wang L., Wang X., and Feng J., "On Image Matrix Based Feature Extraction Algorithms," *IEEE Transactions on Systems*, *Man, and Cybernetics, Part B: Cybernetics*, vol. 36, no. 1, pp. 194-197, 2006.

[51] Wang X., Huang C., Fang X., and Liu J., "2DPCA vs. 2DLDA: Face Recognition using

Two-Dimensional Method," *in Proceedings of International Conference on Artificial Intelligence and Computational Intelligence*, Shanghai, vol. 2, pp. 357-360, 2009.

[52] Wang Y. and Zhang Y., "Facial Recognition Based on Kernel PCA," *in Proceedings of the 3rd International Conference on Intelligent Networks and Intelligent Systems*, Shenyang, pp. 88-91, 2010.

[53] Wang Z. and Li X., "Face Recognition Based on Improved PCA Reconstruction," *in Proceedings of the 8th World Congress on Intelligent Control and Automation*, Jinan, pp. 6272-6276, 2010.

[54] Wei X., Sheng Y., Qi W., and Ming L., "Face Recognition Based on Wavelet Transform and PCA," *in Proceedings of Pacific-Asia Conference on Knowledge Engineering and Software Engineering*, Shenzhen, pp. 136-138, 2009.

[55] Xiao-Jie W., "Modular PCA Based on Within-Class Median for Face Recognition," *in Proceedings of the 3rd IEEE International Conference on Computer Science and Information Technology*, Chengdu, vol. 1, pp. 52-56, 2010.

[56] Xu A., Jin X., Jiang Y., and Guo P., "Complete Two-Dimensional PCA for Face Recognition," *in Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, vol. 3, pp. 481-484, 2006.

[57] Yang J., Zhang D., Frangi F., and Yang J., "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131-137, 2004.

[58] Yang M., "Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition using Kernel Methods," *in Proceedings of the 5th IEEE International Conference on Automtic Face and Gesture Recognition*, USA, pp. 215-220, 2002.

[59] Zhang D., Zhou Z., and Chen S., "Diagonal Principal Component Analysis for Face Recognition," *Pattern Recognition*, vol. 39, no. 1, pp. 140-142, 2006.

[60] Zhang J., Fei X., and Zhang Y., "A New Method for Face Recognition Based on PCA Optimize Strategy," *in Proceedings of International Conference on Computer Application and System Modeling*, Taiyuan, vol. 10, pp. 417-420, 2010.

[61] Zhao W., "Discriminant Component Analysis for Face Recognition," *in Proceedings of the 15th International Conference on Pattern Recognition*, Barcelona, vol. 2, pp. 818-821, 2000.

[62] Zheng W., Lai H., and Li Z., "1D-LDA vs. 2D-LDA: When is Vector-Based Linear Discriminant Analysis Better Than Matrix-Based?," *Pattern Recognition*, vol. 41, no. 7, pp. 2156-2172, 2008.

Jamal Hussain Shah is Lecturer in Computer Science Department at COMSATS Institute of Information Technology, Pakistan. His research areas are digital image processing and networking. He are include has more than 3 years experience in IT-related projects, he developed and designed ERP systems for different organizations of Pakistan.

Muhammad Sharif has been an assistant professor at Department of Computer Science, COMSATS Institute of Information Technology Pakistan. He is also PhD Scholar at COMSATS Institute of Information Technology. He has more than 17 years of experience including teaching graduate and undergraduate classes.

Mudassar Raza is an assistant professor at COMSATS Institute of Information Technology, Pakistan. He has more than seven years of experience of teaching undergraduate classes at CIIT Wah. He has been supervising final year projects to undergraduate students. His interest include are digital image processing, and parallel and distributed computing.

Aisha Azeem is a Lecturer in University of Wah, Pakistan. She completed her BS and MS degrees from CIIT Wah in 2008 and 2011 respectively. Her research interests include digital image processing and software engineering.