A Robust Skin Color Based Face Detection Algorithm

Sanjay Kr. Singh[1], D. S. Chauhan[2], Mayank Vatsa[3], Richa Singh[3]*

[1]*Department of Computer Science and Engineering*
Institute of Engineering and Technology
Jaunpur−222002, India
E-mail: sksiet@yahoo.com
[2] *Uttar Pradesh Technical University*
Institute of Engineering and Technology
Lucknow−226021, India
[3] *Department of Computer Science & Engineering*
Indian Institute of Technology
Kanpur−208016, India

Abstract

In this paper, a detailed experimental study of face detection algorithms based on "Skin Color" has been made. Three color spaces, RGB, YCbCr and HSI are of main concern. We have compared the algorithms based on these color spaces and have combined them to get a new skin color based face detection algorithm which gives higher accuracy. Experimental results show that the proposed algorithm is good enough to localize a human face in an image with an accuracy of 95.18%.

Key Words: Face Detection, Skin Color Classification, Color Space, Thresholding, Feature Extraction

1. Introduction

Human face perception is currently an active research area in the computer vision community. Human face localization and detection is often the first step in applications such as video surveillance, human computer interface, face recognition and image database management. Locating and tracking human faces is a prerequisite for face recognition and/or facial expressions analysis, although it is often assumed that a normalized face image is available. In order to locate a human face, the system needs to capture an image using a camera and a frame-grabber to process the image, search the image for important features and then use these features to determine the location of the face.

For detecting face there are various algorithms including skin color based algorithms. Color is an important feature of human faces. Using skin-color as a feature for tracking a face has several advantages. Color processing is much faster than processing other facial features. Under certain lighting conditions, color is orientation invariant. This property makes motion estimation much easier because only a translation model is needed for motion estimation. However, color is not a physical phenomenon; it is a perceptual phenomenon that is related to the spectral characteristics of electro-magnetic radiation in the visible wavelengths striking the retina. Tracking human faces using color as a feature has several problems like the color representation of a face obtained by a camera is influenced by many factors (ambient light, object movement, etc.), different cameras produce significantly different color values even for the same person under the same lighting conditions and skin color differs from person to person. In order to use color as a feature for face tracking, we have to solve these problems. It is also robust towards changes in orientation and scaling and can tolerate occlusion well. A disadvantage of the color cue is its sensitivity to illumination color changes and, especially in the case of RGB, sensitivity to illumination intensity. One way to increase tolerance toward intensity changes in images is to transform the RGB image into a color space whose intensity and chromaticity are separate and use only chromaticity part for detection.

In this paper we have presented a comparative study of three well known skin color face localization/detection algorithms and have given a new algorithm based on skin color classification in RGB, YCbCr and HSI color models. Results of comparative experiments show that there are some pros and cons in each of the algorithms. So we have taken the combination of the three results to find the skin region and then from the skin region facial features have been extracted to get the face from the skin region. Our results show that we can localize the face more effectively by using the proposed algorithm.

In the next section we have explained the color spaces used to classify the skin color. Third section explains the well known algorithms based on RGB, YCbCr and HSI color spaces and the proposed algorithm. Experimental results have been explained in fourth section and fifth section is conclusion.

2. Color Models for Skin Color Classification

The study on skin color classification has gained increasing attention in recent years due to the active research in content-based image representation. For instance, the ability to locate image object as a face can be exploited for image coding, editing, indexing or other user interactivity purposes. Moreover, face localization also provides a good stepping stone in facial expression studies.

It would be fair to say that the most popular algorithm to face localization is the use of color information, whereby estimating areas with skin color is often the first vital step of such strategy. Hence, skin color classification has become an important task. Much of the research in skin color based face localization and detection is based on RGB, YCbCr and HSI color spaces. In this section the color spaces are being described.

2.1 RGB Color Space

The RGB color space consists of the three additive primaries: red, green and blue. Spectral components of these colors combine additively to produce a resultant color.

The RGB model is represented by a 3-dimensional cube with red green and blue at the corners on each axis (Figure 1). Black is at the origin. White is at the opposite end of the cube. The gray scale follows the line from black to white. In a 24-bit color graphics system with 8 bits per color channel, red is (255, 0, 0). On the color cube, it is (1, 0, 0).

The RGB model simplifies the design of computer graphics systems but is not ideal for

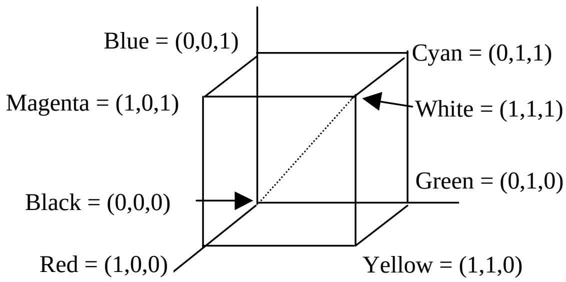

Figure 1. RGB color cube

all applications. The red, green and blue color components are highly correlated.

This makes it difficult to execute some image processing algorithms. Many processing techniques, such as histogram equalization, work on the intensity component of an image only.

2.2 YCbCr Color Space

YCbCr color space has been defined in response to increasing demands for digital algorithms in handling video information, and has since become a widely used model in a digital video.

It belongs to the family of television transmission color spaces. The family includes others such as YUV and YIQ. YCbCr is a digital color system, while YUV and YIQ are analog spaces for the respective PAL and NTSC systems. These color spaces separate RGB (Red-Green-Blue) into luminance and chrominance information and are useful in compression applications however the specification of colors is somewhat unintuitive.

The Recommendation 601 specifies 8 bit (i.e. 0 to 255) coding of YCbCr, whereby the luminance component Y has an excursion of 219 and an offset of +16. This coding places black at code 16 and white at code 235. In doing so, it reserves the extremes of the range for signal processing footroom and headroom. On the other hand, the chrominance components Cb and Cr have excursions of +112 and offset of +128, producing a range from 16 to 240 inclusively.

2.3 HSI Color Space

Since hue, saturation and intensity are three properties used to describe color, it seems logical that there be a corresponding color model, HSI. When using the HSI color space, you don't need to know what percentage of blue or green is required to produce a color. You simply adjust the hue to get the color you wish. To change a deep red to pink, adjust the saturation. To make it darker or lighter, alter the intensity.

Many applications use the HSI color model. Machine vision uses HSI color space in identifying the color of different objects. Image processing applications such as histogram operations, intensity transformations and convolutions operate only on an intensity image. These operations are performed with much ease on an image in the HSI color space.

For the HSI being modeled with cylindrical coordinates, see Figure 2. The hue (*H*) is represented as the angle 0, varying from 0^o to 360^o. Saturation (*S*) corresponds to the radius, varying from 0 to 1. Intensity (*I*) varies along the *z* axis with 0 being black and 1 being white.

When $S = 0$, color is a gray value of intensity 1. When $S = 1$, color is on the boundary of top cone base. The greater the saturation, the farther the color is from white/gray/black (depending on the intensity). Adjusting the hue will vary the color from red at 0^o, through green at 120^o, blue at 240^o, and back to red at 360^o. When $I = 0$, the color is black and therefore *H* is undefined. When $S = 0$, the color is grayscale. *H* is also undefined in this case.

By adjusting *I*, a color can be made darker or lighter. By maintaining $S = 1$ and adjusting *I*, shades of that color are created.

3. Algorithms

In this section we have described the three algorithms [1-3] which are based on RGB, YCbCr and HIS respectively and a new algorithm based on the combination of these three algorithms. It has been stated that the above three algorithms work very well under the condition that there is only one face is present in the image. In the implementation of the algorithms there are three main steps viz. (1) Classify the skin region in the color space, (2) apply threshold to mask the skin region and (3) draw bounding box to extract the face image.

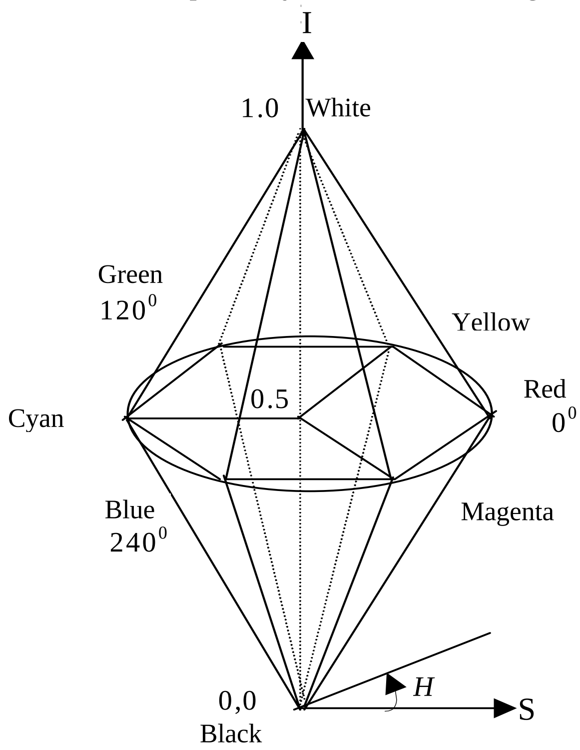

Figure 2. Double cone model of HSI color space

3.1 Skin Color Based Face Detection in RGB Color Space

Crowley and Coutaz [1] said one of the simplest algorithms for detecting skin pixels is to use skin color algorithm. The perceived human color varies as a function of the relative direction to the illumination. The pixels for skin region can be detected using a normalized color histogram, and can be further normalized for changes in intensity on dividing by luminance. And thus converted an [R, G, B] vector is converted into an [*r, g*] vector of normalized color which provides a fast means of skin detection. This gives the skin color region which localizes face. As in [1], the output is a face detected image which is from the skin region. This algorithm fails when there are some more skin region like legs, arms, etc.

3.2 Skin Color Based Face Detection in YCbCr Color Space

We have implemented a skin color classification algorithm [2] with color statistics gathered from YCbCr color space. Studies have found that pixels belonging to skin region exhibit similar Cb and Cr values. Furthermore, it has been shown that skin color model based on the Cb and Cr values can provide good coverage of different human races. The thresholds be chosen as $[Cr_1, Cr_2]$ and $[Cb_1, Cb_2]$, a pixel is classified to have skin tone if the values $[Cr, Cb]$ fall within the thresholds The skin color distribution gives the face portion in the color image. This algorithm is also having the constraint that the image should be having only face as the skin region.

3.3 Skin Color Based Face Detection in HSI Color Space

Kjeldson and Kender defined a color predicate in HSV color space to separate skin regions from background [3]. Skin color classification in HSI color space is the same as YCbCr color space but here the responsible values are hue (H) and saturation (S). Similar to above the threshold be chosen as [H_1, S_1] and [H_2, S_2], and a pixel is classified to have skin tone if the values [H,S] fall within the threshold and this distribution gives the localized face image. Similar to above two algorithm this algorithm is also having the same constraint.

3.4 Proposed Algorithm for Face Detection

It is assumed that by combining the detected regions from all the three algorithms, skin region is extracted. Thus, the three algorithms are combined assuming that their combination gives the skin region from the image and from the skin detected image face is extracted by first extracting facial features and then drawing a bounding box around the face region with the help of facial features. The assumption can thus be stated as, if the skin image is detected by one or more algorithm(s) and for the same image other algorithm gives the false result, then also the face is extracted using the combination algorithm. This assumption is based on the basic idea of Venn diagram from Set Theory. If we state that the result from RGB color space is "A", the result from YCbCr color space is "B" and the result from HSI color space is "C" and if any of the result contains a skin image then the union of the three will surely be a skin image. This can be further explained as follows: let, A = {a, b, c, d}, B = {b, e} and C = {b, f}. If "a" = "skin image" then by the first algorithm the result will be true but next two will be failed similarly, if we take the case if "e" or "f" = "skin image". But if we take the union of the three then D = (AUBUC) = {a, b, c, d, e, f} and if either of "a", "e" or "f" is skin image then this algorithm can detect it. Now the case is that if all the three algorithms have the same result then this is the case if "b" = "skin image", then also D will be having the element b. By this we can detect the skin region from a color image. Figure 3 shows the steps of the proposed algorithm for skin color classification.

After getting the skin region, facial features viz. Eyes and Mouth are extracted. The image obtained after applying skin color statistics is subjected to binarization i.e., it is transformed to gray-scale image and then to a binary image by applying suitable threshold. This is done to eliminate the hue and saturation values and consider only the luminance part. This luminance part is then transformed to binary image with some threshold because the features we want to consider further for face extraction are darker than the background colors. After thresholding, opening and closing operations are performed to remove noise. These are the morphological operations, opening operation is erosion followed by dilation to remove noise and closing operation is dilation followed by erosion which is done to remove holes. Now we extract the eyes, ears and mouth from the binary image by considering the threshold for areas which are darker in the mouth than a given threshold.

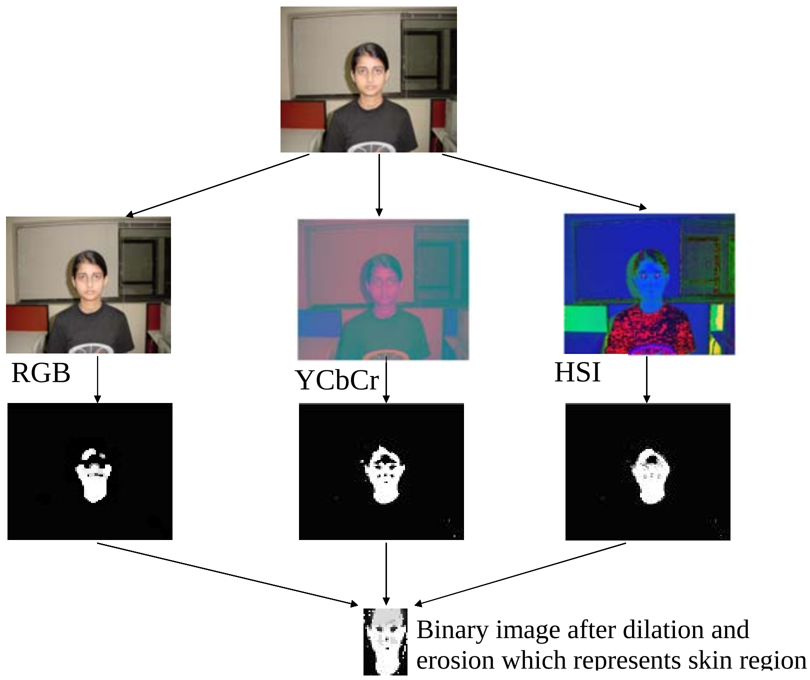

Figure 3. Skin color detection from proposed algorithm

A triangle drawn with the two eyes and a mouth as the three points in case of a frontal face then we see that we get an isosceles triangle (*i j k*) in which the Euclidean distance between two eyes is about 90-110% of the Euclidean distance between the centre of the right/left eye and the mouth.

After getting the triangle, it is easy to get the coordinates of the four corner points that form the potential facial region. Since the real facial region should cover the eyebrows, two eyes, mouth and some area below the mouth, the coordinates can be calculated as follows: Assume that (Xi, Yi), (Xj, Yj) and (Xk, Yk) are the three center points of blocks i, j, and k, that form an isosceles triangle. ($X1$, Y1), ($X2$, Y2), ($X3$, Y3) and ($X4$, Y4) are the four corner points of the face region. $X1$ and $X4$ locate at the same coordinate of (Xi – 1/3D(i, k)); $X2$ and $X3$ locate at the same coordinate of (Xk + 1/3D(i, k)); Y1 and Y2 locate at the same coordinate of (Yi + 1/3D(i, k)); Y3 and Y4 locate at the same coordinate of (Yj – 1/3D(i, k)); where D(i, k) is the Euclidean distance between the centers of block i (right eye) and block k (left eye).

$$X1 = X4 = Xi - 1/3D(i, k) \quad (1)$$

$$X2 = X3 = Xk + 1/3D(i, k) \quad (2)$$

$$Y1 = Y2 = Yi + 1/3D(i, k) \quad (3)$$

$$Y3 = Y4 = Yj - 1/3D(i, k) \quad (4)$$

This is done for a frontal face; similarly for the side (left/right) view. For the side view a right triangle has obtained with the characteristic that the Euclidean distance of line *i k* is equal to 2 times of the Euclidean distance of line *j k*, and the Euclidean distance of line *i j* is equal to 1.732 times of the Euclidean distance of line *j k*. The 4 rules to get the face boundary for right side view are as follows

$$X1 = X4 = Xi - 1/6D(i, j) \quad (5)$$

$$X2 = X3 = Xi + 1.2D(i, j) \quad (6)$$

$$Y1 = Y2 = Yi + 1/4D(i, j) \quad (7)$$

$$Y3 = Y4 = Yi - 1.0D(i, j) \quad (8)$$

Similarly, the 4 rules to get the face boundary for left side view are:

$$X1 = X4 = Xj - 1/6D(j, k) \quad (9)$$

$$X2 = X3 = Xj + 1.2D(j, k) \quad (10)$$

$$Y1 = Y2 = Yj + 1/4D(j, k) \quad (11)$$

$$Y3 = Y4 = Yj - 1.0D(j, k) \quad (12)$$

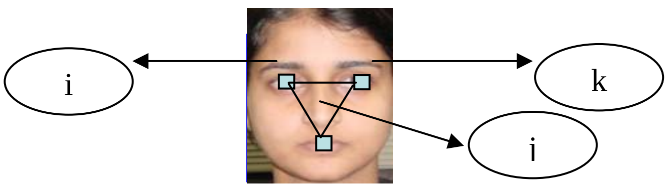

Figure 4. Three points (*i*, *j*, and *k*) satisfy the matching rules, which will form an isosceles triangle.

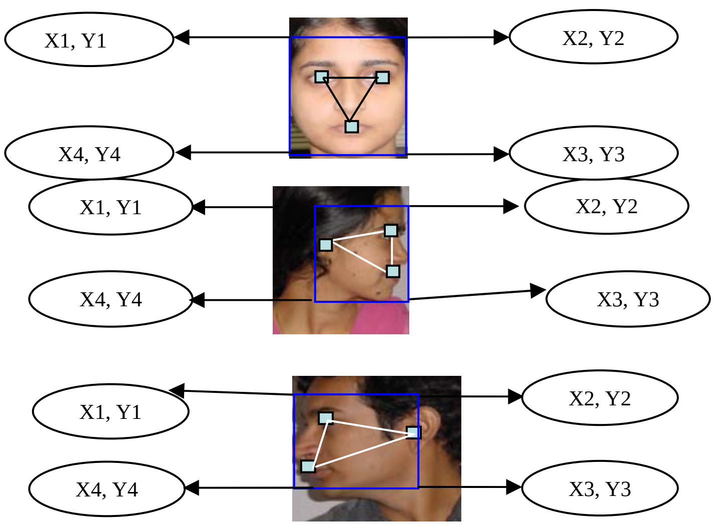

Figure 5. Drawing bounding box

4. Experimental Results

In this section a detailed experimental comparison of the above stated algorithms has been presented. We have used two color image databases: (1) database prepared in our conditions, (2) IITK face database [4] and some other images obtained from internet. The accuracy is obtained in all the four cases by using the following equation

% Accuracy = 100 – (False Detection Rate + False Dismissal Rate)

4.1 Results of Skin Color Based Face Detection in RGB Color Space

Results of this experiment show that RGB color space is not very much friendly with face detection based on skin color classification. The accuracy of this experiment is found to be 56.46%. Following table shows that the false detection rate and dismissal rate are very high, thus causing very low accuracy in detecting the face.

In Figure 6, the colored part shows the skin color detection in an image. We have represented the image as a web and the colored part is represented as skin region. In RGB color space it is found that it represents some non-skin region also as the skin region and hence false detection rate is quite high.

4.2 Results of Skin Color Based Face Detection in YCbCr Color Space

Similar experiments have been performed on the YCbCr color space as on RGB. The accuracy is found to be 83.91% which is far better than results from RGB color space. Following table shows that the false detection rate is high and dismissal rate is low, thus causing a bit low accuracy in detecting the face.

In Figure 7, image representation in web form for YCbCr color space is shown; it is found here that skin color region is more effectively extracted. This is because Cb and Cr have some distinct color range for skin region. Thus the accuracy of this algorithm is quite good.

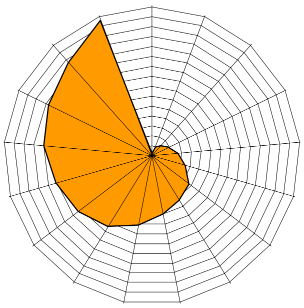

Figure 6. Skin color representation in RGB color space

Table 1. Skin color classification results for RGB color space

No. of Images	False Detection Rate (%)	False Dismissal Rate (%)
1100	25.45	18.09

Table 2. Skin color classification results for YcbCr color space

No. of Images	False Detection Rate (%)	False Dismissal Rate (%)
1100	12.82	3.27

Table 3. Skin color classification results for HSI color space

No. of Images	False Detection Rate (%)	False Dismissal Rate (%)
1100	14.55	3.18

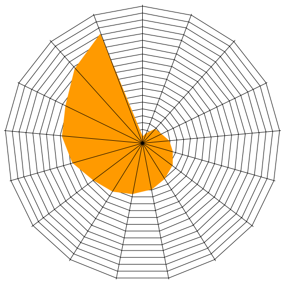

Figure 7. Skin color representation in YCbCr color space

4.3 Results of Skin Color Based Face Detection in HSI Color Space

Experiments show that HSI color space is also good in classifying the skin color region. The accuracy is found to be 82.27% which is nearly equivalent to results from YCbCr color space. Similar to YCbCr, in this color space also false detection rate is high and dismissal rate is low.

Image representation in Figure 8 is in HSI color space, it is found that skin color region is more effectively extracted as in YCbCr color space because H and S (similar to Cb and Cr) have some distinct color range for skin region.

4.4 Results of Proposed Algorithm

The experiments show very good results for proposed algorithm show in Figure 9. The results obtained using the previous conditional probabilities and threshold value. The false detection and false dismissal rate are shown in Table 4. The low false dismissal rate shows the robustness of the algorithm. The relative false detection is also low and this shows that it is able to distinguish between

Table 4. Skin color classification results for combination algorithm

No. of Images	False Detection Rate (%)	False Dismissal Rate (%)
1100	3.09	1.73

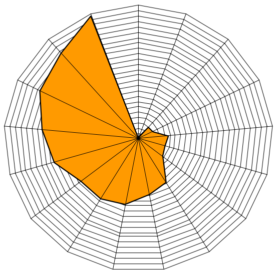

Figure 8. Skin color representation in HSI color space

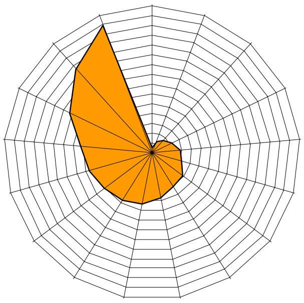

Figure 9. Skin color representation – proposed algorithm

actual skin and background color with skin color appearance. The accuracy is found to be 95.18%.

The web representation also shows that this algorithm is capable of classifying skin region in complex colored image. Sample results from proposed algorithm are shown in Figure 10.

4.5 Comparison of Algorithms

In this section comparison of the four algorithms is given. Table 5 shows the comparison of the algorithms. This comparison shows that the proposed algorithm gives good results but the time consumption is more than the others. Figure 11 shows a comparative graph based on the experimental results of the comparative performance of all the algorithms.

5. Conclusion

In this paper a comparison has been made for detecting faces in the controlled background, using skin color detection on RGB, YCbCr and HSI color spaces. We have found that YCbCr and HSI color space are more efficient in comparison to RGB to classify the skin region. But still both are not able to give very good results. Based on the results of the three algorithms, we have combined from this region(s) facial feature (eyes, ear and

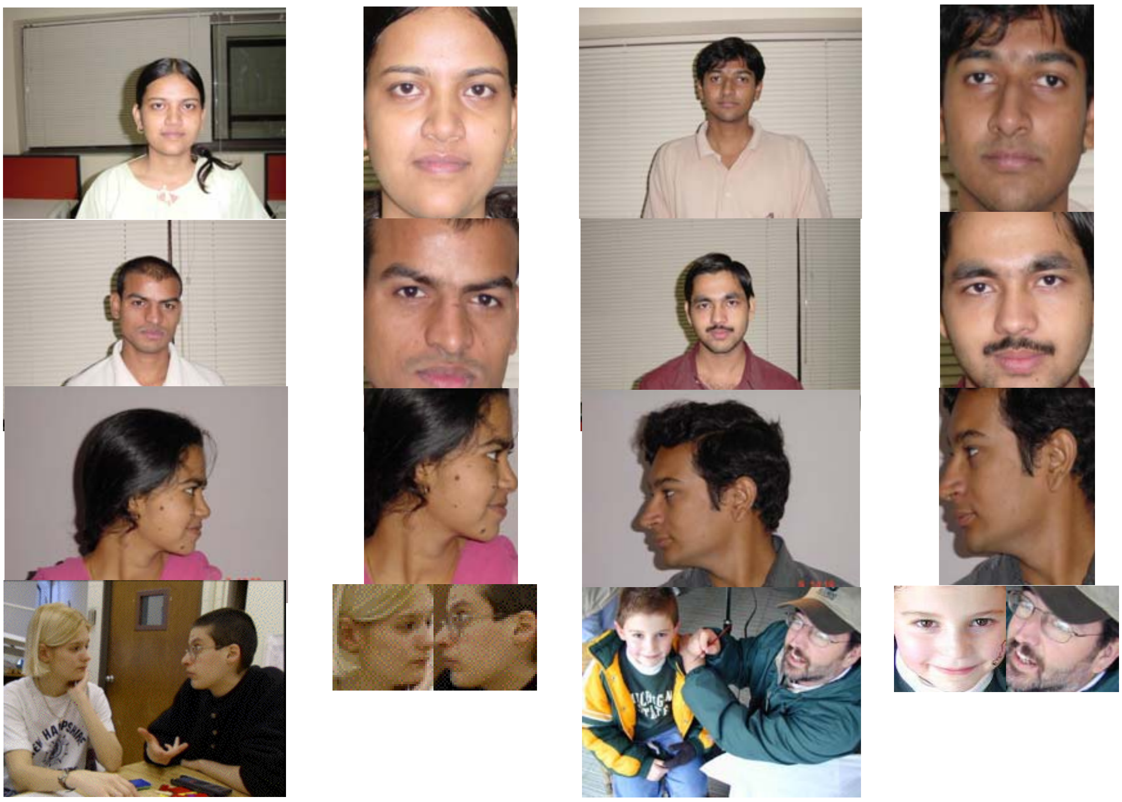

Figure 10. Sample results of proposed face detection algorithm

Table 5. Comparison Chart of the Algorithms

Criterion	RGB Color Space	YcbCr Color Space	HIS Color Space	Proposed Algorithm
Frontal Face	56.46 %	83.91 %	82.27 %	96.01 %
Tilted/ Rotated Face	54.57 %	80.14 %	80.09 %	92.42 %
Profile Face	47.84 %	80.11 %	79.92 %	91.29 %
Complex Background Image	42.62 %	73.72 %	71.19 %	95.18 %
Time Consumption	2.09 sec	3.46 sec	3.52 sec	6.38 sec

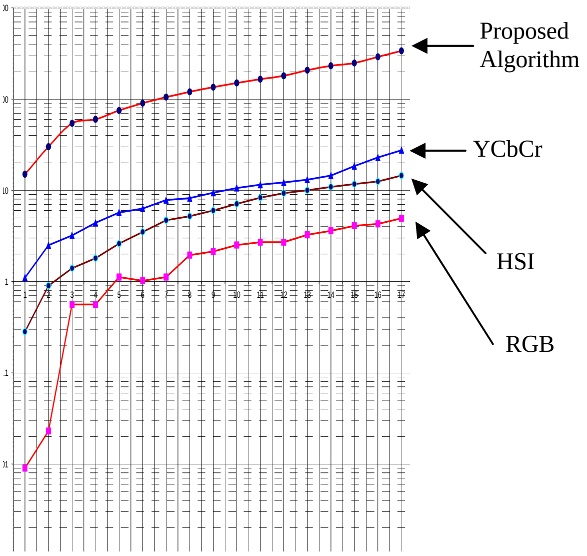

Figure 11. Comparative chart (The lowest dotted line represents the results of RGB color space, 2 and 3 represent the results of HSI and YCbCr respectively, and the upper one shows the results of proposed algorithm.)

mouth) has been extracted from a proposed facial feature extraction algorithm which finally gives the detected face. It is robust and efficient in classifying the skin color region and face region. Accuracy of proposed algorithm is 95.18%.

Acknowledgment

Authors wish to acknowledge Dr. P. Gupta for providing the IITK face database.

References

[1] Crowley, J. L. and Coutaz, J., "Vision for Man Machine Interaction," *Robotics and Autonomous Systems*, Vol. 19, pp. 347-358 (1997).

[2] Cahi, D. and Ngan, K. N., "Face Segmentation Using Skin-Color Map in Videophone Applications," *IEEE Transaction on Circuit and Systems for Video Technology*, Vol. 9, pp. 551-564 (1999).

[3] Kjeldsen, R. and Kender., J., "Finding Skin in Color Images," *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, pp. 312-317 (1996).

[4] http://www.cse.iitk.ac.in/users/pg/

[5] Cai, J., Goshtasby, A. and Yu, C., "Detecting Human Faces in Color Images," *Proceedings of International Workshop on Multi-Media Database Management Systems*, pp. 124-131 (1998).

[6] Yang, J., Lu, W. and Waibel A., "Skin Color Modeling and Adaptation," (CMU-CS-97-146,) CS Department, CMU, PA, U.S.A. (1997).

[7] Yang, M. H. and Ahuja., N., "Detecting Human Faces in Color Images", *Proceedings of IEEE International Conference on Image Processing*, Vol. 1, pp. 127-130 (1998).

[8] Yang, M. H., Kriegman, J. and Ahuja, N., "Detecting Faces in Images: A Survey," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 24, pp. 34-58 (2002).

Manuscript Received: Jun. 2, 2003
Revision Received: Jul. 25, 2003
Accepted: Sep. 15, 2003