Feature tracking and extraction of video images using Wavelet transform and PCA

[1]Gajanan Joshi, [2]Prakash Khanale
[1]Vidya Pratishthans Arts, Science & Commerce college, Baramati,
[2]DSM College, Parbhani.

Abstract: Today the word is moving towards the economic process in area of biometrics as a personal identification methodology. The techniques that are established for an distinctive the individual using face as a biometric has become additional importance in field of biometrics. The face database extracted leads the numerous application like photography, security surveillance, database identification etc. This paper includes the study of facial feature extraction techniques that are Principal part Analysis (PCA) and discrete wavelet Transforms (DWT), hear the comparison of two given algorithms are created with concerned to the rate of feature extraction for face recognition using the Principal part Analysis (PCA) and therefore the PCA using discrete wavelet Transforms (DWT). The projected algorithm uses the thought of DWT for the image compression and PCA for the feature extraction and identification methodology. The limitations of the sole PCA rule are a poor recognition speed and complex mathematical calculating load. To eliminate these limitations we are applying the DWT with completely different decomposition levels, i.e. from level zero to level three to facial image by using Daubechies transform and applying the PCA for feature extraction method. The euclidean Distance Measures system is used to notice the closest matching features within the whole database. In this paper the mentioned algorithms are compared with their feature extraction and recognition time, the second parameter is that the percentage of recognition of a test image. The results shows that the PCA with DWT applied gives higher recognition rate up to 93 than solely PCA ,with very less access time.

IndexTerms **- Face Recognition (FR), Discrete Wavelet Transform (DWT), Principal Components Analysis (PCA).**

I. Introduction

Face recognition has recently received significant attention (Zhao et al. 2003 and jain et al. 2004). It plays a vital role in several application areas, such as human-machine interaction, authentication and surveillance. However, the wide-range variations of human face, thanks to cause, illumination, and expression, end in a extremely complex distribution and deteriorate the recognition performance. Additionally, the matter of machine recognition of human faces continues to draw in researchers from disciplines such as image process, pattern recognition, neural networks, computer vision, computer graphics, and psychology. A general statement of the matter of machine recognition of faces may be formulated as follows: Given still or video images of a scene, determine or verify one or many persons within the scene employing a stored database of faces. In identification problems, the input to the system is an unknown face, and the system reports the determined identity from a database of known individuals, whereas in verification issues, the system needs to confirm or reject the claimed identity of the input face[1].

The solution to the matter involves segmentation of faces (face detection) from cluttered scenes, feature extraction from the face regions, recognition or verification. Robust and reliable face illustration is crucial for the effective performance of face recognition system and still a challenging problem. Feature extraction is complete through some linear or nonlinear transform of the data with subsequent feature choice for reducing the dimensionality of facial image in order that the extracted feature is as representative as possible. Wavelets have been successfully utilized in image processing. Its ability to capture localized time-frequency information of image motivates its use for feature extraction. The decomposition of the info into completely different frequency ranges allows us to isolate the frequency components introduced by intrinsic deformations due to expression or extrinsic factors (like illumination) into certain subbands. Wavelet-based methods prune away these variable subbands, and focus on the subbands that contain the most relevant information to better represent the data[1].

II. Steps involved in face recognition

- **Face detection/segmentation**

This section is concerned about the literature using the so-called “Top-Down” approach for Face Recognition, meaning a face is detected initial then its features are strained within the segmented face-like region. The most recent (State of the art) approach that attracted several researchers, is initial mentioned.

a) Based on RGB data

“Most strategies of color image analysis don't differ significantly from those applied to gray-scale images, they simply entail application of identical strategies as those used for one graylevel image, however applied threefold to the various color images.”

(M.Seul et al., 2000, p 52). A special transformation map called (IHS), that stands for Intensity, Hue and Saturation will be obtained from the RGB bases, see figure 3.

- *Intensity is a measure of brightness:* $I=(R+G+B)/3$ *……….(1)*
- *Hue represents the color value:* $H=cos-1\{[(R-G)+(R-B)]/2[(R-G)2+(R-B)(G-B)]-1/2\}$ *………… (2)*
- *Saturation refers to the depth of the color:* $S=1-min(R,G,B)/I$ *………… (3)*

YCbCr is another transformation that belongs to the family of television transmission color spaces. Skin detection for face location in color images has benefited from these. R. Hsu et al. (2002) introduced a skin detection formula that starts with lighting compensation "reference white" which may be chosen from the top 5-hitter of the luma if the add of its pixels (>100). They detect faces based on the cluster in the (Cb/Y)-(Cr/Y) mathematical space. A sample of their result is shown next figure 1. T. chang et al. (1994) followed identical approach, while H. Wang and S. chang (1997) select the following system to convert form (RGB) to (Y, Cb, Cr):

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} (0.299)(0.587)(0.114) \\ (-0.169)(-0.331)(0.500) \\ (0.500)(-0.419)(-0.081) \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

b) Based on Boundary

Frequently, a very important visual element considered in image segmentation is that the distinction between the face region and its background. The approaches for detection such high distinction regions are referred to as edge detection operators. From reviewing the literature it's found that the foremost relied on operators are Sobel and Canny. The latter is proven to be the most efficient among all since it detects strong as weak edges and minimizes the noise, while it consumes more computational time than others do.

c) Based on eigen faces

In K. Wong et al. (2001), a pair of eye candidates are selected by means of the genetic algorithm to form a possible face candidate. The fitness value of every candidate is measured based on its projection on the eigenfaces. In order for them to enhance the level of detection reliability, each possible face region is normalized for illumination. once a number of iterations, all the face candidates with a high fitness price area unit selected for further verification. Turk and Pentland (1991) earlier developed this method for face recognition. Their method exploits the distinct nature of the weights of eigenfaces in individual face representation. Since the face reconstruction by its principal components is an approximation, a residual error is defined in the algorithm as a preliminary measure of "faceness" This residual error that they termed "distance-from-face-space" (DFFS) provides a good indication of face existence through the observation of global minima in the distance map. each PCA vector is called eigenvector, and when converted back to matrices these vectors will viewed because the eigenfaces of the dataset figure below[3].

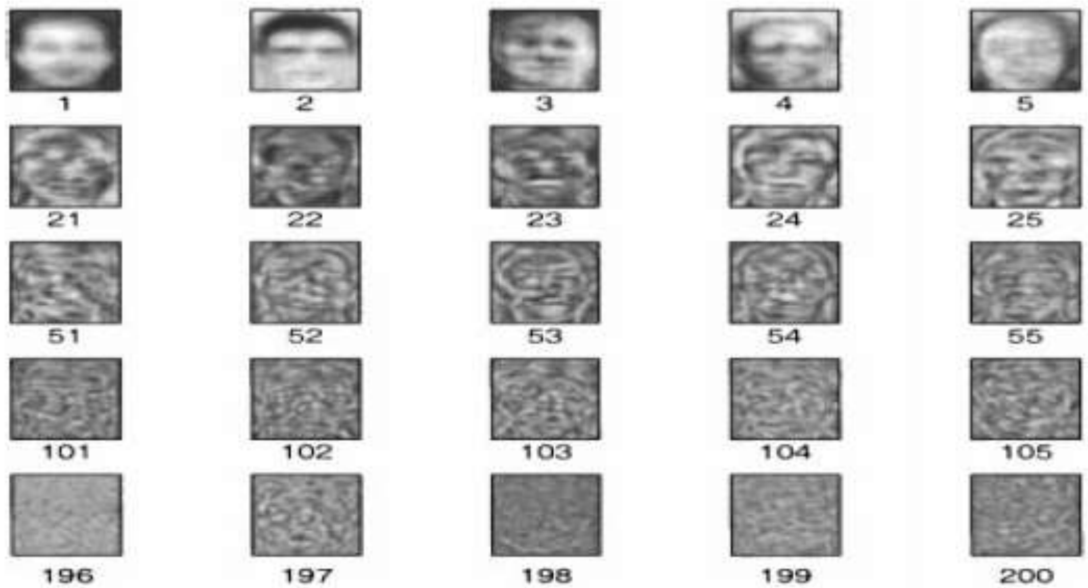


Figure 1 Some examples of eigenfaces computed from the ORL dataset (the number below each image indicates the principal component number, ordered according to eigenvalues). (E. Hjelmas, 2001)

Eigenfaces and Color based approaches are among the most well known methods, and they are the state of the art in this field. Therefore, a comparison between the two is shown next (table 1).

Criterion	**Skin color based approach**	**Eigen-Faces approach**
Handling more than one face	It detects each face	It fails
Rotation, profile and tilted face	It is dependent on the skin color not on the orientation	In case if the eigen space contains such information then it works
Size of the face	Not necessarily, since neck can be included	It depends on the size of the eigenface size
Complex background	If it has color similar to the skin color than it suffers	Usually it handle this situation

Table 1. A comparison between Skin color and Eigen based approaches

- **Preprocessing**

The first step in any face recognition system is Face Detection. Several face detection techniques are proposed within the past decade. They will be classified into geometry-based face detectors and color-based face detectors. Among the geometry-based face detectors, a method examines the triangle-relationship between eye and mouth regions to identify the face region. Additionally to, the normal eye detection methods are often merely and efficiently implemented for frontal face pictures however can be difficult for complicated images. Moreover, skin color has been proven a good image feature for face detection. The automated face detectors supported skin color square measure comparatively quick and so are beneficial for consumer applications such as real-time face recognition embedded in a very smart home environment. In the preprocessing stage of the proposed system, a facial region based on skin color detection is cropped from an input image. The obtained facial region is then resized into an 8×8 pixel image to make the face recognition system scale invariant. Then after, histogram equalization is applied to enhance the image brightness and contrast.

- **Cropping of images**

For the purpose of aligning and cropping detected faces, it is therefore important to understand the following:

- The centre of the box is deemed to be the centre of the face.
- The size of the box is deemed the size of the face.

The purpose of the face detection algorithm is to then enable automatic cropping of images so that faces are a consistent size and position. To this, end three adjustments are available.

It is not necessary to detect faces before cropping. To perform both steps in one cation simply select the images the click 'Detect and Crop'

- Align faces to the left or right
- Align faces up or down

- Size faces

The three parameters above are set with the corresponding tools on the Face Detect toolbar and are expressed as percentages of the total canvas.
To crop the images to the entered values, select the images first then click Detect and Crop.

- **Histogram equalization**

Histogram equalization is a technique for adjusting image intensities to enhance contrast. Let f be a given image represented as a mr by mc matrix of integer pixel intensities ranging from 0 to L − 1. L is the number of possible intensity values, often 256. Let P denote the normalized histogram of f with a bin for each possible intensity. So

$$p_n = \frac{\text{number of pixels with intensity } n}{\text{total number of pixels}} \qquad n = 0, 1, ..., L-1.$$

The histogram equalized image g will be defined by

$$g_{i,j} = \text{floor}((L-1)\sum_{n=0}^{f_{i,j}} p_n),$$

Where floor() rounds down to the nearest integer. This is equivalent to transforming the pixel intensities, k, of f by the function

$$T(k) = \text{floor}((L-1)\sum_{n=0}^{k} p_n).$$

The motivation for this transformation comes from thinking of the intensities of f and g as continuous random variables X, Y on [0, L−1] with Y defined by

$$Y = T(X) = (L-1)\int_0^X p_X(x)dx,$$

Where pX is the probability density function of f. T is the cumulative distributive function of X multiplied by (L−1). Assume for simplicity that T is differentiable and invertible. It can then be shown that Y defined by T(X) is uniformly distributed on [0, L−1], namely that

$$\begin{aligned} \int_0^y p_Y(z)dz &= \text{ probability that } 0 \le Y \le y \\ &= \text{ probability that } 0 \le X \le T^{-1}(y) \\ &= \int_0^{T^{-1}(y)} p_X(w)dw \end{aligned}$$

$$\frac{d}{dy}\left(\int_0^y p_Y(z)dz\right) = p_Y(y) = p_X(T^{-1}(y))\frac{d}{dy}(T^{-1}(y)).$$

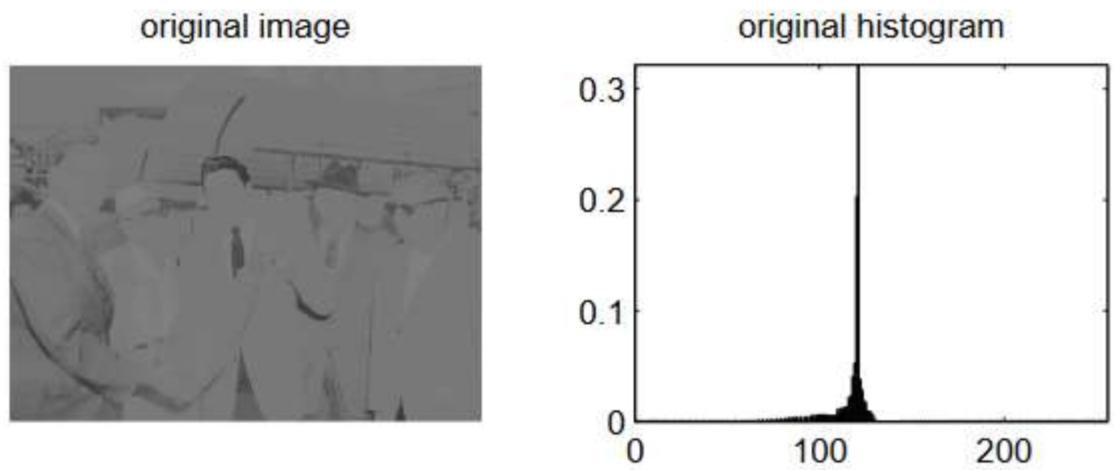


Figure 2 Histogram of image

- **Use of wavelet transformation**

Recently, within the field of image process the DWT becomes a powerful tool for practical study of the facial image primarily based applications. The image compression and the multiresolution of the image are the two dependent factors for the image processing. Multiresolution and therefore the compression provide the way to generate the decomposed data. The image compression could leads to increase the storing capacity of the images. The DWT is that the extremely economical and flexible method to decompose the image into completely different sub-bands or decomposed levels. This sub bands or the idea also are known as the wavelets. Hear, the DWT decomposes the image within the wavelet coefficients. These coefficients are helpful for the feature extraction and compression of the image. In our study, we are going to implementation of 3-wavelet sub bands using the Daubechies transform[2].

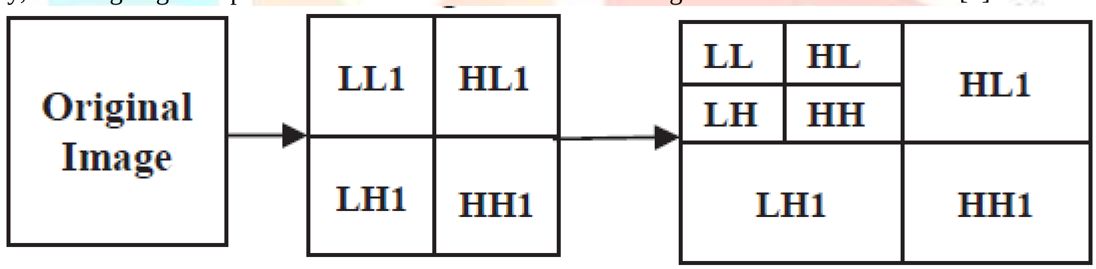


Figure 3 Decomposition at 3-level Wavelet

In this, the initial facial image is decomposed into three wavelet levels by victimization Daubechies Transform as shown in Fig above. The primary band that's LL could be a approximated compressed of the initial facial image. The bands lh and hl record the edges on horizontal and vertical directions severally. Whereas the HH band records, the diagonal edges of the image. This can be the primary level decomposition. Further 3-level decomposition are often conducted on the LL sub-band. During this paper, we preferred the Daubechies transform for the transformation of image in to the 3- level wavelets.

The following are applications of wavelet transforms:

- data and image compression
- partial differential equation solving
- transient detection
- pattern recognition
- texture analysis
- noise/trend reduction

- **PCA for Eigenvalues – Eigenvector**

Principal component Analysis is planned by turk and pent land in 1991, that is commonly used for extracting features of image. Principal element Analysis is most widely used method considering with the facial feature extraction in image processing. Basic plan behind the PCA is, the set of pictures area unit initially transformed into Eigenfaces i.e. lower information space by exploitation the K-L transform method. This method includes the linear transformation of the higher data space into the lower information area using linear transformation method. This extracted lower dimensional image preserves the most of the data} or information from the initial higher dimensional facial image. This mapped lower information space is called as the Eigenface. Then the test Eigenfaces vector from the database is projected on the trainee Eigenfaces vector to get the correct match[3]. For PCA, two-dimensional image matrix should be first transformed to a one-dimensional vector with high order. While the number of training sample is little, it is very difficult to calculate covariance matrix of training sample accurately. Moreover, structure information are lost during processing. The Eigenfaces vector as thought-about as the vector for constructing the covariance matrix .Hear, the pixel information of every image is used to construct the eigen vector. This eigen vector information is used to select the Principal component having the higher eigen value. Each image location Contributes to every eigen vector, so that we can show the eigen vector as a form of face. Every face image can be represented exactly in terms of linear combination of the eigen faces. The quantity of attainable Eigen faces is equal to the quantity of face image within the training set. The faces will also be approximated by using best eigen face, those that have the largest eigen values, and that therefore account for many variance between the set of face images. The primary reason for using fewer eigen faces is procedure efficiency[4].

A. Brief Discussion of PCA-

1) First, consider the set of images in the column matrix or the row matrix format, named A

A=(I1,I2 ,I3 ,I4 ,… … … .IM)

Where, M is the total number of objects present in total database.

2) find the average of the defined matrix A

$$\mu = \frac{1}{M} \sum_{n=1}^{M} I_n$$

Hear, n =is the total number of images in single object of Database μ =Mean of the defined matrix A

3) Then find the differential distance between the trainee images and the mean calculated

$$\alpha = I_i - \mu$$

Hear, we will get the α, scatter matrix for each image.

4) find the covariance matrix C as follows,

$$C = \frac{1}{M} \sum_{n=1}^{M} \alpha_n \alpha_n^T$$

5) Find the eigenvector and eigenvalue of the covariance matrix C. These eigenvectors are arranged in descending order and the weighted vector is selected having the highest eigenvalue for the feature extraction.

6) This feature vector is consist of the extracted data of all the images present in database and is compared with the vector of test image[3].

- **Feature Extraction**

For gaining the best feature vector from the training dataset, at first, all images are normalized. The following steps are performed for feature extraction.

- RGB image is converted into gray scale image and resized it 64x64 pixels.
- Filtering is applied to the image for smoothing and removing the noise.
- 2-D Discrete Wavelet Transform of first level is applied on image. “Bior 3.7” is used as a base wavelet of the images. Each image is decomposed in to the Approximation, Horizontal, Vertical and Diagonal Coefficients. These coefficients are used to reorder the column vector Ii of the images.
- The Feature image matrix I= [I1, I2 , I3 ..... IP] is constructed from the coefficients column vector Ii. Where i represent the no of image.
- This Feature matrix I is transformed to lower dimension subspace Tw using PCA.
- Tw consists of Weight calculated for each image of the respective Dataset[6].

Methods of Feature Extraction:

These options square measure usually a synthesis of low-level features for a selected domain.

1 Color:

The color feature is one of the most wide used visual options in image retrieval. Images characterized by color options have several advantages:

• Hardiness: The color bar chart is invariant to rotation of the image on the read axis, and changes in little steps once turned otherwise or scaled. It is additionally insensitive to changes in image and bar chart resolution and occlusion.

• Effectiveness: There is high proportion of relevance between the question image and the extracted matching pictures.

• Implementation simplicity: The development of the color bar chart may be an easy process, together with scanning the image, distribution color values to the resolution of the bar chart, and building the bar chart exploitation color elements as indices.

2 Texture:

Texture is another vital property of pictures. Texture is a powerful regional descriptor that helps within the retrieval method. Texture, on its own doesn't have the aptitude of finding similar images, however it will be used to classify unsmooth images from non-textured ones and then be combined with another visual attribute like color to create the retrieval simpler. Texture has been one of the most vital characteristic that has been used to classify and recognize objects and have been used in finding similarities between pictures in multimedia system databases.

Basically, texture illustration strategies will be classified into 2 categories: structural; and statistical. Applied math strategies, together with Fourier power spectra, co-occurrence matrices, shift-invariant principal element analysis (SPCA), Tamura options, World decomposition, Markov random field, form model, and multi-resolution filtering techniques cherish Dennis Gabor and ripple transform, characterize texture by the applied math distribution of the image intensity.

3 Shape:

Shape primarily based image retrieval is the mensuration of similarity between shapes delineate by their features. form is associate vital visual feature and it is one of the primitive options for image content description. form content description is difficult to outline as a result of mensuration the similarity between shapes is tough. Therefore, two steps square measure essential in form primarily based image retrieval; they are feature extraction and similarity measure between the extracted options.

Shape descriptors will be divided into 2 main categories: region primarily based and contour-based methods.

Region-based strategies use the complete space of associate object for form description, whereas contour-based methods used solely the knowledge gift within the contour of associate object[5][7].

III. SYSTEM DESCRIPTION

Generally, feature extraction and classification criterion are the two basic operations of any face recognition system. As a result, to improve the recognition performance of such systems one has to enhance these operations. Prior to classification, the features should be extracted from the human face images. Feature extraction in the sense of some linear or nonlinear transforms of the face images with subsequent feature selection is commonly used for reducing the dimensionality of these images so that the extracted features are as representative as possible. However, the problem of extracting features from a human face remains a barrier to apply the practical applications, since the lighting condition, illumination changes, various backgrounds and individual variations influence it.

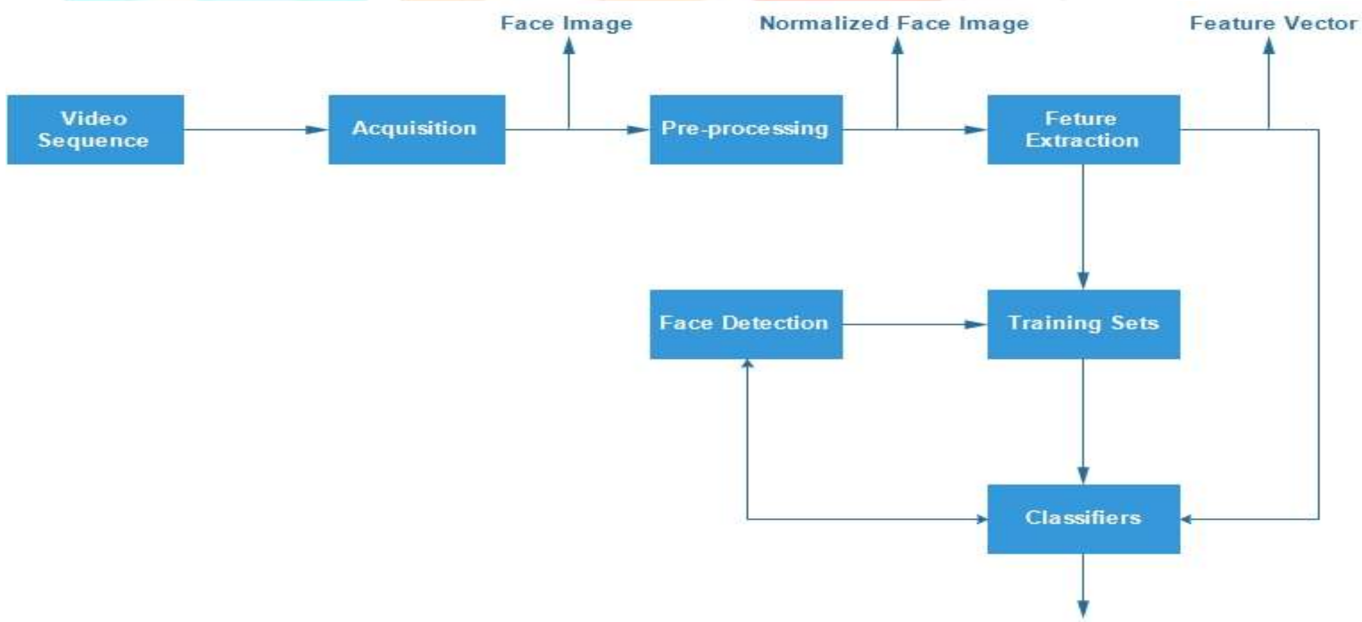


Figure 4 Block diagram

Euclidian Distance Classifier.

In this classification method, each image is transformed to a lower order subspace by wavelet-PCA using the above steps. Upon observing an unknown test image X, the weights are calculated for that particular image and stored in the vector WX. WX is compared with the weights of training set Tw using the Euclidian distance. If average distance exceeds some threshold value, then the weight vector of the unknown image WX lies ”far apart” from the weights of the training images. The optimal threshold value has to be determined empirically[9].

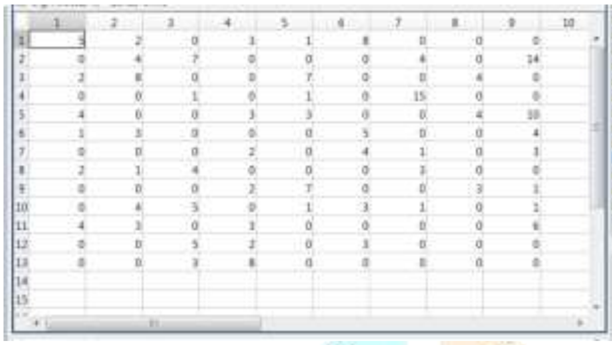
IV. EXPERIMENTAL WORK

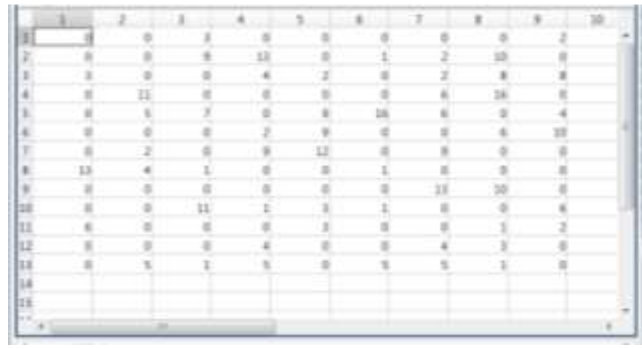
MATLAB 2013a is used for face extraction. Standard video sequence from ’NRC-IIT Facial video database and video captured from camera are taken as an input.

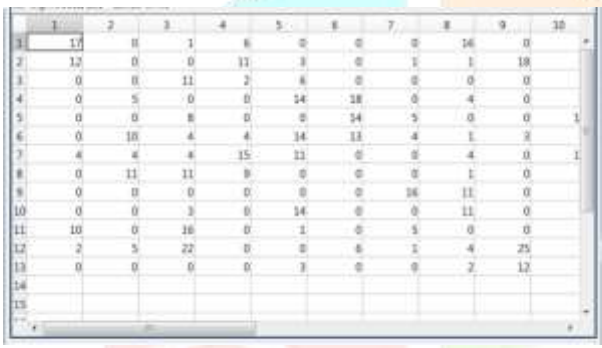
From input video, we get different Eigenfaces that was shown below from that we get average face value.

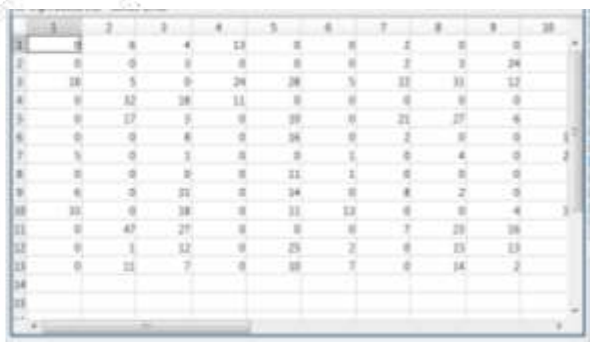
After get face average value, we are going to apply skin detection process on that image. After that, we are going to apply face detection process on that skin-detected part. Because of that, we able to detect face from that image. After getting face coordinates we are applying segmentation method on that, as shown in Face1 figure below. Now we convert that segmented image to grayscale image because of this we can find histogram equalization for adjusting image intensities to enhance contrast. At last, we had applied wavelet decomposition method on that image for data and image compression, partial differential equation solving, transient detection, pattern recognition, texture analysis, noise/trend reduction.

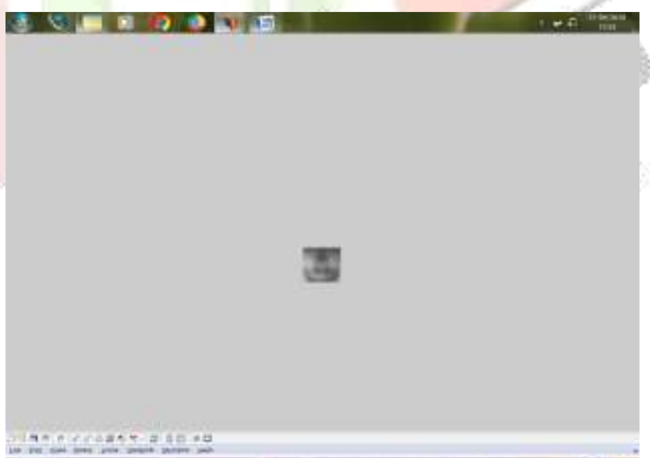
Different result of each stages shown in below figures.


Eigenface 1

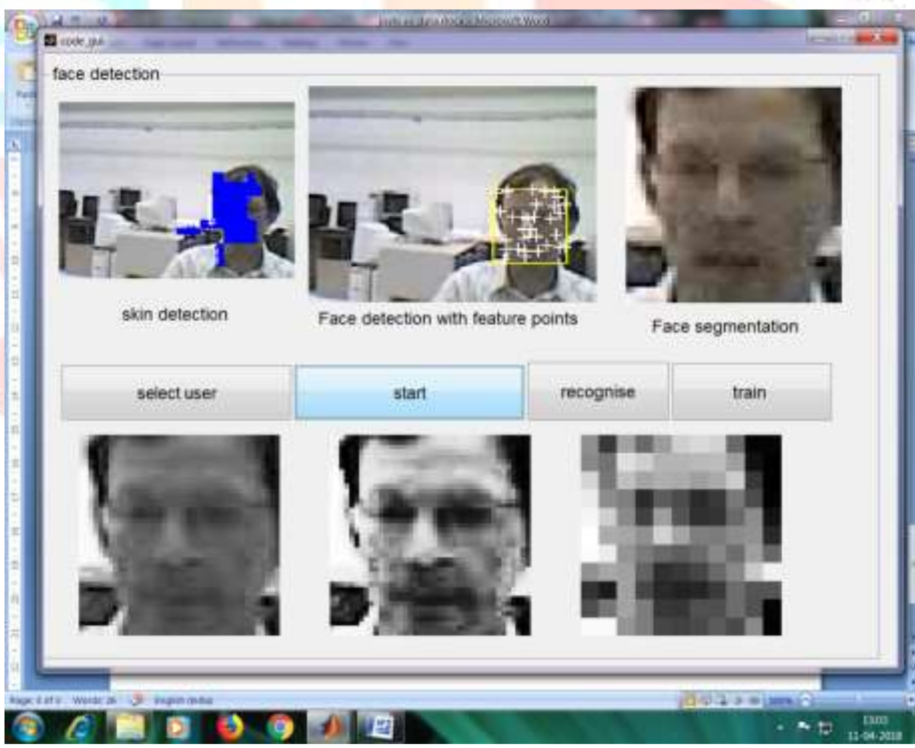

Eigenface 2


Eigenface 3

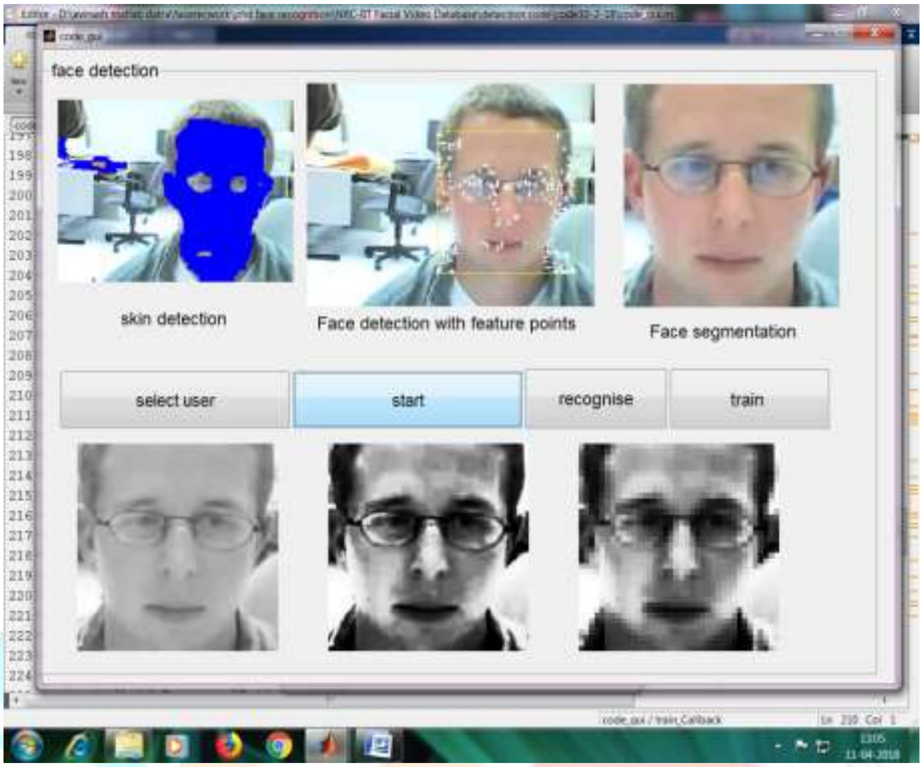

Eigenface 4


Average face

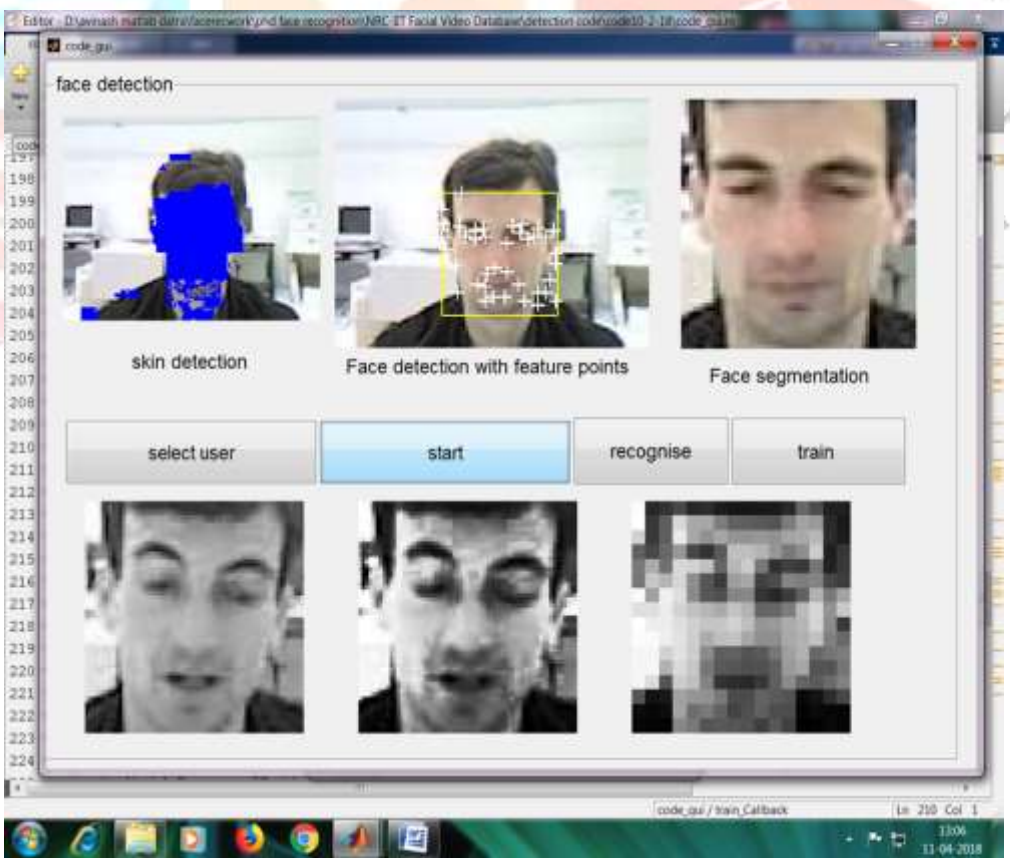
Face 1: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition:



Face 2: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition



Face 3: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition



Face 4: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition

From Real time database, following result generated:

	1	2	3	4	5	6	7	8	9
1	15	0	12	0	0	2	4	12	16
2	0	0	0	0	13	21	0	0	0
3	0	16	0	0	0	0	6	9	16
4	0	19	3	13	0	0	12	0	0
5	0	15	17	13	0	29	0	0	3
6	0	0	10	0	8	0	1	7	0
7	17	2	0	0	0	16	0	15	6
8	9	26	0	0	0	10	7	20	0
9	21	0	0	0	0	7	16	13	0
10	0	0	0	0	0	0	15	21	3
11	11	0	0	9	0	0	0	3	2
12	0	25	0	5	0	0	0	0	0
13	0	8	0	21	0	0	3	0	4
14									

Eigenface 1

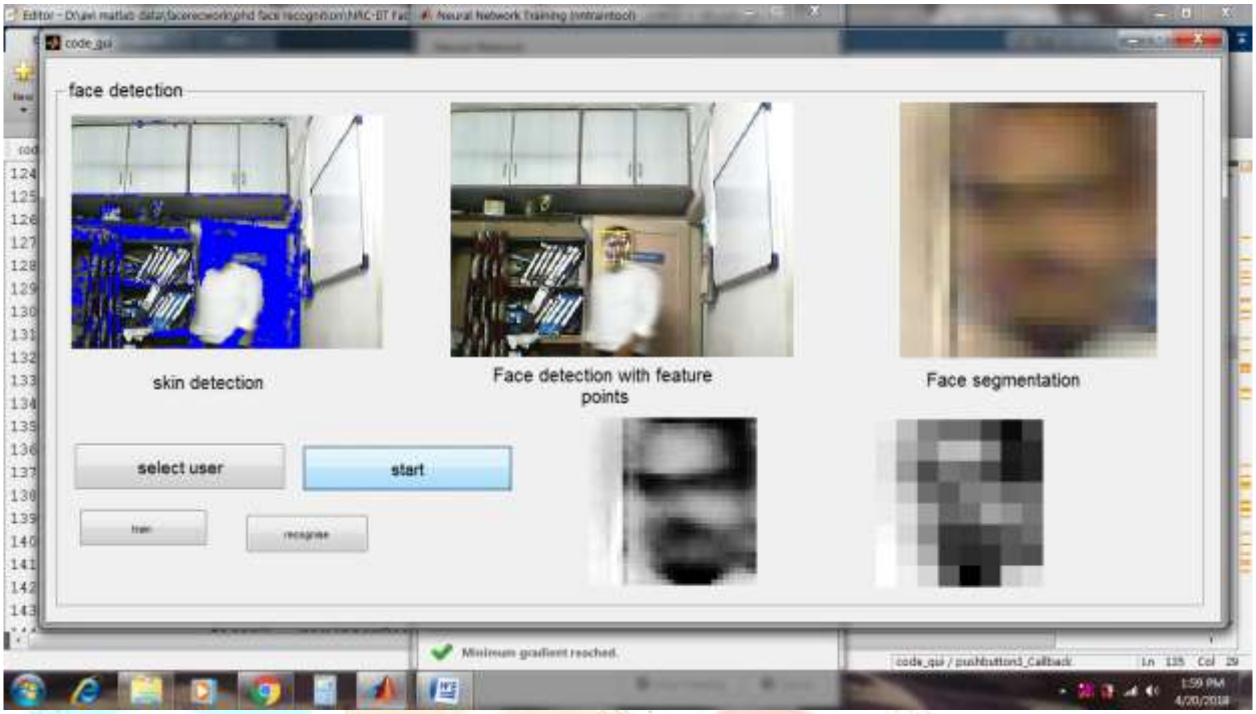
	1	2	3	4	5	6	7	8	9
1	0	0	0	0	0	13	11	0	2
2	0	24	0	1	0	0	0	0	0
3	0	17	32	0	0	0	12	16	14
4	0	0	4	0	0	0	22	13	14
5	0	11	24	3	21	0	0	0	0
6	0	14	1	0	30	0	0	12	21
7	0	2	0	0	18	22	11	0	11
8	0	0	0	0	0	0	0	1	10
9	0	0	13	0	0	0	0	4	0
10	0	12	23	0	0	0	0	0	0
11	0	6	0	0	0	0	0	0	0
12	9	6	1	0	0	0	0	6	12
13	2	9	5	0	21	27	27	18	14
14									

Eigenface 2

	1	2	3	4	5	6	7	8	9
1	0	36	155	245	226	150	117	136	177
2	0	103	245	55	0	0	24	13	1
3	0	60	106	7	0	0	0	0	0
4	0	21	49	0	0	0	0	0	0
5	17	0	0	0	26	42	0	0	0
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	2	0	8	0
9	0	0	0	0	0	0	0	32	43
10	0	0	0	0	0	0	0	20	64
11	0	0	31	0	0	0	12	15	0
12	3	69	30	0	0	0	0	0	0
13	51	77	30	0	0	0	0	0	0
14									

Eigenface 3

Face 1 form real time data: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition



Face 2 form real time data: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition

Face 3 form real time data: Skin detection--face detection—face segmentation—gray conversion—histogram equalization—wavelet decomposition

CONCLUSION

A face recognition approach supported DWT and PCA has been planned. It eliminates the necessity to manually phase input pictures and creates a fashionable feature vector, based on LL quadrant DWT coefficients and L2-norm values from other quadrants. It optimizes the feature vector supported the discriminatory power of every constant. Recognition results show that this approach considerably outperforms recognition using normal DWT/PCA recognition using LL options only.

The performance of the two algorithms were compared and also the performance of DWT-PCA is more than that of Eigenface alone since it converge with comparatively constant recognition rate however the previous converge quicker than the latter. Additionally the number of levels in DWT affects potency.

REFERENCES

[1] Smt.M.P.Satone, Dr. G.K.Kharate "Face Recognition Based on PCA on Wavelet Subband,"2012 IEEE Students' Conference on Electrical, Electronics and Computer Science.

[2] Paul Nicholl, Afandi Ahmad, and Abbes Amira, "Optimal Discrete Wavelet Transform (DWT) Features for Face Recognition," 2010 IEEE

[3] Kishor S Kinage and S. G. Bhirud, "Face Recognition based on Two-Dimensional PCA on Wavelet Subband," International Journal of Recent Trends in Engineering, Vol 2, No. 2, November 2009

[4] Muzammil Abdulrahman, Yusuf G. Dambatta, A. S. Muhammad and Abubakar S. Muhammad, "Face Recognition Using Eigenface and Discrete Wavelet Transform," 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009, pp. 1-6

[5] Dipalee Gupta1, Siddhartha Choubey2, "Discrete Wavelet Transform for Image Processing", International Journal of Emerging Technology and Advanced Engineering Volume 4, Issue 3, March 2015.

[6] Z. M. Hafed and Martin D. Levin, "Face Recognition Using the Discrete Cosine Transform", International Journal of Computer Vision, 43(3), 2001, pp 167-188.

[7] C. Cruz, L.E. Sucar, and E.F. Morales, "Real-Time Face Recognition for Human-Robot Interaction," 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, Sep. 2008, pp. 1-6.

[8] P. Michel and R. El Kaliouby, "Real Time Facial Expression Recognition in Video Using Support Vector Machines," Proceedings of the 5th international conference on Multimodal interfaces - ICMI'03, 2003, p. 258.

[9] F. Perronnin and J.-L. Dugelay, "An introduction to biometrics and face recognition," In: Proc. Of IMAGE 2003: Learning, Understanding, Information Retrieval, Medical, Cagliari, Italy, pp 1-20, June 2003.

[10] Swati Jadon et. Al.,"Face Recognition Using SOM Neural Network with DDCT Facial Feature Extraction Techniques", 4th IEEE International Conference on Communication and Signal Processing (ICCSP 2015) 02-04 April 2015 - Melmaruvathur, TN, IND.

[11] NRC-IIT Facial video database.