A REVIEW ON FACE FEATURE EXTRACTION TECHNIQUE FOR FACE RECOGNITION SYSTEM

[1]Rohan Naik, [2]Kalpesh Lad

[1]Assistant Professor, Naran Lala College of Professional & Applied Science, Navsari, India.-396445.
naikrohan@gmail.com
[2]Assistant Professor, Shrimad Rajchandra Inst. of Management & Comp. Appl., UTU, Bardoli, India–394350.
kalpesh.lad@utu.ac.in

***Abstract*: Face recognition applications are widely used in variety of areas like human authentication, surveillance system etc. In Front face recognition, we may compare whole face template but for side view face recognition it is advisable to extract individual face features for further investigation. So, face feature extraction is vital step of face recognition because recognition result depends on how accurately features are extracted. An accurate classifier gets failed if face features are not accurately extracted. Authors' review existing methods of face feature extraction with their advantages and limitations.**

***Keywords*: Face Recognition, Gabor, PCA, SIFT, Eigen Space, Edge detection.**

I. INTRODUCTION

Face recognition has always been a challenge in computer vision in recent years; efficient face recognition has keen focus of research for human identification. Over the recent years, face recognition system consisting of mainly three steps i) face detection, ii) feature extraction, iii) classification (feature matching).

In the detection stage, the portion of the image having face is located. In the feature extraction step, the detected face is described in terms of various features. If feature extraction step does not perform well then best classifier cannot produce accurate results. Good feature extraction is those that on one hand, minimizes intra-person dissimilarities and on the other hand maximizes differences between persons. Finally, the classification step determines whether a certain feature corresponds with a target face or a model [1]. Feature extraction is very essential if face is side view, it is not advisable to compare whole side face temple with database front face because rim of face is different [5, 25]. So it requires efficient feature extraction technique.

2. REVIEW OF FACE FEATURE EXTRACTION TECHNIQUES

Following are popular technique used for feature extraction for front and side view face recognition up to certain angel with restraint.

2.1 Gabor Filter Approach

Gabor filters have proved to be a powerful tool for facial feature extraction for front face. This method is capable of deriving multi-orientation information from a face image at several scales, with the derived information being of local in nature [2, 3, 4]. The usual approach while using Gabor filters for face recognition is to construct a filter bank with filters of different scales and orientations. In the spatial domain, A 2-D Gabor filter in time domain is represented as

$$\varphi_{u,v}(x,y) = \frac{{f_u}^2}{\pi\gamma\delta}.\exp\left(\frac{f^2}{\gamma^2}x'^2 + \frac{f^2}{\delta^2}y'^2\right)\exp(j2\pi f_u x')$$

Where $x' = x\cos\theta v + y\sin\theta v$, $y' = -x\sin\theta v + y\cos\theta v$, and the parameters fu and θv are defined as $fu = fmax/2(u/2)$ and $\theta v = v\pi/8$.

Gabor filters represent Gaussian kernel functions modulated by a complex plane wave whose center frequency and orientation are defined by fu and θv, respectively. The factors γ and η define the ratio between the center frequency and the size of the Gaussian envelope. Commonly the values of γ and η are set to $\gamma = \eta = \sqrt{2}$.

Let $I(x, y)$ denote a grey-scale face image defined on a grid of size $m\times n$ and let $\psi u, v(x, y)$ represent a Gabor filter determined by the parameters fu and θv. The filtering operation with the Gabor filter can then be written as follows

$$Gu,v(x, y) = I(x, y) * \psi u,v(x, y).$$

Where $Gu,v(x, y)$ denotes the complex convolution result which can be decomposed into a real and an imaginary part:

$$Eu,v(x, y) = Re[Gu,v(x, y)]$$
$$Ou,v(x, y) = Im[Gu,v(x, y)].$$

Based on the decomposed filtering result both the phase $\varphi u,v(x, y)$ as well as the magnitude $Au,v(x, y)$ filter responses can be computed as:

$$A_{u,v}(x,y) = \sqrt{{E_{u,v}}^2(x,y) + {O_{u,v}}^2(x,y)}$$

$$\phi_{u,v}(x,y) = arctan\left(O_{u,v}(x,y)/E_{u,v}(x,y)\right)$$

While exhibiting desirable belongings, such as orientation selectivity or spatial locality, Gabor filters have also some shortcomings too. These limitations crucially affect the characteristics and size of the Gabor illustration of a given face image. Amongst these demerits the use of non-orthogonal filters is the most important one. This redundant information makes the Gabor face representation affects the size of the face representation.

Gabor filter works well with front face orientation but not good support in side view face and face expression. Its execution speed is low as it require more no of iterations.

2.2. SIFT Based Face Recognition

Scale-invariant feature transform (SIFT) is an algorithm in computer vision to detect and describe local features in images. Due to its excellent performance, SIFT was widely accepted for many face recognition applications, but the implementation of SIFT was complicated because of detecting false key-points in the face image due to background objects. Reliable and accurate matching of image features has become a basic problem in computer vision applications. Scale-Invariant Feature Transform (SIFT) was published in 1999 and upgraded in 2004 [6] which was used to detect and describe local image features. SIFT is an outstanding feature descriptor, because it is invariant to linear scaling, orientation changes, and partially Invariant to affine distortion and illumination changes.

Various steps in SIFT algorithm are:

Step 1: Constructing a Scale Space:
We begin with an image and generate progressively blurred copies of it. Then original image will resize to half of its size and again create blurred images and the process continues. The number of octaves, octave number of blurred images in each octave and the amount of blurring depends on size of the image. Mathematically, the blur can be obtained by Gaussian operator and is shown below

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

$$L(x, y, \sigma) = G(x, y, \sigma) * L(x, y)$$

L = Blurred image.
G = the Gaussian blur operator.
I = Image to get blurred.
x, y = coordinates.
σ = the amount of blur.

Step 2: Application of Gaussian:
While finding Difference of Gaussian, an image can be blurred a little and a second order derivative can be calculated on it. This locates edges and corners on the image. These edges and corners are good for finding key points. It should be noted that the second order derivative is extremely sensitive to noise. The blur smoothes the noise and stabilizes the second order derivative.

Fig.1: DoG extraction from the octaves

Step 3: Finding Key-points:
Finding key points is a two part process as mentioned below.
a. locate of maxima/minima in DoG
b. find sub pixel maxima/minima

Step 4: Getting rid of Bad Key-points:
Bad key points are removed by using following processes.
a. removing low contrast features:
b. removing edges:

Step 5: Assigning Orientation to Key-points:

Here in this step we collect gradient directions and magnitudes around each key-point. The gradient and magnitude will calculate by these formulas:

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

$$\theta = tan^{-1}((L(x,y+1) - L(x,y-1))/(L(x+1,y) - L(x-1,y)))$$

The magnitude and orientation is calculated for all pixels around the key-point.

Step 6: Generate Histogram based on features:
The paper [7] proposes Volume- SIFT (VSIFT) and Partial-Descriptor-SIFT (PDSIFT) for face recognition based on the original SIFT algorithm. Here, Cong Geng et.al. performs comparison of holistic approaches: Fisherface (FLDA), the null space approach (NLDA) and Eigen feature Regularization and Extraction (ERE) with feature based approaches: SIFT and PDSIFT. Experiments on the ORL and AR databases show that the performance of PDSIFT is significantly better than the original SIFT approach. Moreover, PDSIFT can achieve comparable performance as the most successful holistic approach ERE and significantly outperforms FLDA and NLDA. Volume-SIFT (VSIFT) approach eliminates unreliable keypoints not based on the contrast value of $\left|O_{(x,y,\sigma^2)}\right|$at the locationof each candidate keypoint but based on the volume of thestructure:

$$V_{(x,y,\sigma^2)} = \sigma^2 \left|O_{(x,y,\sigma^2)}\right| = \sigma^4 \left|\nabla^2 G * I_{(x,y,)}\right|$$

Where σ is the scale of the keypoint.

$$V_{1(x,\ y,\sigma^2)} = \sigma 1^2 \left|O_{1(x,\ y,\sigma 1^2)}\right|$$

$$V_{2(x,\ y,\sigma^2)} = \sigma 2^2 \left|O_{2(x,\ y,\sigma 2^2)}\right|$$

Here, the values of $V_{(x,\ y,\sigma^2)}$at location of each candidate keypoint are evaluated, and, if this value is below some threshold, the keypoint is removed. In order to compute the descriptor for each keypoint, a support of 16 × 16 pixel neighborhoods around each keypoint in the Gaussian blurred image at the key point's scale is needed. For an input image with a size 200 × 200 pixels, detect some key points around the scale σ = 32. And for the keypoint to be detected on a large scale, we need a support with size
(16 × 24) × (16 × 24) = 256 × 256 pixels according to the original input image. In this case, the keypoint are not described due to the limitation of image size. Hence, here proposing the Partial-Descriptor-SIFT (PDSIFT) to enable the keypoints detected at large scales or near face boundaries,

descriptor F is a 128 dimensional vector. And we define a diagonal matrix M as a descriptor mask with size 128 × 128

$$m_{ii} = \begin{cases} 1, if\ the\ block\ is\ valid \\ 0, if\ the\ block\ is\ invalid \end{cases}$$

Where I ∈ (1,128).

And the similarity between gallery descriptor

$$F_{mg} = M_g F_g$$

And probe descriptor,

$$F_{mp} = M_p F_p$$

can be calculated as:

$$S_{gp} = \frac{{F_{mg}}^T . F_{mg}}{|M_p F_{mg}||M_g F_{mp}|}$$

Where Mg and Mp are mask matrices for Fg and Fp, respectively. S_{gp} actually represents the similarity between gallery and probe descriptors in the common valid subspace.

2.3 KERNAL PCA Based Extraction

Different color spaces are applicable for different applications. This paper [8] investigates the performance of face recognition with some color spaces using kernel-based Principal Component Analysis (Kernel-PCA). Kernel-PCA is a non-linear extension from the normal PCA algorithm. Experiments are performed with the Gaussian kernel function. Color spaces are linear or non-linear transform from RGB. In this paper, the RGB, YCbCr, and HSV color spaces are compared with the gray image (luminance information -Y). Kernel-PCA is used to extract features from individual color components or from combining the three components of every color space in one vector. The experiments are performed on FEI color database [9]. FEI database is frontal face images with seven profile images rotation of up to about 180 degrees and two different facial expression images. The experimental results show that the V color component of the HSV colorspace outperforms all the used color organization.

Schölkopf et al. [10] have developed a nonlinear PCA called Kernel-PCA. The kernel-PCA is not interested in principal components in input space, but rather in principal components of variables which are nonlinearly related to the input variables [11].

In brief, the following steps were necessary to compute the principal components: first, compute the dot product matrix, the associate eigenvector. Secondly, compute its eigenvectors and normalize them. Third, compute projections of a test point onto the eigenvectors. The k-nearest neighbor classifier (k-NN) is a method for classifying objects by finding the closest k neighbors in the feature space. Different parameters are used with k-NN, such as value of k nearest neighbors and distance model. In our work, the Euclidean norm distance is used find the class of the closest query point.

For face recognition using Kernel-PCA, the use of one individual color components is more effective than the use of the combination of three color components. The Kernel-PCA method with gray images performed a recognition rate of 89.25%. The V from the HSV color space outperforms all the used color components, 91.75%. The individual color components can provide an improvement from 0.5% to 2.5% recognition rate while using the gray image.

2.4 MB LBP Based Face Recognition

Ahonen *et al.* [12] proposed Local Binary Patterns (LBP) method for extract local information, which is invariant to illumination and orientation conditions. Since then many researchers have applied LBP for Face recognition. Petpon *et al.* [13] applied LBP for Face recognition on Yale face database B and obtained average recognition rate of 87 %. Further, variants of LBP features have been proposed namely Elongated LBP (ELBP), Multi Scale block LBP (MBLBP), VLBP, LGBP, SLBP [14]. Petpon *et al.* [13] applied LLBP for face recognition on Yale face database B and obtained 89.71% average recognition rate, SLBP on the same database and reported 62.28% average recognition rate. Applying MB-LBP yielded 89.68% average recognition rate.

The original Local Binary Patterns (LBP) operator labels the pixels of an image with decimal numbers, which are called as LBPs or LBP codes by accounting the local structure around that pixel [12]. Here the central pixel's gray-value is compared with its eight neighbors and pixels values lesser than the center pixel are encoded with 0 and others are encoded as 1. For a given pixel, concatenating all these binary values in a clockwise direction, which starts from the one of its top-left neighbor, a binary number is obtained and the corresponding decimal value of this binary number is then used to label the given pixel.

Features calculated from the 3x3 local neighborhoods cannot capture macro-structures (larger scale structure) that may be dominant features of a face. In order to capture the micro structures and as well as the macro structures, multi scale block LBP (MB-LBP) was proposed [15]. Multi scale block LBP is a variant to the basic LBP with neighborhoods of different operator sizes (3 x 3, 6 x 6, 9 x 9, 12 x 12, 15 x 15 and 18 x 18, to name a few). MB-LBP labels a pixel of the image by comparing the average gray-values of neighborhood sub-regions with the average of center sub regions gray value. The MB-LBP operator consists of 9 blocks. The scale of the MB-LBP operator is denoted as *s* x *s*, where *s* denotes the size of each block in the operator. The scalar values of averages over blocks are computed using the integral image so MB-LBP feature extraction process is faster compared to other variants of LBP.

MB-LBP operator gives better recognition rate than the holistic approach using Mahalanobis cosine distance metric. The performance of face recognition system is affected by the operator size and the grid that the face image is divided. It is found that as the operator size increases, the recognition rate decreases. If the number of grids per image is increased to some extent, the recognition rate improves due to increased spatial information found in the facial images. There is a trade-off in operator size and maximum grids per image to get better recognition rate.

2.5 EIGEN SPACE APPROACH

Eigenspace-based methods, mostly derived from the Eigenface-Algorithm, project input faces onto a dimensionally reduced space where the recognition is carried out, performing a holistic analysis of the faces. Different eigen space-based methods have been proposed. Eigenspace-based face recognition [16 - 19] corresponds to one of the most successful methodologies for the computer aided recognition of faces in digital images. Starting with the Eigenface-Algorithm, different eigenspace based approaches for the recognition of faces have been proposed in the last few years. They differ mostly in the kind of projection method used, in the projection algorithm employed, or in the use of simple or differential images before/after projection, and in the similarity matching criterion or classification method employed.

The first step is in Eigen space algorithm is to create database and load database, the database can have M number of images that each image have size of N*N. Once we have the training set, convert face images in training set to face vectors. Each image in data set should convert to a column vector. If image presents by the size of $N*N$ converts to a $1*N^2$ vector columnform.

Fig 2: convert image to vector colum.

Once we have converted all images to face vectors, the next step is normalisethe face vectors.For the normalization first we calculate the average vector(Mean) of face vectors. Mathematically Mean of each observation that is shown in equation (1).

$$Mi = \sum_{i=0}^{mn} \sum_{i=1}^{p} face\_Db(i,j) \qquad (1)$$

The mean vector will have the dimension $(M)_{nm*1}$

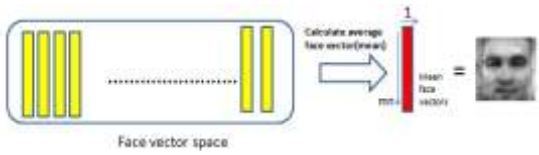

Fig 3: calculate mean vectors.

After calculate the average face vectors(means) then we calculate the normalization of face vectors ,for normalized face vector easily we subtract face vectors from average face vector(mean).

Mathematically is shown as

$$(\Delta(M))_{nm*1} = (face\_Db(i)_{nm*1} - (M)_{nm*1} \qquad (2)$$

Where i∈ $1,2,3.....p$

After normalize the face vector we calculate eigenvectors that we need to calculate covariance matrix C, The covariance matrix has given by

$C=AA^T$ [3]

Where A= $\{N_{f1}, N_{f2}, N_{f3}, \ldots\ldots\ldots\ldots N_{fm}\}$

Matrix A has size of $N^2 * M$, which each normalized face vector has size of N^2 and M is the number of face vectors. The size of the Covariance matrix s very large

$C=A_{N^2*M} \cdot A^T{}_{M*N^2} C_{N^2*}$

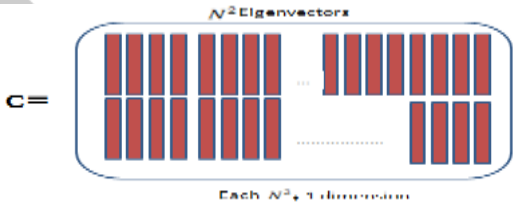

Fig 4: high dimensional eigenvector matrix.

With the above calculation the covariance matrix system will run slowly or run out of memory because the computationalcomplexity will be so large. So dimensionaly reduction is applied to the Co variance matrix

$C=A^TA$

Then

$C=A_{M*N^2} \cdot A^T{}_{N^2*M} C_{M*M}$

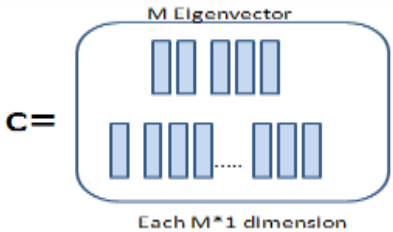

Fig 5: low dimensional eigenvector matrix.

With the above formula we reduce the calculation and effect of noise on the needed eigenvectors.If you assume N^2=2500 and M =100 then the first covariance has the size of 2500*2500 but the second fourmula will give 100 * 100 it means we reduce noise and calculation from eigenvectors.

The next step after finding eigenvector is to find the best orthoganal noise free Eigen-Faces.

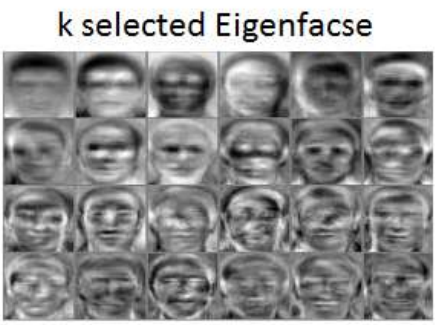

Fig 6: selected Eigen-faces.

The above Eigen faces should represent each image in training set. We avoided lot of calculation and effect of noise reduces dimensionality. After finding the best K Eigen-Faces the next step is to find Weight vector for each Eigen-Face. It's shown below.

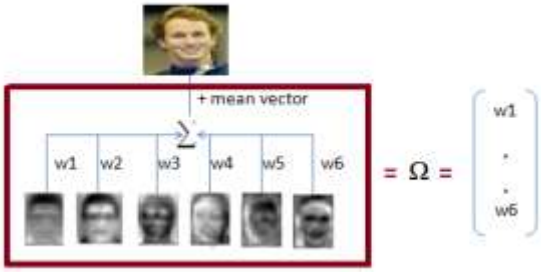

Fig 7:shows the weight vectors.

Ω= weight vector, the weight vector represents by PCA result.Each w is the percentage of each eigenface that contribute to make each image in training set.

The projection methods employed in the paper [20] for the reduction of dimensionality are PCA [21], FLD [22], and EP [23].PCA is a general method for identifying the linear directions in which a set of vectors are best represented in a least-squares sense, allowing a dimensional reduction by choosing the directions of largest variance. The theoretical solution of this problem is well known and is obtained by solving the eigensystem of the correlation matrix **R**. On the other hand, FLD searches for the projection axes on which the input vectors of different classes are far away from each other (similar to PCA), and at the same time input vectors of a same class are close to each other. The solution of this problem is obtained by solving the general eigensystem for the so-called within-class and between-class scatter matrices. When using the standard approach, the best results are obtained with the FLD-Whitening-Euclidean combination. Using other FLD combinations very similar results are obtained. These results confirm the better theoretical discrimination ability of FLD over PCA, and the results reported when Fisherfaces was proposed for the first time . It also ealuated that the FLD algorithm obtains projection axes that best separates the input data in a least-squares sense.The Bayesian classifier and SVM give similar results while using small and large datasets.

2.6 Improved PCA with DWT

In this paper [24], an improved PCA face recognition algorithm based on the Discrete Wavelet Transform and the Support Vector Machines is analyzed. The 2-D Discrete Wavelet Transform has been used to process the ORL standard face images to form the low frequency sub images by extracting the low frequency component. Then the PCA method is used to obtain the characterizations of these sub images. At last, the extracted eigenvectors are put into the SVM classifier for training and recognition. The experimental results indicate that this algorithm reduces the computational quantity because the dimension of the total population scatter matrix of the source images has deduced a lot and the performance of the SVM classifier is superior to many other classifiers. Compared with the traditional PCA face recognition algorithm, the calculation speed and the recognition efficiency are enhanced.

The main steps in this algorithm are

1. Wavelet decomposing

Related experiments manifest that the information of the face images have been decomposed in some degrees after the 2-D Discrete Wavelet Transform and the sum of every sub image is close to the entire energy. The size of the image decreased to 1/16 of the original image. In this paper, the ORL face images have been through the 2-D Discrete Wavelet Transform first, the low frequency component LL2 been extracted then so that the resolution rate reduced to 28×23.

2. Characteristic extraction

Calculate the total population scatter matrix

$$S_w = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)(x_i - \mu)^T$$

of the processed images above, here x_i refers to the image vector of number i low frequency face image training set, μ refers to the average image vector of the set, N refers to the total number of the image. Then do the K-L transform to Sw. Owning to the S_w can be expressed

$$S_w = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)(x_i - \mu)^T = \frac{1}{N}XX^T.$$

By a matrix, R=X^TX can be constructed. Then calculate the eigenvalue and the corresponding eigenvector through SVD decomposing and then the K-L orthogonal normalized eigenvector

$$u_i = \frac{1}{\sqrt{\lambda_i}}Xv_i$$

3) SVM classification and the reorganization

For the multiple class pattern recognition, there are many ways. Suppose there are k classes.

a) One-Against-The-Rest:

Construct k sub classifiers, earmark the sample belongs to the class j as the positive class and the other samples as the negative class. Choose the class which has the maximum discriminate function as the class of the test data.

b) One-Against-One:

Construct k*(k- 1) /2 subclassifiers, every classifier trains two different kinds of data and the voting strategy has been used in this classifier.

c) The way to resolve one time:

Different from the above two methods, another method is presented that can solve the multiple class problems for one time; it evaluates k SVM classifiers simultaneously from an optimized problem.

In this paper [24], the 2-level 2-D Discrete Wavelet Transform with the sym2 wavelet base has been used to preprocess the face images in the ORL standard facedatabase, then the low frequency component has been extracted so that the total population scatter matrix of the face images can be trailed off and the most energy of the image is reversed. Then use the classical PCA algorithm to calculate the eigenvector and put the effective eigenvectors into the SVM Classifier fortraining and recognizing. The experiments prove that this algorithm can improve the calculation speed and recognition rate. The further research-orientation in thefuture is to choose the wavelet base which can lead to faster calculation speed and higher recognition rate to improve the preprocessing efficiency of the image, and to ameliorate the classifying strategy of the SVM classifier to solve the problem that recognition speed slows down when the amount of the extracted eigenvector is too large.

III. CONCLUSION

Various Face feature extraction methods are discussed and its performance is analyzed. Most of this popular face recognition algorithms are limited in its performance due to various reasons like illumination changes, pose variations etc. So, it is difficult to develop generalise approach that work in all environment of face recognition.

IV. FUTURE WORK

Many approaches of front face recognition is developed but still it requires to develop approach of side view face recognition with specific angle in control environment with normak face expression behaviour.

REFERENCES

[1] R. Naik, K. Lad "FACE FEATURES BASED SIDE VIEW FACE RECOGNITION SYSTEM" NATIONAL JOURNAL OF SYSTEM AND INFORMATION TECHNOLOGY,VOL. 9, NO. 1, ISSN:0974-3308, Impact Factor: 3.894, JUNE 2016.

[2] Vitomir ˇ Struc, RokGajˇsek and Nikola Paveˇsi´c, “Principal Gabor Filters for Face Recognition”. Biometrics: Theory, Applications, and Systems, 2009. BTAS '09. IEEE 3rd International Conference on 28-30 Sept. 2009.

[3] Ki-Chung Chung; SeokCheolKee; Sang Ryong Kim, "Face recognition using principal component analysis of Gabor filter responses," Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 1999. Proceedings. International Workshop on , vol., no., pp.53,57, 1999

[4] Rahman M.T.; Bhuiyan, M.A., "Face recognition using Gabor Filters," Computer and Information Technology, 2008. ICCIT 2008. 11th International Conference on , vol., no., pp.510,515, 24-27 Dec. 2008

[5] R. Naik, K. Lad "A Review on Side-View Face Recognition Methods" International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) ISSN(Online): 2320-9801 ISSN (Print) : 2320-9798 Vol. 4, Issue 3, March 2016.

[6] Lowe D. Distictive Image Features From Scale-InvariantKeypoints[J], International Journal of Computer Vision, 2004,60(2):91-110.

[7] Cong Geng and Xudong Jiang, “ Face Recognition Using SIFT Features ” in Image processing (ICIP), 2009 16th IEEE International Conference on ,vol. no. pp.3313-3316,7-2`10, Nov. 2009

[8] Ebied, H.M., "Kernel-PCA for face recognition in different color spaces," in Computer Engineering & Systems (ICCES), 2012 Seventh International Conference on , vol., no., pp.201-206, 27-29 Nov. 2012

[9] FEI Color face database

[10] B. Schölkopf, A. Smola, and K. R. Müller, “Nonlinear component analysis as a kernel eigenvalue problem,” Neural computing Journal, vol. 10, no. 5, pp. 1299-1319, 1998.

[11] G. Jarillo, W. Pedrycz, and M. Reformat, “Aggregation of classifiersbased on image transformations in biometric face recognition,”Machine Vision and Applications J., vol. 19, pp. 125–140, 2008.

[12] T. Ahonen, A. Hadid, and M. Pietikainen, “Face description with local binary patterns: Application to face recognition,” IEEE Transactions onPattern Analysis and Machine Intelligence, vol. 28, no. 12, pp. 2037–2041, 2006.

[13] A. Petpon and S. Srisuk, “Face recognition with local line binary pattern,” in Proceedings of Fifth International Conference on Image andGraphics, 2009, pp. 533–539.

[14] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, “Local binary patterns and its application to facial image analysis: A survey, ”IEEE Transactions on Systems, Man, and Cybernetics, Part C:Applications and Reviews (TSMC-C), vol. 41, no. 6, pp. 765–781, 2011.

[15] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Z. Li, “Learning multi-scale block local binary patterns for face recognition,” in Proceedings ofInternational Conference on Advances in Biometrics, ICB 2007, Seoul,Korea, August 27-29, 2007, ser. Lecture Notes in Computer Science, S.- W. Lee and S. Z. Li, Eds., vol. 4642. Springer, 2007, pp. 828–837.

[16] Ruiz-del-Solar, J.; Navarrete, P., "Eigenspace-based face recognition: a comparative study of different approaches," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on , vol.35, no.3, pp.315,325, Aug. 2005.

[17] Chen-Ta Hsieh; Chin-Chuan Han; Chang-Hsing Lee; Kou-Chin Fan, "Pattern Classification Using Eigenspace

Projection," Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2012 Eighth International Conference on , vol., no., pp.154,157, 18-20 July 2012

[18] L. SirovichandM.Kirby, "A low-dimensional procedure for the characterization of human faces," J. Opt. Soc. Amer.A, vol. 4, no. 3, pp. 519–524, 1987.

[19] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cogn.Neurosci., vol. 3, no. 1, pp. 71–86, 1991.

[20] Ruiz-del-Solar, J.; Navarrete, P., "Eigenspace-based face recognition: a comparative study of different approaches," in Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on , vol.35, no.3, pp.315-325, Aug. 2005

[21] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cogn.Neurosci., vol. 3, no. 1, pp. 71–86, 1991.

[22] R. A. Fisher, "The use of multiple measures in taxonomic problems," Ann. Eugenics, II, vol. 7, pp. 179–188, 1936.

[23] C. Liu and H. Wechsler, "Evolutionary pursuit and its application to face recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 6, pp. 570–582, Jun. 2000.

[24]Hong Wang; Su Yang; Wei Liao, "An Improved PCA Face Recognition Algorithm Based on the Discrete Wavelet Transform and the Support Vector Machines," in Computational Intelligence and Security Workshops, 2007.CISW 2007. International Conference on , vol., no., pp.308-311, 15-19 Dec. 2007.

[25] R. Naik, K. Lad "Face Recognition from Multi Angled Images" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 IJERTV4IS020537 www.ijert.org Vol. 4 Issue 02, February-2015.