

Modelling Uncertainty in Representation of Facial Features for Face Recognition

Hiremath P.S., Ajit Danti and Prabhakar C.J.
*Gulbarga University, Gulbarga; JNN College of Engineering, Shimoga & Kuvempu University, Shimoga
 India*

1. Introduction

Face is one of the important biometric identifier used for human recognition. The face recognition involves the computation of similarity between face images belonging to the determination of the identity of the face. The accurate recognition of face images is essential for the applications including credit card authentication, passport identification, internet security, criminal databases, biometric cryptosystems etc. Due to the increasing need for the surveillance and security related applications in access control, law enforcement, and information safety due to criminal activities, the research interest in the face recognition has grown considerably in the domain of the pattern recognition and image analysis. A number of approaches for face recognition have been proposed in the literature (Zhao et al. 2000), (Chellappa et al. 1995). Many researchers have addressed face recognition based on geometrical features and template matching (Brunelli and Poggio, 1993). There are several well known face recognition methods such as Eigenfaces (Turk and Pentland 1991), Fisherfaces (Belhumeur et al. 1997), (Kim and Kitter 2005), Laplacianfaces (He et al. 2005). The wavelet based Gabor function provide a favorable trade off between spatial resolution and frequency resolution (Gabor 1946). Gabor wavelets render superior representation for face recognition (Zhang, et al. 2005), (Shan, et al. 2004), (Olugbenga and Yang 2002).

In recent survey, various potential problems and challenges in the face detection are explored (Yang, M.H., et al., 2002). Recent face detection methods based on data-driven learning techniques, such as the statistical modeling methods (Moghaddam and Pentland 1997), (Schneiderman, and Kanade, 2000), (Shih and Liu 2004), the statistical learning theory and SVM based methods (Mohan et al., 2001). Schneiderman and Kanade have developed the first algorithm that can reliably detect human faces with out-of-plane rotation and the first algorithm that can reliably detect passenger cars over a wide range of viewpoints (Schneiderman and Kanade 2000). The segmentation of potential face region in a digital image is a prelude to the face detection, since the search for the facial features is confined to the segmented face region. Several approaches have been used so far for the detection of face regions using skin color information. In (Wu, H.Q., et al., 1999), a face is detected using a fuzzy pattern matching method based on skin and hair color. This method has high detection rate, but it fails if the hair is not black and the face region is not elliptic. A face detection algorithm for color images using a skin-tone color model and facial features is

presented in (Hsu et al. 2002). Face recognition can be defined as the identification of individuals from images of their faces by using a stored database of faces labeled with people's identities. This task is complex and can be decomposed into the smaller steps of detection of faces in a cluttered background, localization of these faces followed by extraction of features from the face regions, and finally, recognition and verification. It is a difficult problem as there are numerous factors such as 3D pose, facial expression, hair style, make up etc., which affect the appearance of an individual's facial features. In addition to these facial variations, the lighting, background, and scale changes also make this task even more challenging. Additional problematic conditions include noise, occlusion, and many other possible factors.

Many methods have been proposed for face recognition within the last two decades. Among all the techniques, the appearance-based methods are very popular because of their efficiency in handling these problems (Chellappa et. al. 1995). In particular, the linear appearance based face recognition method known as eigenfaces (Turk & Pentland 1991) is based on the principal component analysis of facial image ensembles (Kirbi & Sirovich 1990). The defining characteristic of appearance-based algorithms is that they directly use the pixel intensity values in a face image as the features on which to base the recognition decision. The pixel intensities that are used as features are represented using single valued variables. However, in many situations same face is captured in different orientation, lighting, expression and background, which lead to image variations. The pixel intensities do change because of image variations. The use of single valued variables may not be able to capture the variation of feature values of the images of the same subject. In such a case, we need to consider the symbolic data analysis (SDA) (Bock & Diday 2000; Diday 1993), in which the interval-valued data are analyzed. Therefore, there is a need to focus the research efforts towards extracting features, which are robust to variations due to illumination, orientation and facial expression changes by representing the face images as symbolic objects of interval type variables (Hiremath & Prabhakar 2005). The representation of face images as symbolic objects (symbolic faces) accounts for image variations of human faces under different lighting conditions, orientation and facial expression. It also drastically reduces the dimension of the image space. In (Hiremath & Prabhakar 2005), a symbolic PCA approach for face recognition is presented, in which symbolic PCA is employed to compute a set of subspace basis vectors for symbolic faces and then project the symbolic faces into the compressed subspace. This method requires less number of features to achieve the same recognition rate as compared to eigenface method. The symbolic PCA technique, however, encodes only for second order statistics, i.e., pixel wise covariance among the pixels, and is insensitive to the dependencies of multiple (more than two) pixels in the patterns. As these second order statistics provide only partial information on the statistics of both natural images and human faces, it might become necessary to incorporate higher order statistics as well. The kernel PCA (Scholkopf et. al. 1998) is capable of deriving low dimensional features that incorporate higher order statistics. Higher order dependencies in an image include nonlinear relations among the pixel intensity values, such as the relationships among three or more pixels in an edge or a curve, which can capture important information for recognition. The kernel PCA is extended to symbolic data analysis as symbolic kernel PCA (Hiremath & Prabhakar 2006) for face recognition and the experimental results show improved recognition rate as compared to the symbolic PCA method. The extension of symbolic analysis to face recognition techniques using methods based on linear discriminant

analysis, two-dimensional discriminant analysis, Independent component analysis, factorial discriminant analysis and kernel discriminant analysis has been attempted in (Hiremath and Prabhakar Dec 2006, Jan 2006, Aug 2006, Sept 2006, 2007).

It is quite obvious that the literature on face recognition is replete with a wide spectrum of methods addressing a broad range of issues of face detection and recognition. However, the objective of the study in the present chapter is the modeling of uncertainty in the representation of facial features, typically arising due to the variations in the conditions under which face images of a person are captured as well as the variations in the personal information such as age, race, sex, expression or mood of the person at the time of capturing the face image. Two approaches, namely, fuzzy-geometric approach and symbolic data analysis, for face recognition are considered for the modeling of uncertainty of information about facial features.

2. Fuzzy face Mode for Face Detection

In (Hiremath and Danti, Dec 2005), the detection of the multiple frontal human faces based on the facial feature extraction, using the fuzzy face model and the fuzzy rules, is proposed and it is described in this section. The input color image is searched for the possible skin regions using the skin color segmentation method. In which, 2D chromatic space CbCr using the sigma control limits on the chromatic components Cb and Cr, derived by applying the statistical sampling technique. Each potential face region is then verified for a face in which, initially, the eyes are searched and then the fuzzy face model is constructed by dividing the human facial area into quadrants by two reference lines drawn with respect to the eyes. Further, other facial features such as mouth, nose and eyebrows are searched in the fuzzy face model using the fuzzy rules and then face is detected by the process of defuzzification. Overview of this fuzzy-geometric approach is shown in the Figure 3.

2.1 Skin Color Segmentation

Face detection based on skin color is invariant of facial expressions, rotations, scaling and translation (Hsu et al. 2002). Human skin color, with the exception of very black complexion, is found in a relatively narrow color space. Taking advantage of this knowledge, skin regions are segmented using the skin color space as follows.

Skin Color Space

The YCbCr color model is used to build the skin color space. It includes all possible skin colors. We are able to extract more facial skin color regions excluding the non-skin regions. The skin color space uses only the chromatic color components Cb and Cr for skin color segmentation using the sigma control limits (Hiremath and Danti, Feb 2006). The procedure to build skin color space is described as following.

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 in which luminance is partially separated (Jain A.K. 2001). Skin color space is developed by considering the large sample of facial skins cropped manually from the color face images of the multi racial people. Skin samples are then filtered using low pass filter (Jain 2001) to remove noises. The lower and

upper control limits of the pixel values for the chromatic red and blue color components are determined based on one-and-half sigma limits using the equation (1).

$$\mu_i = \frac{1}{(m \times n)} \sum_{x=1}^m \sum_{y=1}^n c(x,y), \quad \bar{\mu} = \frac{1}{k} \sum_{i=1}^k \mu_i, \quad \sigma = \sqrt{\frac{\sum_{i=1}^k (\mu_i - \bar{\mu})^2}{k}} \quad (1)$$

$$lcl = \bar{\mu} - 1.5\sigma, \quad ucl = \bar{\mu} + 1.5\sigma$$

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

$$P(x,y) = \begin{cases} 1, & \text{if } (lcl_r \leq Cr(x,y) \leq ucl_r) \ \& \ (lcl_b \leq Cb(x,y) \leq ucl_b), \\ 0, & \text{Otherwise,} \end{cases} \quad (2)$$

where $Cr(x,y)$ and $Cb(x,y)$ are the chromatic red and blue component values of the pixel at (x,y) in the red and blue planes of the test image respectively. Hence, the lower and upper sigma control limits lcl_r and ucl_r for red and lcl_b and ucl_b for blue colors, can transform a color image into a binary skin image P , such that the white pixels belong to the skin regions and the black pixels belong to the non skin region as shown in the Figure 1(b). In the computation of the lower and upper control limits, experimental results show that, in the 3σ limits, the probability of inclusion of non-skin pixels in the face area is high. On the contrary, in the σ limits, the probability of omission of facial skin pixels in the face area is high. It is found that 1.5σ limits are the optimal limits, which yield a suitable trade off between the inclusion of facial skin pixels and the omission of non-skin pixels in the face area. In the experiments, the values of the mean $\bar{\mu}$ and the standard deviation σ , and lower and upper control limits of the chromatic color components are quantified based on the several sample skin images of the multiracial people and are mentioned in the Table 1. The sigma control limits are flexible enough to absorb the moderate variations of lighting conditions in the image to some extent. The results of the skin color segmentation are shown in the Figure 1(b). The skin color segmentation leads to a faster face detection process as the search area for the facial features is comparatively less. The comparative analysis of the different skin color segmentation methods is shown in the Table 2.

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

Table 1. Statistical values for the skin color space

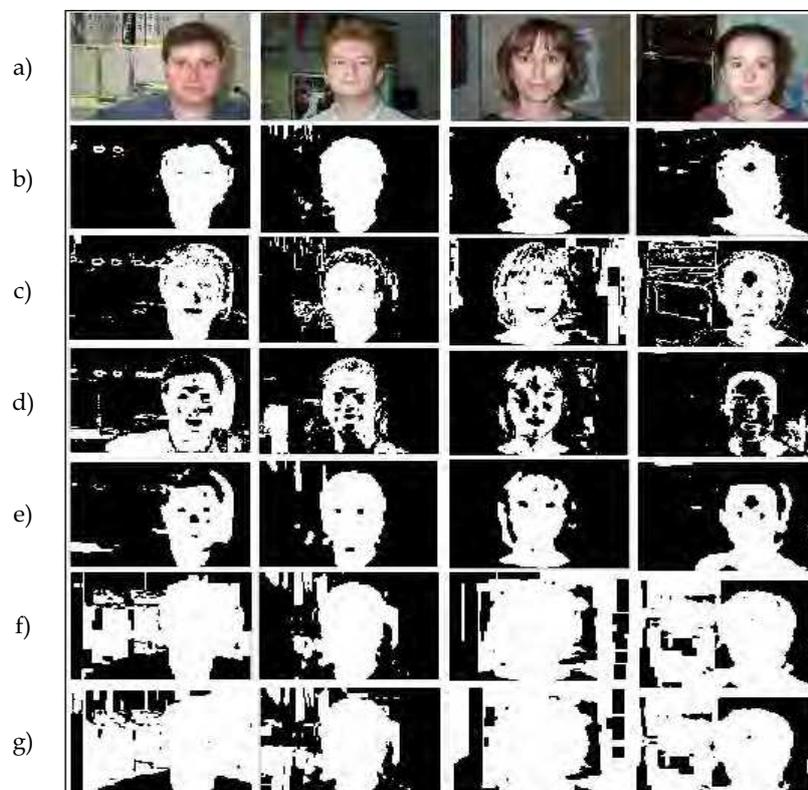


Figure 1. Comparison of skin segmentation results. a) Original Image, b) YCbCr (Hiremath-Danti, Feb 2006), c) RGB (Wang-Yuan method), d) HSV (Bojic method), e) YCbCr (Chai method), f) YUV (Yao method), g) YIQ (Yao method)

Skin Color spaces based on	Avg. time (In secs)	Std. Dev.	% Avg segmented skin area	Avg No. of facial feature blocks
RGB Model (Wang & Yuan 2001)	1.04	0.0332	29.00	67
HSV Model (Bojic & Pang 2000)	0.59	0.0395	32.83	84
YCbCr Model (Chai & Ngan 1999)	2.12	0.0145	26.31	26
YUV Model (Yao and Gao 2001)	1.01	0.0136	52.85	99
YIQ Model (Yao and Gao 2001)	1.05	0.0143	66.07	105
YCbCr (Hiremath & Danti, Feb 2006)	0.82	0.0137	25.28	21

Table 2. Comparison of time, segmented skin area, and number of candidate facial feature blocks for the various skin color segmentation methods

Pre processing of Skin Segmented Image

The binary skin segmented image obtained above is preprocessed by performing binary morphological opening operation to remove isolated noisy pixels. Further, white regions

may contain black holes these black holes may be of any size and are filled completely. The binary skin image is labeled using the region labeling algorithm and their feature moments, such as center of mass (\bar{x}, \bar{y}) , orientation θ , major axis length, minor axis length and area, are computed (Jain, A.K., 2001; Gonzalez, R.C., et al., 2002). By the observation of several face regions under analysis, it is found that the face regions are oriented in the range of $\pm 45^\circ$ degrees in the case of frontal view of the face images. Only such regions are retained in the binary skin image for further consideration. The remaining regions are considered as non face regions and are removed from the binary skin image. The observation of several real faces also revealed that the ratio of height to width of each face region is approximately 2, only such regions are retained. Further, though the skin regions of different sizes are successfully segmented, it is found that the potential facial features are miss-detected whenever the face area is less than 500 pixels. Hence, the regions, whose area is more than the 500 pixels are considered for the face detection process. The resulting binary skin image after the preprocessing and applying the above constraints is expected to contain potential face regions (Fig 2(a), (b)).

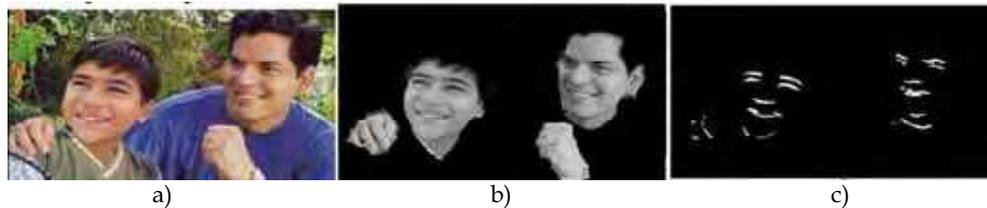


Figure 2. Results of Skin color segmentation a) Original Image b) Potential face regions in gray scale image c) Sobel Filtered Binary image

2.2 Face Detection

Each potential face region in the binary image is converted into gray scale image as shown in Figure 2.(b) and then each face region is passed on to our fuzzy face model to decide whether the face is present in that region or not, by the process of facial feature extraction using the fuzzy rules (Hiremath & Danti Dec. 2005). The detailed face detection process, which detects multiple faces in an input image, is described in Figure 3.

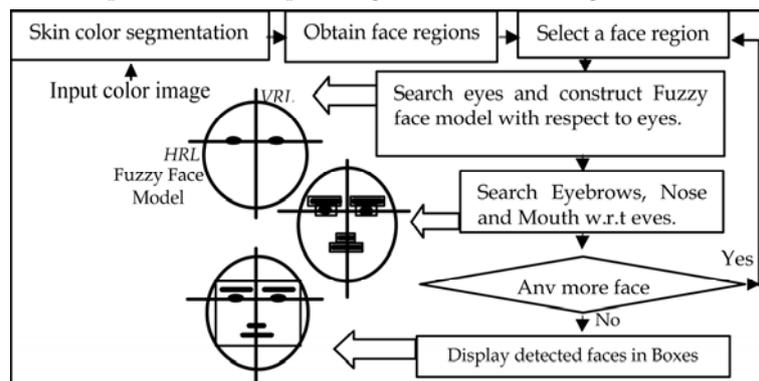


Figure 3. Overview of the multiple face detection process

Preprocessing of Face Regions

The each gray scale version of the potential face region is filtered using the Sobel edge filter and binarized using a simple global thresholding and then labeled. In the labeled image, the essential facial feature blocks are clearly visible in the potential face region under consideration Figure 2(c). Further, for each facial feature block, its center of mass (\bar{x}, \bar{y}) , orientation θ , bounding rectangle and the length of semi major axis are computed (Jain, A.K., 2001).

Feature Extraction

The feature blocks of the potential face region in the labeled image are evaluated in order to determine which combination of feature blocks is a potential face and the procedure is explained as follows:

Searching Eyes

The eyes are detected by exploiting the geometrical configuration of the human face. All the feature blocks are evaluated for eyes. Initially, any two feature blocks are selected arbitrarily and assume them as probable eye candidates. Let (x_1, y_1) and (x_2, y_2) be respectively, the centers of right feature block and left feature block. The line passing through the center of both the feature blocks is called as the *horizontal-reference-line (HRL)* as shown in Figure 4 and is given by the equation (3) and the slope angle θ_{HRL} between the HRL and x-axis is given by equation (4).

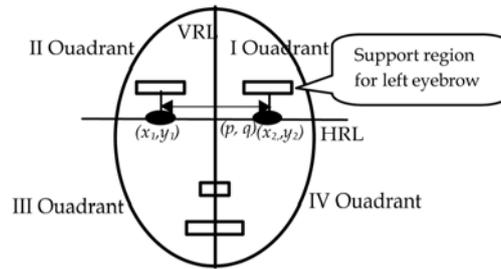


Figure 4. Fuzzy face model with support regions for eyebrows, nose and mouth shown in rectangles

$$ax + by + c_{HRL} = 0$$

$$\text{where, } a = y_2 - y_1, \quad b = x_1 - x_2, \quad c_{HRL} = x_2 y_1 - x_1 y_2 \quad (3)$$

The slope angle θ_{HRL} between the HRL and x-axis is given by:

$$\theta_{HRL} = \tan^{-1}(-a/b), \quad -\pi/2 \leq \theta_{HRL} \leq \pi/2 \quad (4)$$

Since the fuzzy face model is a frontal view model, a face in a too skewed orientation is not considered in this model. Hence, the slope angle θ_{HRL} is constrained within the range of $\pm 45^\circ$. If the current pair of feature blocks does not satisfy this orientation constraint, then they are rejected and another pair of feature blocks from the remaining feature blocks is taken for matching. Only for the accepted pairs of features, the normalized lengths of the semi major axis l_1 and l_2 are computed by dividing the length of the semi major axis by the distance D between these two features. The distance D is given by the equation (5).

$$D = \left[(x_1 - x_2)^2 + (y_1 - y_2)^2 \right]^{1/2} \quad (5)$$

Let θ_1 and θ_2 are the orientations of the above accepted feature blocks. The evaluation function E_{Eye} is computed using the equation (6) to check whether the current pair of features is a potential eye pair or not.

$$E_{Eye} = \exp \left[-1.2 \left((l_1 - l_2)^2 + (l_1 + l_2 - 1)^2 + (\theta_1 - \theta_{HRL})^2 + (\theta_2 - \theta_{HRL})^2 \right) \right] \quad (6)$$

This evaluation function value ranges from 0 to 1 and it can be given the interpretation of a probability value. The constant 1.2 is the mean of the negative exponential distribution, which is determined empirically with respect to the sample images considered for experimentation to optimize higher detection rate with lower false detections. Hence, higher the evaluation value E_{Eye} higher is the probability of the two selected feature blocks to be eyes. If this evaluation value is greater than an empirical threshold value 0.7, then these two feature blocks are accepted as the *potential eye pair candidate*. Otherwise, this pair of blocks is rejected and another pair of feature blocks is selected. For potential eye pair candidate, the fuzzy face model is constructed and the other facial features are searched as follows.

Construction of Fuzzy Face Model

It is assumed that every human face is having the same geometrical configuration and the relative distances between the facial features are less sensitive to poses and expressions (Yang et al. 2002). The fuzzy face model is constructed with respect to the above potential eye candidates. A line perpendicular to the *HRL* at the mid point of the two eyes is called as vertical reference line (*VRL*). Let (p, q) be the mid point of the line segment joining the centers of the two eye candidates. Then the equation of the *VRL* is given by equation (7).

$$bx - ay + c_{VRL} = 0 \quad (7)$$

These two reference lines (*HRL* and *VRL*) are used to partition the facial area into quadrants as shown in Figure 4. The vertical and horizontal distances of the facial features namely, eyebrows, nose and mouth are empirically estimated in terms of the distance D between the centers of the two eyes on the basis of the observations from several face images. The notations $V_{Eyebrows}$, V_{Nose} and V_{Mouth} denote the vertical distances of the centers of eyebrows, nose and mouth from the *HRL* which are estimated as $0.3D$, $0.6D$ and $1.0D$ respectively. The notations $H_{Eyebrows}$, H_{Nose} and H_{Mouth} denote the horizontal distances of the centers of eyebrows, nose and mouth from the *VRL* which are estimated as $0.5D$, $0.05D$ and $0.1D$ respectively. The facial features are enclosed by the rectangles to represent the support regions, which confine the search area for facial features. This completes the construction of the fuzzy face model with respect to the selected potential eye pair candidate in the given face region as shown in Figure 4. Further, the fuzzy face model is used to determine which combination of the feature blocks is a face.

Searching Eyebrows, Nose and Mouth

The searching process proceeds to locate the other potential facial features, namely eyebrows, nose and mouth with respect to the above potential eye pair candidate. The support regions for eyebrows, nose and mouth are empirically determined using fuzzy rules as given in Table 3. Then these support regions are searched for facial features. For illustration, we take the left eyebrow feature as an example to search. Let a feature block K

be a potential left eyebrow feature. The horizontal distance h_{Leb} and the vertical distance v_{Leb} of the centroid of the K^{th} feature from the VRL and HRL, respectively, are computed using the equation (8).

Feature(j)	Vertical distances				Horizontal distances			
	\min_{v_j}	\max_{v_j}	\bar{v}_j	σ_{v_j}	\min_{h_j}	\max_{h_j}	\bar{h}_j	σ_{h_j}
Eyebrows	0.02	0.38	0.2	0.06	0.24	0.65	0.45	0.07
Nose	0.30	0.90	0.6	0.10	-0.2	0.2	0.0	0.07
Mouth	0.45	1.35	0.9	0.15	-0.3	0.3	0.0	0.10

Table 3. Empirically determined distances of the facial features (normalized by D)

$$h_{Leb} = \frac{|b\bar{x}_K - a\bar{y}_K + c_{VRL}|}{(a^2 + b^2)^{1/2}} \quad \text{and} \quad v_{Leb} = \frac{|a\bar{x}_K + b\bar{y}_K + c_{HRL}|}{(a^2 + b^2)^{1/2}}, \quad (8)$$

Treating h_{Leb} and v_{Leb} as the fuzzy quantities to represent the possible location of the potential left eyebrow feature, the fuzzy membership values $\mu_{h_{Leb}}$ and $\mu_{v_{Leb}}$, respectively, are defined using the trapezoidal fuzzy membership function (Hines & Douglas 1990). In particular, the membership function $\mu_{v_{Leb}}$ is defined using the equation (9) and Table 3.

$$\mu_{v_{Leb}}(v_{Leb}) = \begin{cases} 0, & \text{if } v_{Leb} \leq \min v_{Leb} \\ \frac{(v_{Leb} - \min v_{Leb})}{(\alpha - \min v_{Leb})}, & \text{if } (\min v_{Leb} \leq v_{Leb} \leq \alpha) \\ 1, & \text{if } (\alpha \leq v_{Leb} \leq \beta) \\ \frac{(\max v_{Leb} - v_{Leb})}{(\max v_{Leb} - \beta)}, & \text{if } (\beta \leq v_{Leb} \leq \max v_{Leb}) \\ 0, & \text{if } (v_{Leb} \geq \max v_{Leb}) \end{cases} \quad (9)$$

Similarly, the membership function $\mu_{h_{Leb}}$ is defined. The support region for the potential left eyebrow feature is the set of values h_{Leb} and v_{Leb} whose fuzzy membership values are non-zero. The Figure 5(a) shows the graph of the trapezoidal fuzzy membership function μ_{v_j} for the vertical distance of the j^{th} feature and the support region for the left eyebrow is shown in Figure 5(b). To evaluate K^{th} feature block in the support region for left eyebrow, the value of the evaluation function E_K is given by the equation (10). The E_K value ranges from 0 to 1 and represents the probability that the feature block K is a left eyebrow.

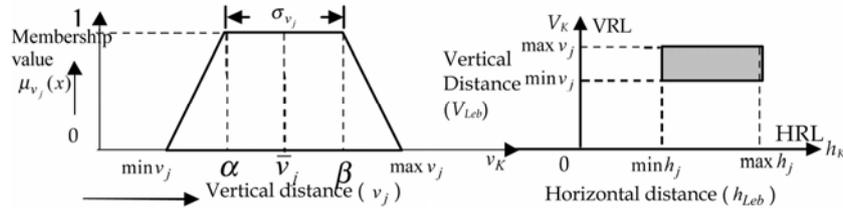


Figure 5. Trapezoidal fuzzy membership function μ_{v_j} for the vertical distance of the j^{th} facial feature b) Support region for left eyebrow in the I quadrant of face model

$$E_K = \frac{1}{2} \left(\exp \left[-1.2 \left(\frac{v_{Leb} - V_{Eyebrows}}{D/2} \right)^2 \right] + \exp \left[-1.2 \left(\frac{h_{Leb} - H_{Eyebrows}}{D/2} \right)^2 \right] \right) \quad (10)$$

Similarly, evaluation value is computed for all the feature blocks present in that support region of the left eyebrow. The evaluation value E_{Leb} is a fuzzy quantity represented by the set of E_K values with their corresponding fuzzy membership values μ_K . The membership value μ_{Leb} corresponding to E_{Leb} is obtained by the *min-max* fuzzy composition rule (Klir & Yuan 2000) given by the equations (11) and (12). The feature block having the evaluation value E_{Leb} with the corresponding μ_{Leb} found in the support region of the left eyebrow is the *potential left eyebrow* feature with respect to the current pair of potential eye candidates.

$$\mu_K = \min(\mu_{h_K}, \mu_{v_K}), \text{ for each } K \quad (11)$$

$$\mu_{Leb} = \max_K \{\mu_K\} \quad (12)$$

Similarly, the right eyebrow, nose and mouth are searched in their respective support regions determined by appropriately defining the membership functions for the fuzzy distances (horizontal and vertical) from the centroid of these facial features, and their fuzzy evaluation values are computed by applying the *min-max* fuzzy composition rule. The overall fuzzy evaluation E for the fuzzy face model is defined as the weighted sum of the fuzzy evaluation values of the potential facial features namely, for the eye, left eyebrow, right eyebrow, nose and mouth, respectively. The weights are adjusted to sum to unity as given in the equation (13). The membership value μ_E corresponding to E is obtained by the fuzzy composition rule as given by the equation (14).

$$E = 0.4E_{Eye} + 0.3E_{Mouth} + 0.2E_{Nose} + 0.05E_{Leb} + 0.05E_{Reb} \quad (13)$$

$$\mu_E = \min\{\mu_{Mouth}, \mu_{Nose}, \mu_{Leb}, \mu_{Reb}\} \quad (14)$$

Above procedure is repeated for every potential eye pair candidate and get the set of fuzzy faces. These fuzzy faces are represented by the set of E values with their corresponding membership values μ_E . Finally, the most probable face is obtained by the defuzzification process as given by the equation (15).

$$\mu_{E_{max}} = \max_{E \in \Omega} \{\mu_E\} \quad (15)$$

Then the E value corresponding to $\mu_{E_{max}}$ is the defuzzified evaluation value E_D of the face. If there are more than one E value corresponding to $\mu_{E_{max}}$, the maximum among those values is the defuzzified evaluation value E_D of the face. Finally, the potential eyes, eyebrows, nose and mouth features corresponding to the overall evaluation value E_D constitute the most probable face in the given face region, provided E_D is greater than the empirical threshold value 0.7. Otherwise this face region is rejected. The face detection results are shown in Figure 6, where (a) display the feature extraction in which facial features are shown in bounding boxes (Jain 2001) and (b) shows detected face in rectangular box. (Hiremath P.S. & Danti A. Feb 2006). The above procedure is repeated for every potential face region to detect possible faces in the input image.



Figure 6. Results of Face Detection a) Facial Feature extraction b) Detected face in box

2.3 Experimental Results

The MATLAB 6.0 implementation of the above described procedure on Pentium IV @ 2.6 GHz yields the success rate of 96.16%. The average time taken to detect one face is about 0.78 sec, which depends on the size of the potential face region. The search area for the facial feature extraction is confined to only the total area covered by the support regions i.e. $0.67D^2$, (D is distance between eyes) which is considerably very small compared to that of the image size. This reduced search area leads to the reduction in the detection time to a great extent. Sample detection results are shown in Figure 7 and Figure 8 with detected faces enclosed in rectangular boxes. Due to the constraints of the face model, miss-detection occurs due to several reasons i.e. profile (side) view faces, abnormal lighting conditions, face occluded by hair, very small face sizes, face occluded by hand and too dark shadow on faces as shown in Figure 9.

The comparison of different state of the art detectors proposed by (Shih and Liu 2004, we refer as S-L method) and (Schneiderman and Kanade 2000, we refer as S-K method) and (Hiremath and Danti, Dec. 2005, we refer as H-D method) is given in Table 4. It is observed that, fuzzy face model approach based on skin color segmentation (H-D method) is comparable to others in terms of detection rate and very low in both detection time and false detections.

Method	Det. Rate (%)	False detection	Det. Time (secs)	Dataset	No. of images	No. of faces
S-L method	98.2	2	not reported	MIT-CMU	92	282
S-K method	94.4	65	5	MIT-CMU	125	483
H-D method	96.1	02	0.78	CIT, FERET, Internet	650	725

Table 4. Comparison of performance



Figure 7. Sample detection results for single as well as multiple human faces with sizes, poses, expressions and complex backgrounds



Figure 8. Sample images with expressions, lighting conditions, complex background & beards



Figure 9. Sample images with miss-detections

3. Optimization of feature sets

A small set of geometrical features is sufficient for the face recognition task, which requires less computational time and less memory due to their low dimension. In this approach, facial features detected based on the Fuzzy face model are considered. The normalized geometrical feature vector is constructed with the distances, areas, evaluation values and fuzzy membership values. Normalization is done with respect to the distance between eyes. Further, the feature vector is optimized and demonstrated that the resultant vector is invariant of scale, rotation, and facial expressions. This vector uniquely characterizes each human face despite changes in rotation, scale and facial expressions. Hence, it can be effectively used for the face recognition system. Further, it is a 1-dimensional feature vector space which has reduced dimensionality to a greater extent as compared to the other methods (Turk & Pentland, 1991; Belhumeur et al., 1997) based on the 2-dimensional image intensity space. In (Hiremath and Danti, Dec. 2004), the method of optimization of feature sets for face recognition is presented and it is described as below.

3.1 Geometrical Facial Feature Set

The geometrical facial feature set contains total of about 26 features, in which 12 facial features are obtained from face detector and remaining 14 projected features are determined by the projection of facial features such as eyes, eyebrows, nose, mouth and ears.

Facial Features

Using the face detector based on Lines-of-Separability face model (Hiremath P.S. & Danti A., Feb. 2006) and fuzzy face model (Hiremath P.S. & Danti A., Dec. 2005) respectively, the list of geometrical facial features extracted are given in the Table 5.

Projected Features

The centroid of the facial features obtained by our face detectors are projected perpendicularly to the *Diagonal Reference Line (DRL)* as shown in the Figure 10. The *DRL* is

the line bisecting the first quadrant in the HRL - VRL plane and is a locus of point (x,y) equidistant from HRL and VRL . The equation of the DRL is given by:

$$Ax + By + C = 0, \text{ where the coefficients } A, B, \text{ and } C \text{ are given by:} \quad (16)$$

$$A = (a - b), \quad B = (a + b), \quad C = (c_{HRL} - c_{VRL}) \quad (17)$$

Feature	Description	Feature	Description
E_{Eyes}	Evaluation value of eyes	E_{Rear}	Evaluation value of right ear
E_{Leb}	Evaluation value of left eyebrow	E	Overall evaluation value of the face
E_{Reb}	Evaluation value of right eyebrow	μ_{Leb}	Membership value of left eyebrow
E_{Nose}	Evaluation value of nose	μ_{Reb}	Membership value of right eyebrow
E_{Mouth}	Evaluation value of mouth	μ_{Nose}	Membership value of nose
E_{Lear}	Evaluation value of left ear	μ_{Mouth}	Membership value of mouth

Table 5. List of geometrical features extracted from face detectors

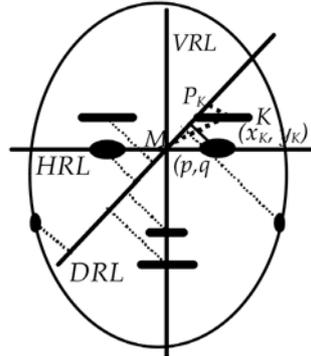


Figure 10. Projection of features onto DRL

Distance Ratio Features

The distance ratios are computed as described in the following. Let (x_K, y_K) be the centroid K of the k^{th} feature (e.g. left eyebrow in the Figure 10). Let P_K be the projection of point K on the DRL . Then, the following distances are computed:

$$KP_K = \left| \frac{Ax_K + By_K + C}{\sqrt{A^2 + B^2}} \right| \quad (\text{Perpendicular distance}) \quad (18)$$

$$MK = \sqrt{(p - x_K)^2 + (q - y_K)^2} \quad (\text{Radial distance}) \quad (19)$$

$$MP_K = \sqrt{MK^2 - KP_K^2} \quad (\text{Diagonal distance}) \quad (20)$$

$$R_{Leb} = \frac{KP_K}{MP_K} \quad (\text{Distance ratio}) \quad (21)$$

The notation, R_{Leb} denote the distance ratio obtained by the projection of left eyebrow. Similarly the distance ratios R_{Le} , R_{Re} , R_{Reb} , R_{Nose} , R_{Mouth} , R_{Lear} and R_{Rear} are determined, respectively for left eye, right eye, right eyebrow, nose, mouth, left ear and right ear.

Distance Ratio Features in Combination

The distances of all the facial features along the DRL are used to compute the distance ratios for the combination of facial features as follows.

$$R_{Leye2Reye} = \frac{MP_{Leye}}{MP_{Reye}} \quad (\text{Left Eye to Right Eye}) \quad (22)$$

$$R_{Leb2Reb} = \frac{MP_{Leb}}{MP_{Reb}} \quad (\text{Left Eyebrow to Right Eyebrow}) \quad (23)$$

$$R_{N2M} = \frac{MP_n}{MP_m} \quad (\text{Nose to Mouth}) \quad (24)$$

$$R_{Lear2Rear} = \frac{MP_{Lear}}{MP_{Rear}} \quad (\text{Left Ear to Right Ear}) \quad (25)$$

Area Features

The centroids of the eyes, eyebrows, nose and mouth are connected in triangles as shown in the Figure 11. The areas covered by the triangles are used to determine the area features. In Figure 11(a), e_1 and e_2 denote right and left eyes respectively; n and m denote nose and mouth respectively. The coordinates (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , and (x_4, y_4) are the centroids of right eye, left eye, nose, and mouth respectively.



Figure 11. Triangular area features (a) Areas formed by eyes, nose, and mouth (b) Areas formed by eyebrows, nose, and mouth

The triangular area A_{en} formed by eyes and nose; and, the triangular area A_{em} formed by eyes and mouth are computed as given below.

$$A_{en} = 0.5 \begin{vmatrix} (x_1 - x_3) & (y_1 - y_3) \\ (x_2 - x_3) & (y_2 - y_3) \end{vmatrix} \quad \& \quad A_{em} = 0.5 \begin{vmatrix} (x_1 - x_4) & (y_1 - y_4) \\ (x_2 - x_4) & (y_2 - y_4) \end{vmatrix} \quad (26)$$

$$A_{Eyes} = \frac{A_{en}}{A_{em}} \quad (27)$$

Then the ratio of areas covered by eyes, nose and mouth is given by the equation (27). Similarly, in Figure 11(b), b_1 and b_2 denote right and left eyebrows respectively, and n and m denote nose and mouth respectively. The coordinates (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , and (x_4, y_4) are the centroids of right eyebrow, left eyebrow, nose, and mouth respectively. The triangular area A_{ebn} formed by eyebrows and nose; and, the triangular area A_{ebm} formed by eyebrows and mouth are computed as given below.

$$A_{ebn} = 0.5 \begin{vmatrix} (x_1 - x_3)(y_1 - y_3) \\ (x_2 - x_3)(y_2 - y_3) \end{vmatrix} \quad \& \quad A_{ebm} = 0.5 \begin{vmatrix} (x_1 - x_4)(y_1 - y_4) \\ (x_2 - x_4)(y_2 - y_4) \end{vmatrix} \quad (28)$$

$$A_{Eyebrows} = \frac{A_{ebn}}{A_{ebm}} \quad (29)$$

Then the ratio of areas covered by eyebrows, nose and mouth is given by the equation (29). The projected features are listed in the Table 6.

Feature	Description	Feature	Description
R_{Leye}	Distance ratio by left eye	R_{Rear}	Distance ratio by right ear
R_{Reye}	Distance ratio by right eye	$R_{Leye2Reye}$	Distance ratio by left and right eyes
R_{Leb}	Distance ratio by left eyebrow	$R_{Reb2Leb}$	Distance ratio by left & right eyebrows
R_{Reb}	Distance ratio by right eyebrow	R_{N2M}	Distance ratio by nose and mouth
R_{Nose}	Distance ratio by nose	$R_{Lear2Rear}$	Distance ratio by left ear and right ear
R_{Mouth}	Distance ratio by mouth	A_{Eyes}	Area ratio by eyes, nose and mouth
R_{Lear}	Distance ratio by left ear	$A_{Eyebrows}$	Area ratio by eyebrows, nose and mouth

Table 6. List of projected features

Final geometrical features include 26 features, in which 12 features are from the Table 5 and 14 features are from the Table 6.

3.2 Optimization of Features Sets

Three subsets of features from 26 features in different combinations are considered for optimization. The subset A, B, C consist of 14, 6, 14 features, respectively as given below.

$$Subset A = (R_{Leye}, R_{Reye}, R_{Leb}, R_{Reb}, R_{Nose}, R_{Mouth}, R_{Lear}, R_{Rear}, R_{Leye2Reye}, R_{Reb2Leb}, R_{N2M}, R_{Lear2Rear}, A_{Eyes}, A_{Eyebrows}) \quad (30)$$

$$Subset B = (E_{Eyes}, E, R_{Reb2Leb}, R_{N2M}, A_{Eyes}, A_{Eyebrows}) \quad (31)$$

$$Subset C = (\mu_{Leb}, \mu_{Reb}, \mu_{Nose}, \mu_{Mouth}, E_{Eyes}, E_{Leb}, E_{Reb}, E_{Nose}, E_{Mouth}, E, R_{Reb2Leb}, R_{Mouth2Nose}, A_{Eyes}, A_{Eyebrows}) \quad (32)$$

The every feature subset is optimized by the maximal distances between the classes and minimal distances between the patterns of one class. Here each class represents one person

and the different images of one person were considered as patterns. The effectiveness of every feature subset is determined by the evaluation function F as given below.

$$F = \frac{D_d}{D_m} = \frac{\sqrt{\sum_{i=1}^k (M_d - D_i)^2}}{\sqrt{\sum_{i=1}^k (M_m - M_i)^2}} \text{ where } M_i = \frac{\sum_{j=1}^n f_{ij}}{n}, M_m = \frac{\sum_{i=1}^k M_i}{k}, D_m = \sqrt{\frac{\sum_{i=1}^k (M_m - M_i)^2}{k-1}}$$

$$D_i = \sqrt{\frac{\sum_{j=1}^n (M_i - f_{ij})^2}{n-1}}, M_d = \frac{\sum_{i=1}^k D_i}{k}, D_d = \sqrt{\frac{\sum_{i=1}^k (M_d - D_i)^2}{k-1}} \quad (33)$$

where M_i and D_i are mean and variance of the feature values f_{ij} for ($j=1$ to k) k images of the i -th person respectively, M_m and M_d are mean of M_i and D_i respectively. The F value is the ratio of the measures of dispersion of sample standard deviations and of the sample means of the feature values in the k sample images of a class. For illustration we have used ORL face database, which contain 40 subjects or classes and each of 10 variations. The Figure 12 shows the optimization of feature subsets in which F values along the y-axis are plotted for 40 classes along the x-axis. The lower F value indicates the stronger invariance property of the feature subset with respect to scale, rotation and facial expressions. In the Figure 12 it shows that the feature subset C is well optimized with the lowest F values compared to other subsets and, hence it corresponds to a better feature subset.

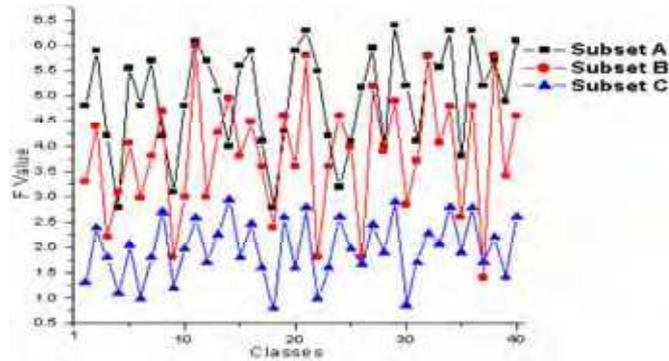


Figure 12. Optimization of subsets of features

Invariance Property

The above feature Subset C is considered as the most optimized geometrical feature vector for face recognition and is invariant to scaling, rotation, and facial expressions, because the relative geometrical distances between the facial features such as eyes, nose, mouth, and eyebrows vary proportionally with respect to scaling, rotation, and facial expressions, and their feature values remain nearly constant. Hence the optimized feature vector characterizes each human face uniquely. The Figure 13 illustrates the invariance property of

feature vectors for the images shown in Figure 13(a). The Figure 13(b), feature vectors exhibit negligible variations in the feature values.

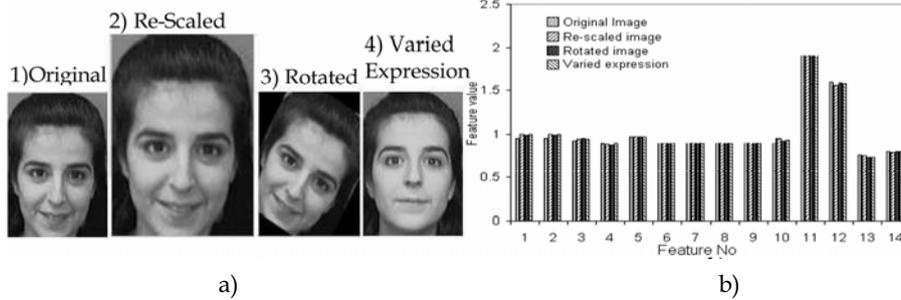


Figure 13. Illustration of invariance property a) Different images of the same person b) Feature vectors for the images in a)

4. Face Recognition

In automated face recognition, a human face can be described by several features, but very few of them are used in combination to improve discrimination ability and different facial features have different contributions in personal identification. The use of geometrical features will always have the credit of reducing huge space that is normally required in face image representation, which in turn increases the recognition speed considerably (Zhao et al. 2000). In (Hiremath and Danti, Jan 2006), the geometric-Gabor features extraction is proposed for face recognition and it is described in this section.

4.1 Gemetric-Gabor feature Extraction

In the human ability of recognizing a face, the local features such as eyes, eyebrows, nose and mouth dominate the face image analysis. In the present study, we have used geometrical features and Gabor features in combination for face recognition. The optimized feature set (Subset C) is considered as *Geometric-Features* for face recognition and the features are listed as below.

$$\begin{aligned} \text{Geometric Features} = (\mu_{Leb}, \mu_{Reb}, \mu_{Nose}, \mu_{Mouth}, E_{Eyes}, E_{Leb}, E_{Reb}, E_{Nose}, \\ E_{Mouth}, E, R_{Reb2Leb}, R_{Mouth2Nose}, A_{Eyes}, A_{Eyebrows}) \end{aligned} \quad (34)$$

The Gabor features are extracted by applying the Gabor filters on the facial feature locations as obtained by our face detector and these locations are considered as highly energized points on the face. We refer these Gabor features as *Geometric-Gabor Features* and the feature extraction process is as given below.

The local information around the locations of the facial features is obtained by the Gabor filter responses at the highly energized points on the face. A Gabor filter is a complex sinusoid modulated by a 2D Gaussian function and it can be designed to be highly selective in frequency. The Gabor filters resemble the receptive field profiles of the simple cells in the visual cortex and they have tunable orientation, radial frequency bandwidth and center frequency. The limited localization in space and frequency yields a certain amount of robustness against translation, distortion, rotation and scaling. The Gabor functions are

Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free /Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

