Biometric Authentication based on Detection and Recognition of Multiple Faces in Image

Jolly D. Shah and S.H. Patil

Computer Engineering Department, Bharati Vidyapeeth College of Engineering,
Bharati Vidyapeeth University, Pune
E-mail : Jolly.M.Shah@gmail.com

ABSTRACT

Biometric Authentication works on some human characteristics, such as, finger print, voice, pattern of lines in the iris of eye or face. I represent a multiple face detection method based on skin color information and the Lines of Separability (LS) face model and recognition method based on principle component analysis and neural network. Face detection method uses YCbCr color model and sigma control limits for variation in its color components. Faces are extracted from the images. The face space is described by a set of eigenfaces. Each face is efficiently represented by its projection onto the space expanded by the eigenfaces and new descriptor. After face detection stage neural network are used for face recognition. A system is implemented which detects multiple faces in an image and recognizes it.

Keywords: *Face detection, Face recognition, Sigma control limits, Skin color space, Principal component analysis, Neural network and Eigenface.*

1. INTRODUCTION

To determine people's identity, the obvious question is what technology is best suited to supply this information? There are many ways that humans can identify each other, and so is for machines. There are many different identification technologies available, many of which have been in commercial use for years. The most common person verification and identification methods today are Password/PIN known as Personal Identification Number, systems. The problem with that or other similar techniques is that they are not unique, and is possible for somebody to forget, loose or even have it stolen for somebody else. In order to overcome all these problems there has been considerable interest in "biometrics" identification systems, which use pattern recognition. some of these methods are fingerprints, retina and iris recognition technique; though these techniques are not easy to use. For example in bank transactions and entry into secure areas, such technologies have the disadvantage that they are intrusive both physically and socially. The user must position the body relative to the sensor, and then pause for a second to declare himself or herself. That does not mean that face recognition does not need specific positioning. As we are going to analyze later on the poses and the appearance of the image taken is very important. While the pause and present interaction are useful in high security, they are exactly the opposite of what is required when building a store that recognize its best customers, or an information kiosk that remembers you, or a house that knows the people who live there. Face recognition from video and voice recognition have a natural place in these next generation smart environments, they are unobtrusive, are usually passive, do not restrict user movement, and are now both low power and inexpensive. Perhaps most important, however, is that humans identify other people by their face and voice, therefore are likely to be comfortable with systems that use face.

The developing of a face detection and recognition system is quite difficult because human face is quite complex, multidimensional and is subject to variations because of environment changes, aging etc. For that reason human face recognition is a challenging problem because of changes in the face identity and variation between images of the same due to illumination and viewing direction. The issues are how the features adopted to represent a face under environmental changes and how we satisfy new face image based on the chosen representation. Computers that recognize human face systems have been applied in many applications such as security system, mug shot matching and model based video coding. The problem can be described as following. Given an image of human face find out who it is if match exists with database model otherwise no match occurs. I divide this problem in two stages. In the very first stage, face which is located in the images is detected. This system detects single face as well as multiple faces in image. After detection of face, in next stage face is identified using neural network architecture. Here image is vertically oriented frontal view in gray scale. Normal expression variations are allowed and image is prepared under constant illumination.

In face detection stage, face detection algorithm for color images using a skin tone color model and facial features is presented. In this method, the skin color bias is corrected by a lighting compensation technique and then skin regions are detected. The face candidates based on spatial arrangement of these skin patches are generated for face detection. The problem of sensitivity of face detection algorithm to illumination conditions, under which the input image is capture, has been addressed. One of the aspects, from which illumination influence can be observed, is the choice of proper color space. A method for human skin color clustering in 2D chromatic space CbCr with luminance(Y) dependent chromatic components Cb and Cr has been proposed. Face is identified and then principle component analysis is applied to find aspects of face which are important for identification. Eigenvectors are calculated from initial face image set. Then new faces are projected onto space expanded by eigenfaces and represented by weighted sum of the eigenfaces. These weights are used to identify faces.

In face recognition stage, face database is created and face is recognized. In neural network approach, a separate network for each person is built. Very first, the detected face is projected onto eigenface space to get a new descriptor. This new descriptor becomes the input and applied to each person's network. The one with maximum output is selected and reported as host if it satisfies threshold. Now, each stage is described in detail in next sections. Procedure to detect and identify face is also explained. Fig. 1 shows flowchart to understand process of face recognition.

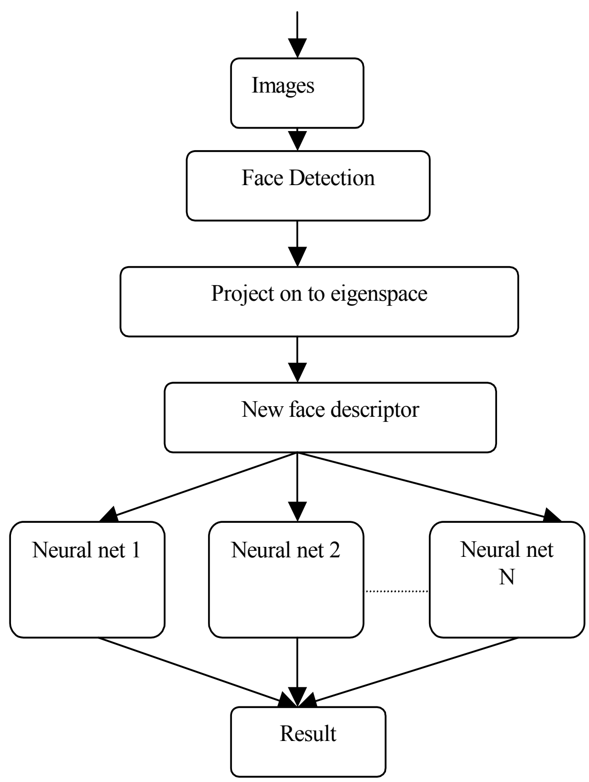

Fig. 1: Face recognition process flow

2. METHOD OF MULTIPLE FACE DEECTION IN IMAGE

Images are usually larger than the actual face. So, it is very important to detect face within image. I use YCbCr color model, which is widely used in digital video, to build proposed skin color space. It includes all possible skin color colors. I am able to extract more facial skin color regions excluding the nonskin regions using sigma control limits for color component variations. For experimentation 600 images are used from the CIT face database and the internet. The proposed skin color space uses only the chromatic color components Cb and Cr for skin color segmentation. The statistical sampling technique using sigma control limits are applied on the chromatic blue and red components.

The sample images are in RGB colors. The RGB color space represents colors with luminance information. Luminance varies from person to person due to different lighting conditions and hence luminance is not a good measure in segmenting the human skin color. The RGB image is converted into YCbCr color model using Eq. (1) where luminance is partially separated.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad ...(1)$$

Skin color space is developed by considering the large sample of facial skins cropped manually from the color face images of the multiracial people. Skin samples are then filtered using low pass filter to remove noises, lower and upper limits of the pixel values for the chromatic red and blue components are determines as below.

$$\mu_i = \left(\frac{1}{m \times n}\right) \sum_{x=1}^{m} \sum_{y=1}^{n} c(x, y) \;, \bar{\mu} = \frac{1}{k} \sum_{i=1}^{k} \mu_i \quad ...(2)$$

$$, \sigma = \sqrt{\frac{\sum_{i=1}^{k} (\mu_i - \bar{\mu})^2}{k}} \quad lcl = \bar{\mu} - 1.5\sigma, \; ucl = \bar{\mu} + 1.5\sigma$$

where μ_i denote the mean of the chromatic color components of the ith sample image c(x,y) of size m × n, where c denotes the color plane (i.e. red and blue). $\bar{\mu}$ and σ denote mean and standard deviation of the color components of the population of all k sample images, respectively. Lower and upper limits, *lcl* and *ucl* of the chromatic color components, respectively are used as threshold values for the segmentation of skin pixels.

From the binary skin segmented regions obtained by the proposed skin color space, the possible faces are detected by the process of facial features extraction such as eyes, nose, mouth and eyebrows using the proposed matching rules and face detection algorithm as follows. Each of the skin regions is preprocessed by the morphological operations to remove isolated noisy pixels and fill up holes if any. The preprocessed binary skin region is then multiplied with the luminance component of the original color image to crop the potential face area in gray scale as shown in Fig. 2. The gray scale skin region is filtered by Sobel edge operator and then binarized using global threshold which is then denoised by morphological operations making the essential facial features clearly visible in the skin region under consideration. The denoised image is labeled to group the active pixels into connected blocks. These blocks are candidate blocks of the facial features as shown in Fig. 2. With each labeled candidate block, its center of mass ($\bar{x}$, $\bar{y}$), orientation θ, bounding rectangle and the length of major axis are computed.

After that facial features are extracted like eye, mouth, eyrbrows, ears etc. The eyes are searched by selecting each pair of facial candidate blocks randomly as probable eyes and they are checked for orientation constraint. Evaluation value E_{Eye} for eye will be computed for all possible pair of blocks that satisfy the orientation constraint, and then a pair of blocks with maximum E_{Eye} will be considered as a potential eye candidate.

Initially we select any two features block randomly and assume them as probable eye candidates. Their corresponding center of mass

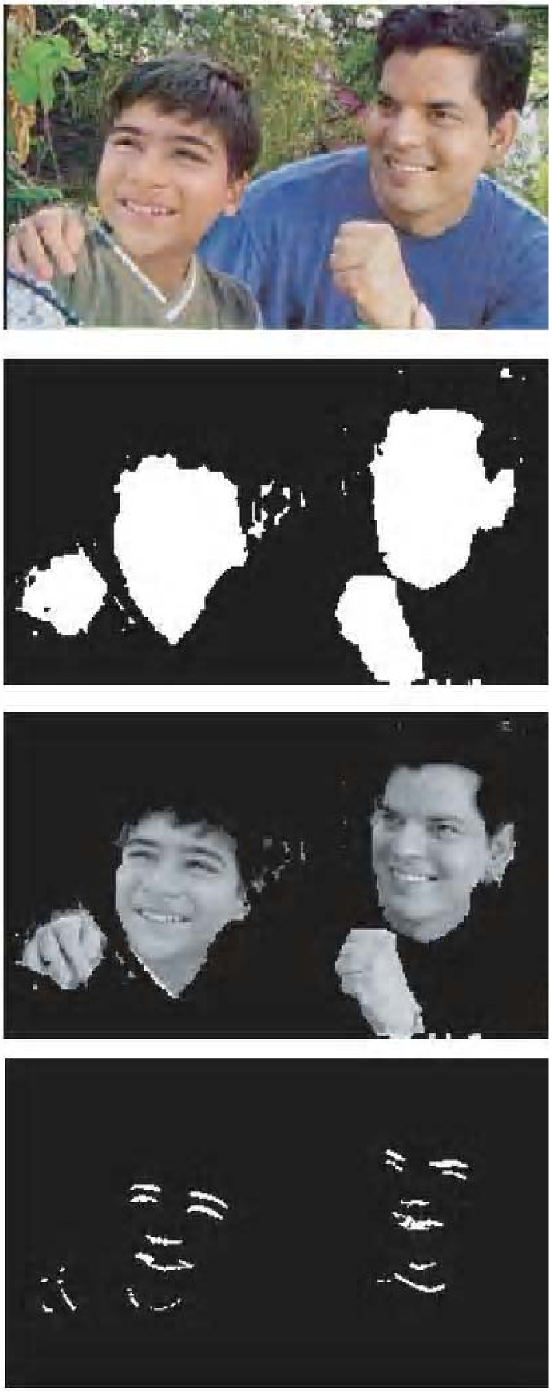

Fig. 2. Results of skin color segmentation: original image, skin segmented binary image, skin segmented gray scale image, and sobel filtered binary image respectively

($\bar{x}$, $\bar{y}$)and orientation θ are computed using the equations;

$$\overset{\rho}{x} = \frac{1}{N}\sum_{(x,y)\in B}\sum x \quad \bar{y} = \frac{1}{N}\sum_{(x,y)\in B}\sum y$$

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right) \qquad \text{...(3)}$$

Where (i, j) order central moments $\mu_{i,j}$ are given by:

$$\mu_{i,j} = \sum_{(x,y)\in B}\sum (x-\bar{x})^i (y-\bar{y})^j$$

and B denotes feature block.

Once the potential eyes are found, distance of the eyebrows, nose and mouth from the horizontal reference line (HRL) and vertical reference line (VRL) with respect to the distance (D) between eyes are estimated. Further LS face model is constructed using these estimated distance. Matching process proceeds to search other facial features like mouth, nose and eyebrows with respect to every potential eye pair candidate. For each pair of potential eye candidates, the relative regions of the other facial features are searched within the LS face model, using the matching rules and the face detection algorithm given below.

Step 1: Input the preprocessed binary image with segmented potential face regions.

Step 2: Choose a potential face region.

Step 3: Select any pair of feature blocks to be probable eye candidate blocks in the potential face region.

Step 4: Compute the slope angle

$$\theta_{HRL} = \tan^{-1}\left(-\frac{a}{b}\right) \qquad \text{...(4)}$$

$$-\pi/2 \le \theta_{HRL} \le \pi/2$$

using Eq. (4) and if it is between - 45° to + 45°, then compute evaluation value E_{Eye} using Eq. (5)

$$E_{Eye} = e^{[-1.2((l_1-l_2)^2+(l_1+l_2-1)^2+(\theta_1-\theta_{HRL})^2+(\theta_2-\theta_{HRL})^2)]} \qquad \text{...(5)}$$

If E_{Eye} is grater than the empirical threshold value 0.7, then above selected pair of feature blocks are accepted as the potential eye candidates. Further, with respect to these potential eye candidates, construct the LS face model by estimating the vertical and horizontal distance of all the other facial features, namely mouth, nose and eyebrows otherwise reject this pair of feature blocks and go to step 3.

Step 5: Search other potential features such as eyebrows, nose and mouth with respect to the above potential eye candidates using corresponding matching rules and compute their evaluation values E_{Leb}, E_{Reb}, E_{Nose} and E_{Mouth}.

Step 6: Compute the overall evaluation value

$$E = 0.4E_{Eye} + 0.3E_{mouth} + 0.2E_{Nose} + 0.05E_{Leb} + 0.05E_{Reb}$$

(5) using Eq. (5) with respect to potential eyes, eyebrows, nose and mouth with respect to the above potential eyes, eyebrows, nose and mouth found in steps 4 and 5.

Step 7: Repeat Step 3 through Step 6 for all other pairs of feature blocks assumed to be probable eye pair and find the maximum overall evaluation value E_{Max} from the set of E values computed in step 6.

Step 8: The potential eyes, eyebrows, nose and mouth corresponding to the maximum overall evaluation value E_{Max} constitute the most probable face in the chosen potential face region, provided E_{Max} is grater than the empirical threshold value 0.7. Otherwise, reject this potential face region and go to step 2.

Step 9: Repeat Step 2 through Step 8 for each potential face region in the input image in order to detect multiple faces present in the input image.

3. PRE PROCESSING

3.1 Image Size Normalization

It is usually done to change the acquired image size to a default image size on which the face recognition system operates.

3.2 Median Filtering

For noisy image especially obtain from cameras or from a frame grabber, median filtering can clean image without losing information.

3.3 Histogram Equalization

It is usually done on too dark or too bright images in order to enhance image quality and to improve recognition performance. It modifies the dynamic range of image and as a result some important facial features become more apparent.

3.4 Illumination Normalization

Face images taken under different illuminations can degrade recognition performance especially for face recognition system based on principal component analysis in which entire face information is used for recognition. The image is treated as a vector in the high dimensional space. Its vector length is adjusted to the vector length of average face in the face space.

4. FACE RECOGNITION

Principal component analysis for face recognition is based on the information theory approach. Here, relevant information in a face image is extracted and encoded as efficiently as possible. Recognition is performed on a face database that consists of models encoded similarly. A simple approach to extracting the information contained in an image of a face is to somehow capture the variation in a collection of face images, independent of any judgment of features and use this information to encode and compare individual term.

After face is detected, in recognition stage the first step is that face is projected on to eigenspace and new descriptor is calculated. Then it is taken as an input to neural network. In face recognition each face is treated as a separate individual. And for each person separate neural network is created. Face image is treated as a two dimensional array having dimensions 60×60 with entries containing intensity values. Principle Component analysis is applied on to face, because it is an efficient way to represent face. Here new coordinate system is created for faces where coordinates are part of the eigenvectors of a set of face images. Faces are encoded by small set of weights corresponding to their projection onto new coordinate system, and recognize by comparing them with known individual.

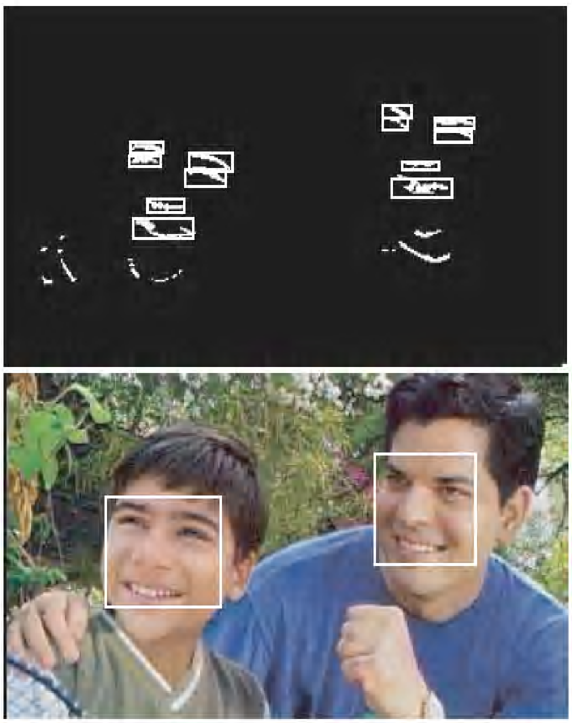

Fig. 3: Results of face detection: Feature extraction, Face detection

4.1 Eigenspace Representation of Face

First an initial set of face images [X1, X2, …, Xn] is prepared. Now the average value of the face is calculated by X=(X1+X2+….+Xn)/n. Then the average value is removed from each face, Y=Xi-X, i=1, 2,…., n. Now, eigenvectors are calculated from each of the new image set [Y1, Y2,….., Yn]. These eigenvectors are orthogonal to each other and they look like sort of face and can be referred as eigenfaces. The dimension of the complete eigenspace is n-1, because the eigenvalue of the remaining eigenface is 0. Fig. 4. shows face window extraction, Fig. 5. shows extracted face, and Fig. 6. shows resize face. Sample face images, the average face of them, eigenfaces of face images are shown in Fig. 7, Fig. 8 and Fig. 9 respectively. With each eigenvector eigenvalue is associated. Eigenvector with higher eigenvalue provide more information on the face image than those with smaller values. This is shown in Fig. 10 and Fig. 11 respectively. Faces can be reconstructed from their eigenvalues. Faces with high eigenvalue can be reconstructed easily than faces with low eigenvalue. Because they provide much more information on face variation. A better approach is to use neural network architecture and do learning in unsupervised manner to identify face.

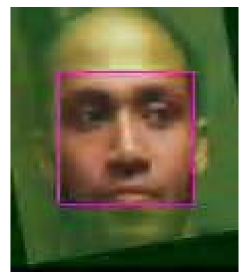

Fig. 4: Example of face window extracted in pink boundary

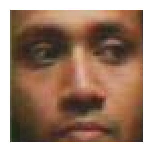

Fig. 5: Extracted face

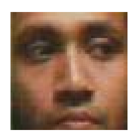

Fig. 6: Resize 60X60

Fig. 7: Sample face

Fig. 8: Mean face

Fig. 9: Eigenfaces of sample face

4.2 Unsupervised Learning

In neural network approach to recognize face neural net is created for each person. Neural network has a backpropogation structure with three layers. First layer has 100 nodes. Middle layer has 10 nodes and output node gives result among 0.1 to 1.0.

As higher the value of output node then input face is likely to be the network's host. Recognition algorithm selects the network with maximum output. And if selected network satisfies predetermine threshold then input face is accepted otherwise it is rejected. Fig. 12 represents neural network architecture .

Fig. 10: Faces with high eigenvalue

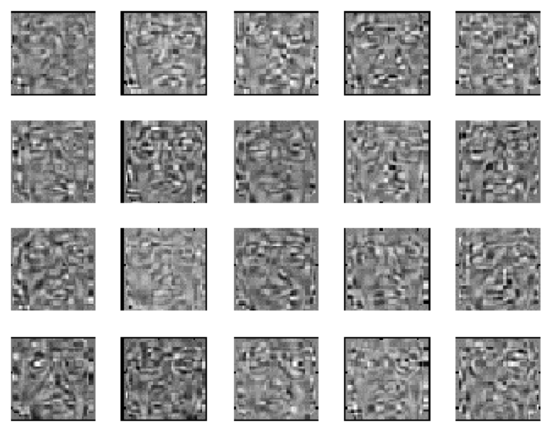

Fig. 11: Faces with low eigenvalue

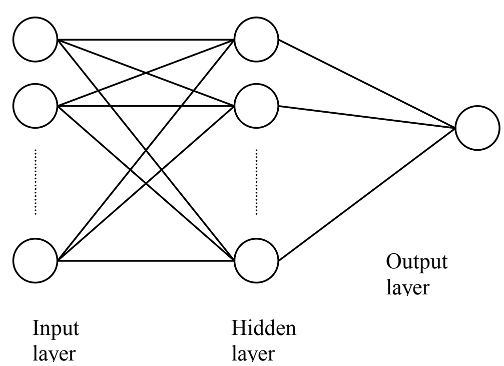

Fig. 12: Neural network architecture

Once neural network structure is created then very next step is to train a network. For that training set should be created. In the beginning of training, a number of face images are selected for each person;

those are well assigned to frontal view. Any of them can represent their host clearly. All the faces are detected or cut by face detection code. These faces can be used as positive examples for their own network and negative examples for other network.

After the basic neural network is created we run them over new faces from individual in our database. If image fails to pass detection code it is ignored. If face detection code detects face in image, it is applied to face recognition code. We check the face recognition result to find more faces for training. Once we get these new faces we add them to our training examples and neural network is retrained. This process will continue until no significant recognition error is found. If we use the original face descriptors from the training examples as neural network input, it is difficult to make coverage. So the average of training set is set to zero and unify its standard derivation.

5. SUMMARY OF EIGENFACE RECOGNITION PROCEDURE

The eigenfaces approach in face recognition is in the following steps:

1. Form a face library that consists of face images of Known individuals.
2. Choose a training set that include a number of images (N) for each person with some variation expression and lightening.
3. Calculate matrix N×N, find its eigenvectors and eigenvalue, and choose the eigenvector with highest associated eigenvalue.
4. Combine normalized training set of images to produce eigenfaces. Store these eigenfaces for later use.
5. For each member in training library, compute and store a feature vector.
6. Create neural network for each person in database.
7. Train these network as the faces are used as a positive examples of their own network and negative examples of all other networks.
8. For each new face image to be identified, calculate its feature vector.
9. Use these feature vectors as network input and simulate all networks with these inputs.
10. Select the network with maximum output. If output selected network passes a predefined threshold, it is reported as the host of input face. Otherwise it is reported as unknown and it is added this member to the face library with its feature vector and network.

6. EXPERIMENTAL RESULTS

The proposed face detection approach is implemented using MATLAB software with a compound dataset of 600 images. Images are taken from CIT face database, FERET face database and internet. The test images are expected to contain single as well as multiple frontal human faces of different sizes with complex background under normal lighting conditions. The faces having different poses and expressions are also detected successfully as they are invariant with respect to skin color. In experimental results, success ratio is more than 93%.

With the choice of YCbCr color model and sigma control limits for variations in skin color components the skin region segmentation in a given image is achieved more accurately and quickly. Further, the LS face model is applied to each of the segmented skin region, after preprocessing, in order to select possible faces in these regions. The LS face model has better success rate of 93.5% and is also faster than the geometrical face model. Hence, the proposed face detection algorithm has the merit of both the efficient skin region segmentation and the robust LS face model. It is invariant to face posses and expressions. It is able to detect multiple faces

in an image with complex background and is race invariant with the exception of block complexion. However, it fails to detect side view faces and occulted faces. This is due to the fact that the LS face model is constrained to detect only the frontal view face, the skin color segmentation yields single region for occulted faces and the face area is constrained to be greater than 500 pixels.

A system which does face detection from the images taken by camera is built and then the face is

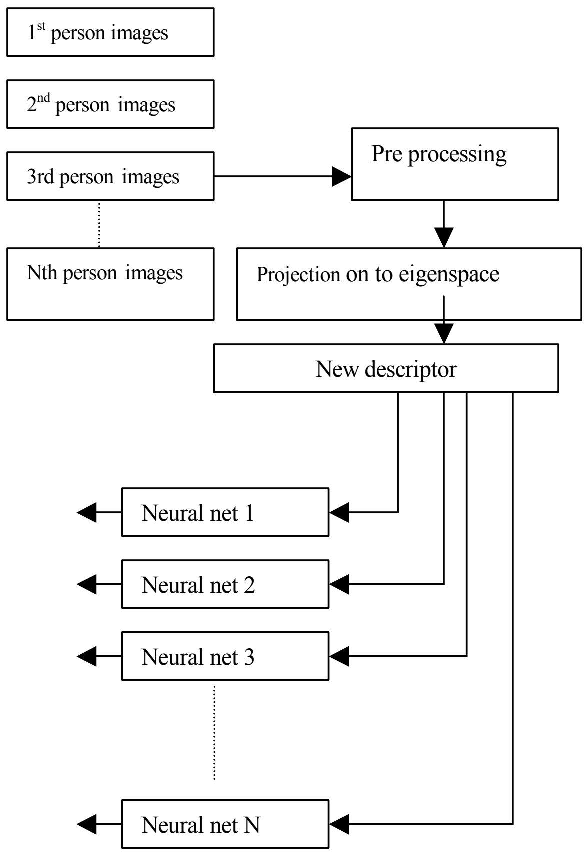

Fig. 13: Training of neural network

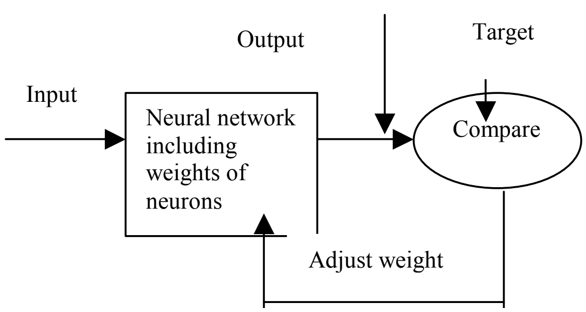

Fig. 14: Neural network operation

identified using face recognition code. Experiments are conducted under roughly the same illumination condition. Wide expression variations are incorporated. The faces are frontal view without significant orientation change. A recognition rate of more than 94% is achieved. Recognition rate is improved by recognizing same person multiple times.

Face recognition has become an important issue in many applications such as security system, credit card verification and criminal identification. For example, the ability to model a particular face and distinguish it from a large of stored face model would make it possible to vastly improve the criminal identification. Even the ability to merely detect faces, as apposed to recognizing them, can be important. Detecting faces in photographs for automating color file development can be very useful, since the effect of many enhancements and noise reduction techniques depends on the image content.

The eigenface approach used in this work has advantage in its speed, simplicity, learning capability and robustness to small changes in face images.

7. COMPARISON

Table 1: Statistical values for the Proposed Skin Color Space

Color Component	Mean($\overline{\mu}$)	Std. Deviation(σ)	lcl	ucl
Cb (Blue)	120	15	97.5	142.5
Cr(Red)	155	14	134	176

Table 2: Comparison of Time, Segment Skin area and Number of Candidate Facial Feature Bloks for the various Skin Color Segmentation Methods

Skin Color Spaces Based on	Avg. time(in secs)	Std. Dev.	% Average Segmented Skin Area	Average No. of Candidate Facial Feature Blocks in the Segmented Skin Area
RGB Model	1.04	0.0331	29.00	67
HSV Model	0.59	0.0395	32.83	84
YCbCr Model	2.12	0.0144	26.31	26
YUV Model	1.01	0.0136	52.85	99
YIQ Model	1.05	0.0143	66.07	105
Proposed Model	0.82	0.0137	25.28	21

8. ACKNOWLEDGEMENT

I would like to thank Prof. S.H. Patil for his guidance on my biometric work and solving my doubts at several stages of my work.

REFERENCES

1. Mathew, J.H.2003. Eigenvalues and Eigenvectors, PP: 612-621.
2. Nazish. et. al, , 2001. Face recognition using neural network, Proc. IEEE INMIC 2001, PP- 277-281.
3. A. pentland, B. Moghaddam and T. starner, “View based and modular eigenfaces for face recognition”, Proceeding IEEE Conf. on Computer vision and pattern recognition, P.84-91, June 1994.
4. Roberto Brunelli and Tomaso Poggio, “Face recognition: Feature versus templates”, IEEE Transaction on pattern Analysis and Machine Intelligence, 15(10):1042-1052, 1993.
5. M. Turk and A. Pentland, “Eigenfaces for recognition “, Journal of Cognitive Neuroscience, vol.3, no.1, PP.71-86, 1991.
6. Manjunath, B.S.Chellappa, R., and Malsburg C., “A feature based approach to face recognition”, Trans. of IEEE, PP. 373-378 (1992) Carnegie Mellon University, November 1995.
7. Ching, C. W., and Huang, L. C., “Human face recognition from single front view”, Int. J. of Pattern Recognition and Artificial Intelligence, Vol. 6(4), pp. 570-593, (1992).
8. C. Garcia and G. Tziritas, Face detection using quantized skin color region merging and wavelet packet analysis, IEEE Trans. Multimedia. 1(3) (1999) 264-277.
9. J. Park, J. Seo, A. Dongun and S. Chung, Detection of human faces using skin color and eyes, Proc. IEEE (2000).
10. H. Wang and S.F.Chang, A highly efficient system for automatic face detection in MPEG video, IEE Trans, CSVT 7(4) (1997) 615-628.
11. G.Yang and T.S.Huang, Human face detection in a complex background, Patt. Recogn.27 (1) (1994)53-63.
12. D.Chai and K.N.Ngan, Face segmentation using skin color map in videophone applications, IEEE Trans. CSVT 9(4) (1999) 551-564.
13. B.S.Manjunath and W.Y.Ma. Texture features for browsing and retrieval of image data. IEEE Trans. Pattern Analysis and Machine Intelligence, 18(8):837-842, 1996.
14. A.Rahardja, A.Sowmya, and W.Wilson. A neural network approach to component versus holistic recognition of facial expression in images. In Proc. SPIE Intelligent Robots and Computer Vision X: Algorithms and Techniques, volume 1607, pages 62-70, 1991.