A Performance Study of SIFT, SIFT-PCA and SIFT-LDA for Face Recognition

Sanket Panda, Shaurya Nigam, Rohit Kumar, Mamatha HR

Abstract*—Humans have the ability to identify faces instantly with minimum effort and inspired by this, Face Recognition (FR) tries to imitate this ability by using numerous effective algorithms and has been extensively developed in the last decade. FR has received a lot of attention because of its wide range of its applications. Since Humans store and retrieve images instantly when needed, FR imitates this procedure by holding images in a database and trains them to recognize faces. Although many impactful algorithms have been developed, they are not entirely effective in unconstrained settings. Hence, we thoroughly compare the SIFT method and its two variations SIFT-PCA and SIFT-LDA to prove that the variations are better alternatives to regular SIFT.

***Keywords*—Face Recognition; SIFT; PCA; LDA.**

I. INTRODUCTION

Face Recognition [12] has progressed in a drastic manner in the past few years, and has been widely utilized to enforce security by a lot of law-regulation authorities. FR has been developed extensively because of its wide range of applications and the availability of face images through licenses, Identity cards and public service facilities. FR finds applications in access control and authentication to prevent fraud and its recent inclusion in smartphone security and Facebook to carryout tagging of faces proves its potential. Because of its ability to be applied to a number of different fields and settings FR technology is constantly developed and implemented in different areas. There are many FR algorithms [10] and some of them are open-source and free from licensing restrictions, while others are patented. In spite of FR being used in a lot of fields it is not effective when the variation is significant in lighting conditions, face expressions, scale of the images etc. These are known as unconstrained settings as the variations are not consistent among the given images. Hence it is very difficult to positively say that two faces are identical but the probability of a match can be determined instead. To deal with the problems associated with unconstrained settings, many excellent techniques were developed. These approaches can be divided into two main types [14]. The first is the Holistic approach that includes methods like Principal Component Analysis [11] and Linear Discriminant Analysis [15] and secondly, the Feature-based methods like Gabor and SIFT [14]. The Holistic approaches utilize the entire face to perform feature extraction and therefore do not have the usual problems that occur during the detection of nose, mouth, eyes etc.

Revised Version Manuscript Received on May 28, 2015.

Sanket Panda, PES Institute of Technology, 100 Feet Ring Road, BSK III Stage, Bangalore-560085, Karnataka, India.

Shaurya Nigam, PES Institute of Technology, 100 Feet Ring Road, BSK III Stage, Bangalore-560085, Karnataka, India.

Rohit Kumar, PES Institute of Technology, 100 Feet Ring Road, BSK III Stage, Bangalore-560085, Karnataka, India.

Dr. Mamatha HR, Professor, Information Science Department, Protocol d'Echanges pour un Systeme Interbancaire de Telecompensation, Bangalore, India.

The Feature-based approaches on the other hand extract local features using particular feature points in the face. Holistic approaches usually give good results on images that are taken according to laboratory specifications taking care to make the scale, expression, background etc. consistent. But the feature-based techniques can handle variations in unconstrained settings very well. The Scale Invariant Feature Transform (SIFT) algorithm was proposed by David Lowe [1] in 1999. It is a feature-based key-point detection/feature detection method that was proposed to be effective in case of scale variations in images (which was not possible earlier). The SIFT approach has a lot of unique properties due to which it is commonly used for building FR systems when a lot of scale variations are present or even in generalized scenarios as it can handle zoom, translation also very well. The PCA-SIFT [6] algorithm works by improving the local image descriptor of SIFT and is capable of encoding the vital aspects of the image gradient in the feature point vicinity. While SIFT uses weighted histograms PCA-SIFT uses Principal Components Analysis (PCA) for the gradient patch. LDA is also another method which works on the same dimensionality reduction principle of PCA and both of them search for linear combinations of the variables in order to describe the features. LDA tries to model the difference between the classes of data in an explicit way.

In this study, we will show that PCA-SIFT and LDA-SIFT are better alternatives to SIFT as they can provide better performance than regular SIFT by reducing the number of key-points required for matching using dimensionality reduction.

II. SCALE INVARIANT FEATURE TRANSFORM (SIFT)

SIFT [1] is feature descriptor method developed by David Lowe in 1999. It was initially developed for object recognition but has recently been effectively implemented for faces. In handling the images for matching, it can tolerate significant variations in scale, zoom, rotation, lighting etc. SIFT [13] involves four main stages. The first is the detection of the Scale-space extrema, the second is key-point localization, the third is orientation assignment and the last stage is key-point descriptor computation.

The first stage is the detection of the scale space extrema and here SIFT first searches for the locations of the scale and image and uses Difference of Gaussian to locate interest points.

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{x^2+y^2/2\sigma^2} \qquad (1)$$

The convolution of the image with the Gaussian filters of the different scales takes place (Equation 2).

$$L(x,y,\sigma)=G(x,y,\sigma)*I(x,y) \tag{2}$$

The Difference of the Gaussian (DOG) images is then performed and the maxima or minima of this DOG are used to locate the key points as follows:

$$D(x,y,\sigma)=(G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y)=L(x,y,k\sigma)-L(x,y,\sigma) \tag{3}$$

Now, detecting the scale-space extrema means convolving the image with Gaussian blurs at various scales (fig 1). We get the DOG images from the adjacent Gaussian blurred images according to individual octaves.

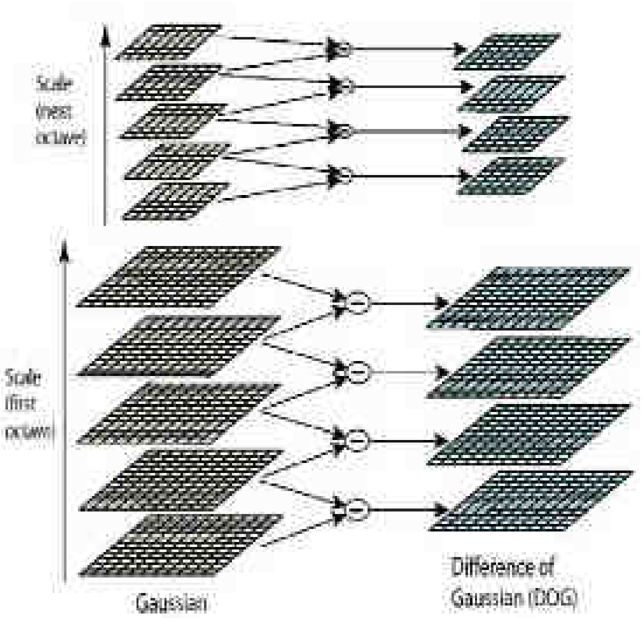


Fig.1: Scale Space Extrema

We then perform Key Point Localization [1]. We calculate the location and scale from key points.

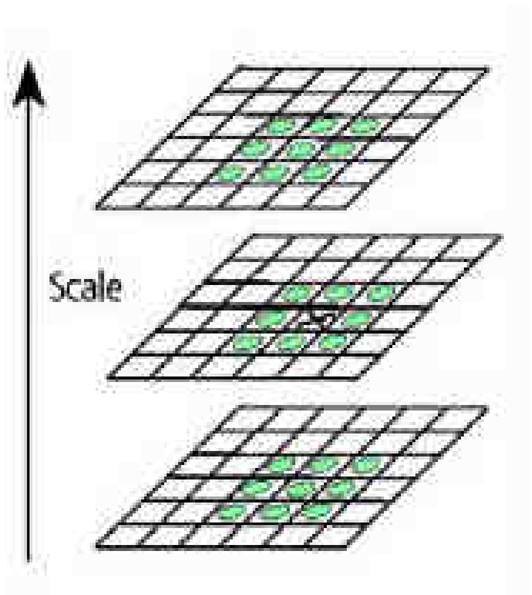


Fig.2 Detection of the Local Extrema

Fig 2 shows the pixel (x) being compared with the 26 neighboring pixels in a 3 x 3 area of the scale space volume.

The assignment of the orientation to the key points is done according to the image gradient. The smoothed image (L) at a scale of σ is used for scale calculation. For the considered image (L) at some scale, the gradient magnitude and orientation is given by m and θ.

They image gradient is calculated as:

$$m(x,y)=\sqrt{((L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1)^2)^2} \tag{4}$$

The orientation is calculated as :

$$\theta(x,y)=\arctan\frac{L(x+1,y)-L(x-1,y)}{L(x,y-1)-L(x,y-1)} \tag{5}$$

Once the key points are localised, they are described by calculating the image gradients at a particular scale around the key-point. Then the gradients are converted into a form that can be represented to allow distortion in shape and illumination.

Finally, the matching is done with the descriptors of the key points of the test and template images. We set a threshold and compare the descriptors of the test and training images, if the distance is less than the threshold, it is considered as matched. In the case that it exceeds the threshold the key point undergoes removal. After this process, the face in the training set that has the maximum number of key point matches is considered a match.

III. PRINCIPLE COMPONENT ANALYSIS - SIFT (PCA-SIFT)

Principal Component Analysis (PCA) [11] is a dimensionality reduction and noise removal technique. It has been applied to object recognition successfully and recently it has been useful in improving SIFT's performance [6]. PCA-SIFT was first proposed by Yan Ke [6] and has all the properties of SIFT such as good performance with scale, zoom etc.

PCA reduces the number of dimensions i.e. it can reduce the redundancy in the key-points of SIFT and also remove noise due to matching of points outside the face. This can improve performance because it can match faster with less key-points and also it can avoid mismatching with points in the background and hence improve accuracy.

PCA-SIFT works as follows: The first stage is the computation of the patch eigen face. Each 41 x 41 patch of the face image is transformed into a vector with 3042 elements. The covariance matrix is constructed for these vectors and PCA is applied to it. Then the matrix with the top eigenvectors is utilized for the projection matrix of PCA-SIFT [6]. Then the features are described by scanning for the feature vector of the considered image patch by constructing a image gradient vector and then it is projected to the feature space (its dimension is set as 20[6]) by using the eigenspace that was already present. Once the features are describes, the matching is then carried out using

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

Euclidean distance to calculate the feature vectors and compare them to see if there is a match.

IV. Linear Discriminant Analysis - SIFT (LDA-SIFT)

Linear Discriminant Analysis [16] or also known as Fisher Discriminant Analysis, easily handles the case where the within-class frequencies are unequal and their performances have been examined on randomly generated test data. The main difference between LDA and PCA is that PCA does more of feature classification and LDA does data classification [8]. LDA helps to better understand the distribution of feature data [9]. It tries to maximize the between-class measure while minimizing the within-class measure [9]. Following are the steps to follow to find the Fisher discriminants for a set of images:

A. Fisher face computation

The images (faces) in the database are read and divided into two sets, one for training and the other for testing. The training set is created and the average of each class each person's training images is calculated. The average of the training images in the database is determined. LDA is applied to find the within-scatter matrix and between scatter matrix. The eigenvectors are found using the between scatter matrix and within scatter matrix. The test images are transformed and all the faces in the database are fed into the face space created in the previous step. Similar to PCA, Euclidean distance method is used to carry out the matching of the test and template images [16].

V. Evaluation and Setup

We use a lot of different measures to evaluate our method and they are described in this section.

Our methods are evaluated using FAR, FRR, Sensitivity, Specificity, and Accuracy. False Acceptance Rate and False Rejection Rate are used to check if the given system is falsely accepting a wrong image or falsely rejecting a correct image. Sensitivity is the true positive rate and specificity is the false positive rate, which tells us how many correct samples and incorrect samples are accurately being identified by the given system. These tell us if the system is deviating from the behavior of an expected efficient system. So an efficient system should have low FRR and FAR and high specificity and sensitivity.

First we perform scale experiments to check if the scale variation is handled well by the algorithms by making one image smaller than the other. Here we check how many key-points are matched. Then we do experiments using lot of database images to see if can match database images well. We use a lot of common databases like AT&T, GRIMACE, FACE95 and FACE96.

The experiments are also performed to compare the database images by using a 20:80 ratio where we hold 20% of the database images as the test set and the rest of them are used for training. Based on this, we compare the test images against the template images to see if the variations: SIFT-PCA and SIFT-LDA can match the image with lesser key points. We also do not use all the key-points, only the good matches are considered according to Lowe's algorithm [1]. Based on the True Positive and True Negatives, we calculate the accuracy and other measures.

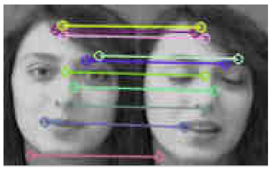
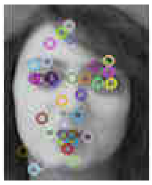
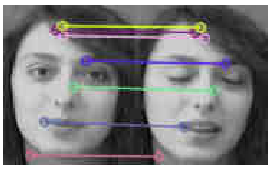
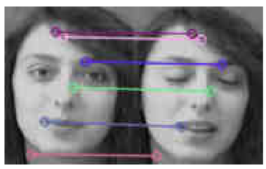
VI. Experiments and Results

We compare the SIFT, SIFT-PCA and SIFT-LDA methods on four databases, AT&T, Grimace, Face95 and Face96.

A. AT&T Database

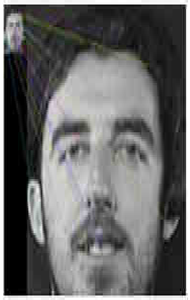
AT&T [2] has four hundred facial images and every individual has with 10 images with variations in many aspects such as their position and lighting. The images are taken with a uniform background and some persons are wearing glasses while the picures are taken.

Table 2: Database Image matching using AT&T

Algorithm	Key-point Detection	Descriptor Matching
SIFT		 10 Good Matches
SIFT-PCA		 7 Good Matches
SIFT-LDA		 6 Good Matches

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

Table 1: Checking Scale Effectiveness using AT&T (ORL) images

Algorithm	Original Image	Test with Original Images	Scale Test
SIFT		 77 keypoint matches	 13 Keypoint matches

First, we check if the change in scale (Table.1) is causing any probem with SIFT detection and matching and hence reduce the size of the test image and match it with the same face regular image (template image) which is not reduced in size, the size of the test image is reduced to about 1/10th of the template image. Table 1 shows that in this test, the key-points are still enough to properly match the face and hence SIFT can handle changes in scale well because it has 13 matches which is enough to know if it is the same face or not.

Then, we check with database images of AT&T to compare the methods and as we see in Table 2 that the SIFT-PCA and SIFT-LDA methods can match the face with less number of key-points than SIFT and therefore they are faster.

Table 3: AT&T Computational Results

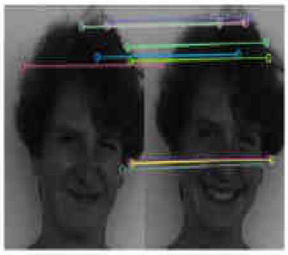
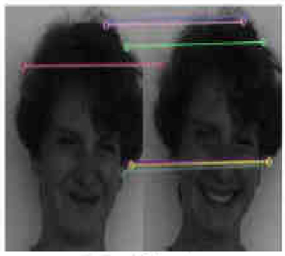
Method	Sensitivity	Specificity	FAR	FRR	Accuracy
SIFT	97.184	66.667	0.334	0.0282	93.75
SIFT-PCA	97.260	71.429	0.286	0.0274	95
SIFT-LDA	98.631	71.4286	0.286	0.0137	96.25

The computational results in Table 3 show that SIFT-PCA is 1.33% more accurate than SIFT and SIFT-LDA is 2.667% more accurate.

B. Grimace Database

Grimace [3] has 360 images and every individual has 20 images with many variations such as head movement, head turn, head scale and sharp expression changes. The expression changes (grimaces) are made by the persons and the image is captured every 0.5 seconds.

Table 4: Database Image matching using AT&T

Algorithm	Key-point Detection	Descriptor Matching
SIFT		 10 Good Matches
SIFT-PCA		 7 Good Matches
SIFT-LDA		 6 Good Matches

we check with other database images to compare the methods and as we see in Table 4 that the SIFT-PCA and SIFT-LDA methods can match the face with less number of key-points than SIFT and also remove noise and therefore they are faster.

The presence of significant expression variations make it difficult to perform key point detection and feature description on this database as we get a lot of false positives or false negatives due to this.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

Table 5 GRIMACE Computational Results

Method	Sensitivity	Specificity	FAR	FRR	Accuracy
SIFT	91.379	35.714	0.642	0.086	80.556
SIFT-PCA	93.548	40	0.6	0.064	86.112
SIFT-LDA	92.187	50	0.5	0.078	87.5

Due to the expression variations, all the methods perform inferior to AT&T with the GRIMACE database.

The computational results in Table 5 show that SIFT-PCA is 6.451% more accurate than SIFT and SIFT-LDA is 7.936% more accurate.

C. *Faces95 Database*

Faces95 [5] has 1440 and every individual has 20 images with variations in head scale due to the forward movement of persons. There are also variations in head turn, head turn, head scale and lighting. The images are captured every 0.5 seconds between variations.

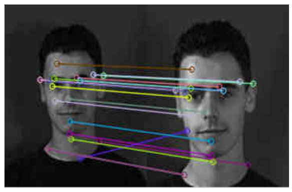
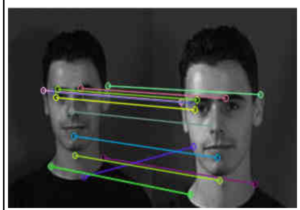
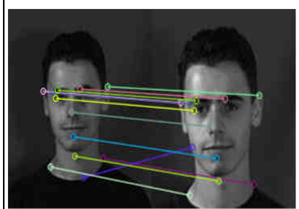
we check with database images of Face95 to compare the methods and as we see in Table 6 that the SIFT-PCA and SIFT-LDA methods can match the face with less number of key-points than SIFT and therefore they are faster.

Table 7: Faces95 Computational Results

Method	Sensitivity	Specificity	FAR	FRR	Accuracy
SIFT	97.647	60	0.4	0.023	93.684
SIFT-PCA	97.647	70	0.3	0.023	94.736
SIFT-LDA	97.674	77.777	0.222	0.023	95.789

The computational results in Table 7 show that SIFT-PCA is 1.111% more accurate than SIFT and SIFT-LDA is 2.197% more accurate.

Table 6: Database Image matching using Face95

Algorithm	Key-point Detection	Descriptor Matching
SIFT		 17 Good Matches
SIFT-PCA		 12 Good Matches
SIFT-LDA		 12 Good Matches

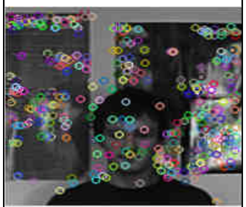
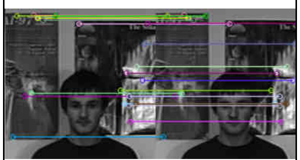
Since Face95 consists of significant head scale variations as the persons move forward for each subsequent image. If we choose non-adjacent images for the tests, we can test how well the system works with the presence of translation (position of the face in the image). Even though there is significant translation in our chosen images, we can see that the SIFT algorithm was able to match the correct face. SIFT-PCA and SIFT-LDA are also good at handling translation in images and were able to do this with lesser number of key-points. SIFT-PCA and SIFT-LDA were equally good at handling this.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

D. *Faces96 Database*

Faces96 database [17] has 3040 images and every individual has 20 images with variations in head movement, head turn, tilt, slant etc. The position of the face also varies as the individual moves forward while the images were captured. The images are captured every 0.5 seconds. The background is complex due to the glossy posters behind the individual.

Table 8: Database Image matching using Face96

Algorithm	Key-point Detection	Descriptor Matching
SIFT		 10 Good Matches
SIFT-PCA		 7 Good Matches
SIFT-LDA		 6 Good Matches

we check with database images of Face96 to compare the methods and as we see in Table 8 that the SIFT-PCA and SIFT-LDA methods can match the face with less number of key-points than SIFT and therefore they are faster.

The computational results in Table 9 show that SIFT-PCA is 2.409% more accurate than SIFT and SIFT-LDA is 0.881% more accurate.

The overall performance falls on this database due to the presence of glossy posters (complex background) but still the SIFT-PCA and SIFT-LDA were capable of handling this better than SIFT. The performance of SIFT-PCA was better than SIFT-LDA on this database, which means that PCA was better at handling the background noise than SIFT-LDA. So in the presence of background noise in a given image, SIFT-LDA can be chosen.

Table 9: Faces96 Computational Results

Method	Sensitivity	Specificity	FAR	FRR	Accuracy
SIFT	88.571	76	0.24	0.114	85.263
SIFT-PCA	90	80	0.2	0.1	87.368
SIFT-LDA	89.855	75	0.25	0.101	86.021

VII. ALGORITHM COMPARISON

The accuracy of the SIFT, SIFT-PCA and SIFT-LDA algorithms on the various databases is shown in Table 10.

Overall, by the summation of the accuracies on various databases, the SIFT-PCA method was by 2.819% more accurate than SIFT and SIFT-LDA was 3.483% more accurate than SIFT.

Table 10: Accuracies of the Algorithms on Various Databases

TECHNIQUE	AT&T	GRIMACE	FACES95	FACES96
SIFT	93.75	80.556	93.684	85.264
SIFT-PCA	95	86.111	94.736	87.369
SIFT-LDA	96.25	87.5	95.789	86.022

We can observe that the SIFT-PCA and SIFT-LDA are capable of matching the given face images with lesser number of key-points than the regular SIFT. The excess key points that are present in SIFT due to redundancy and the noise that often exists, such as matching of the background instead of the face are removed by PCA and LDA. Hence they can provide better performance than regular SIFT.

The overall accuracy showed a reduction in the Grimace and Face96 databases because Grimace has sharp variations in expressions and Face96 has complex background. But even

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

on these, the SIFT-PCA and SIFT-LDA algorithms worked better by matching with lesser number of key points and provided better accuracy.

SIFT-LDA showed better accuracy than SIFT and SIFT-PCA on AT&T, GRIMACE and FACE95 but on FACE96, SIFT-PCA showed better accuracy than SIFT-LDA. This may be due to the way in which LDA handles the noise in the complex background formed due to the presence of the glossy posters.

We can infer that the SIFT-PCA and SIFT-LDA methods are better than regular SIFT.

The accuracies of the algorithms on the various databases are shown in Fig.3.

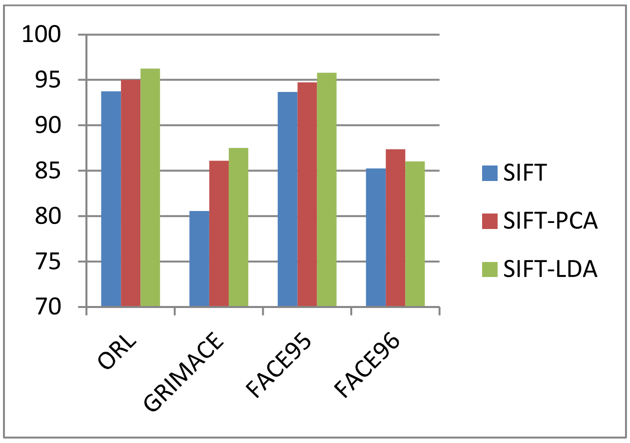


Fig. 3 Accuracies of the Algorithms on Various Databases

VIII. CONCLUSION AND FUTURE DIRECTIONS

We have compared the SIFT algorithms with its variations: SIFT-PCA and SIFT-LDA. Our experiments show that the SIFT-PCA and SIFT-LDA work better than SIFT as they can recognize faces with lesser number of key-points than regular SIFT and they are also capable of removing noise and therefore they provide better accuracy.

For our future work, we are working on comparing SIFT's accuracy with two-dimensional techniques [7] like 2D-PCA and 2D-LDA. 2D methods are shown to be more accurate than 1D methods like PCA and LDA and we will inquire whether they can provide better results than those in this study.

REFERENCES

1. D.G. Lowe. Distinctive image features from scale-invariant key-points. International Journal of Computer Vision, 60(2):91–110, 2004.
2. AT&TDataset:http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.
3. GrimaceDataset: http://cmp.felk.cvut.cz/~spacelib/faces/grimace.html
4. Isra'a Abdul, Ameer Abdul Jabbar, Jieang Tan "Adaptive PCA-SIFT matching approach for Face recognition application", Proceedings of the International Multi Conference of Engineers and Computer Scientists, Vol I, pp 1-5, IMECS 2014.
5. Face 95: http://cmp.felk.cvut.cz/~spacelib/faces/faces95.html
6. Y. Ke and R. Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. In Computer Vision and Pattern Recognition, pages 506–513, 2004.
7. Alsaqre, Falah E., and Saja Al-Rawi. "Symmetry Based 2D Singular Value Decomposition for Face Recognition." In Digital Information Processing and Communications, pp. 486-495. Springer Berlin Heidelberg, 2011
8. LINEAR DISCRIMINANT ANALYSIS - A BRIEF TUTORIAL S. Balakrishnama, A. Ganapathiraju.
9. L-F Chen, H-Y Mark Liao, M-T Ko, J-C Lin, and G-J Yu, "A new LDA-based face recognition system which can solve the small sample size problem", Pattern Recognition, vol. 33, pp. 1713–1726, 2000.
10. De Carrera, Proyecto Fin. "Face recognition algorithms." *Master's thesis in Computer Science, Universidad Euskal Herriko* (2010).1
11. Wold, Svante, Kim Esbensen, and Paul Geladi. "Principal component analysis." *Chemometrics and intelligent laboratory systems* 2, no. 1 (1987): 37-52.
12. Zhao, Wenyi, et al. "Face recognition: A literature survey." *Acm Computing Surveys (CSUR)* 35.4 (2003): 399-458.
13. Bicego, Manuele, Andrea Lagorio, Enrico Grosso, and Massimo Tistarelli. "On the use of SIFT features for face authentication." In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on*, pp. 35-35. IEEE, 2006.
14. Križaj, Janez, Vitomir Štruc, and Nikola Pavešić. "Adaptation of SIFT features for robust face recognition." In *Image Analysis and Recognition*, pp. 394-404. Springer Berlin Heidelberg, 2010.
15. A. M. Martinez, A. C. Kak, "PCA versus LDA", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 23, 2001, pp.228-233
16. R. O. Duda, P. E. Hart, D. G. Stork, Pattern Classification. Wiley, New York (2001)
17. Face 96: http://cmp.felk.cvut.cz/~spacelib/faces/faces96.html
18. R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. In *Proceedings of ComputerVision and Pattern Recognition*, June 2003.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.