Real Time Detection and Tracking of Human Face Using Skin Color Segmentation and Region Properties

Prashanth Kumar G. and Shashidhara M.
Dept. of E&C Channabasaveshwara Institute of Technology, Gubbi, Tumkur-572216, Karnataka, India
Email: prashanth_02@rediffmail.com, mrshashi4u@gmail.com

Abstract—**Real time face detection and face tracking is one of the challenging problem in computer human interaction, video surveillance, biometrics etc. In this paper we are presenting an algorithm for real time face detection and tracking using skin color segmentation and region properties. First segmentation of skin regions from an image is done by using different color models. Skin regions are separated from the image by using thresholding. Then to decide whether these regions contain human face or not we used face features. Our procedure is based on skin color segmentation and human face features (knowledge-based approach). We have used RGB, YCbCr, and HSV color models for skin color segmentation. These color models with thresholds, help to remove non skin like pixel from an image. Each segmented skin regions are tested to know whether region is human face or not, by using human face features based on knowledge of geometrical properties of human face.**

Index Terms—**face detection, color models, bounding box, centroid, euler number, color image segmentation**

I. Introduction

Face detection is a technology that can be applied in a wide variety of fields like image monitoring systems, remote meeting systems, hci(human computer interaction) and face recognition systems. However, a face image contains significant changes due to diversity such as lighting, change in visual point, expression, hair style, make up and glasses. Also, since countless non-facial data that are similar to face data exist, there is a realistic limit in separating facial domain completely from complex background [1]. Existing face detection technology includes the method that uses neural nets [2], a color based method [3], the Gaussian compound model based method [4] and a feature based method [5], [6]. The method that uses neural nets finds a face from a still image.

Several applications, such as face processing (i.e. face, expression, gender classification and gesture recognition), computer human interaction, human crowd surveillance, biometric, video surveillance, artificial intelligence and content-based image retrieval etc. All of these applications, stated above, require face detection, which can be simply viewed as a preprocessing step, for obtaining the "object".

In other words, many of the techniques are proposed for these applications assume that, the location of the face is pre-identified and available for the next step. First problem is come in the way of face detection, is chosen proper color model for skin color segmentation. There are several color models and each has specific work field and strength. We used three color models for skin color segmentation [7], these are RGB, YCbCr and HSV color models. After skin like pixels detection, we convert this segmented image into binary form. This binary image contains skin regions, but we don't know that, where is human face in segmented image. Next step of face detection after segmentation is to extract only the face region by rejecting the non face region in the image based on knowledge of human face. Work of next step is to remove non human face skin area from segmented image, by using Knowledge-based methods or human face features and region properties. This paper is organized as follows, in section 2 we discuss the different 'color spaces and conversions', while section 3 presents 'proposed algorithm', section 4 explains the 'implementation of our method' and section 5 gives the experimental results. Finally the proposed work is concluded in section 6.

II. Color Space Representation

A Skin color modeling the inspiration to use skin color analysis [8] for initial classification of an image into probable face and no face regions stems from a number of simple but powerful characteristics of skin color. Firstly, processing skin color is simpler than processing any other facial feature. Secondly, under certain lighting conditions, color is orientation invariant. The major difference between skin tones is intensity e.g. due to varying lighting conditions and different human race. The color of human skin is different from the color of most other natural objects in the world. An attempt to build comprehensive skin and non-skin models has been done in. One important factor that should be considered while building a statistical model for color is the choice of a Color Space. Segmentation of skin colored regions becomes robust only if the chrominance component is used in analysis. Therefore, we eliminate the variation of luminance component as much as possible by choosing the CbCr plane (chrominance) of the YCbCr color space to build the model [9]. Another reason for the choice of YCbCr

domain is its extensive use in digital video coding applications. Research has shown that skin color is clustered in a small region of the chrominance plane.

The study on skin color classification system 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 for 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 HSV color spaces model [9]. In this section the color spaces model are being described.

A. RGB Color Space

The RGB color space consists of the three additive primaries: Red, Green and Blue. These spectral components are used to produce resultant color. Fig. 1 shows the RGB model represented by a 3-dimensional cube with red, green and blue at the corners on each axis. Black is at the origin and 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 represented as (255, 0, 0) and on the color cube it is (1, 0, 0).

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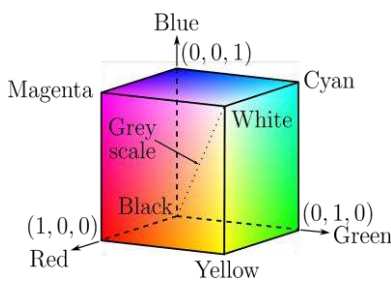

Figure 1. RGB color cube

B. YCbCr Color Space

YCbCr color space shown in Fig. 2 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 colors spaces separate RGB (Red-Green-Blue) into luminance and chrominance information and are useful in compression applications. 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. 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.
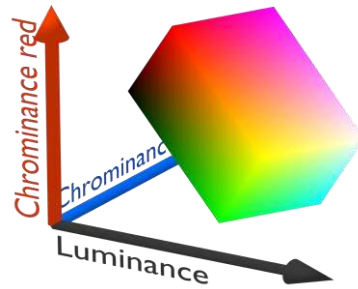

Figure 2. YCbCr color space

C. HSV Color Space

Since hue, saturation and intensity value are three properties used to describe color, it seems logical that there be a corresponding color model, HSV. When using the HSV color space, there is no need to know what percentage of blue or green is required to produce a color. Simply the hue is adjusted to get the desired color. To change a deep red to pink, the saturation is adjusted. To make it darker or lighter the intensity value is changed.

Many applications use the HSV color model. Machine vision uses HSV 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 HSV color space.
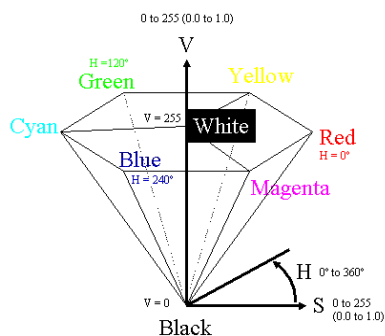

Figure 3. HSV color model

Fig. 3 shows the HSV color space modeled with cylindrical coordinates system. The hue H is represented as the angle varying from 0^0 to 360^0. Saturation S corresponds to the radius, varying from 0 to 1. Value of Intensity V varies along the z axis with 0 being black and 1 being white. When S is 0, color is a gray value of intensity 1. When S is 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 value of

intensity. Adjusting the hue will vary the color from red at 0^0, through green at 120^0, blue at 240^0 and back to red at 360^0.

III. PROPOSED ALGORITHM

The block diagram of real time face detection and tracking system is shown in Fig. 4.

Block diagram description: It consists of Image acquisition, Face segmentation, Face detection, Feature extraction and face tracking.

Image acquisition: The system takes video stream or photographic images from camera as input. The output consists of an array of rectangles which corresponds to the location and scale of faces detected, if it detects no faces it will return an empty array.

Face segmentation: color is an important feature of human faces. Using skin color we can segment the human body parts from the image.

Face detection: It is concerned with find whether any faces are there in a captured frame. For detecting face there are various algorithms available including skin color algorithms based on skin color classification in YC_bC_r and HSV color models.

Feature extraction: Facial features are extracted from the detected face region. After locating the face, the face region is processed for feature extraction to extract facial features positions such as eyes and nose tip position.

Face tracking: to track the face region first we rejected the non face skin regions by using the geometric features such as area, Euler number, Eccentricity, bounding box and cetroid.
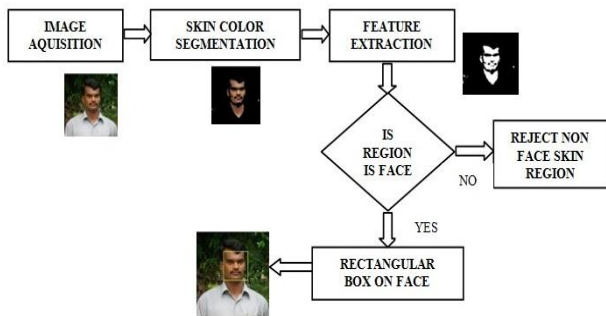

Figure 4. Block diagram of real time face detection and tracking system.

IV. IMPLEMENTATION

A. Image Transformation

The main function of human face tracker is locating the face in the given image which will be solved by using skin color methods mentioned in section 2. After locating the face, the face region is processed for feature extraction to extract facial features positions such as eyes and nose tip positions. In order to obtain those facial features, thresholding process is used. Thresholding process is used here because it is reasonably fast in computation which is compulsory for real-time human face tracker. From the thresholding result, the next process is to determine which area is the real facial feature by computing each object based on its area and circumference from each object. If the computation result is in tolerance range, then the object can be categorized as a feature.

B. Face Detection Using HSV Model

Human skin color ranged from very dark to nearly color less appearing pinkish-white due to the color of the blood under the skin in different people. In general, people whose ancestors come from regions will have darker skin than people whose ancestors come from regions with less sunlight. On average, women have slightly lighter skin than men. However, color changes in one race have been evolved based on genetic variation, skin color and culture changes. In order to process human face tracker, face skin color sample must be taken and RGB value from the image must be changed to HSV color system with the following equations.

After HSV value from each pixel is computed, each pixel is examined whether that pixel can be classified as a skin pixel or non skin pixel. The authors have worked on several images of different persons and found that the H value should be in the range 0 to 0.1 and S should be in the range 0.2 to 0.7 for a pixel to be classified as a skin pixel. By applying this range non skin pixels are removed from the image. Hue value does not change much for different skin colors. Black skin and white skin only influence the saturation of the HSV color system. The reported facial feature detection rate is about 90% with near-frontal facial images.

C. Face Detection Using YCbCr Model

YCbCr is an encoded nonlinear RGB signal, commonly used by European television studios and for image compression work. Color is represented by luma which is luminance, computed from nonlinear RGB, constructed as a weighted sum of the RGB values, and two color difference values Cr and Cb that are formed by subtracting luma from RGB red and blue components.

$$Y=0.299R+0.587G+0.114B \qquad (1)$$

$$Cb=128+(-0.169R+331G+0.500B) \qquad (2)$$

$$Cr=128+(0.500R-0.419G-0.081B) \qquad (3)$$

Thresholding values for skin segmentation are:

$$125<=cr<=165 \qquad (4)$$

$$76<=cb<=126 \qquad (5)$$

$$0.01<=hue<=0.1 \qquad (6)$$

The transformation simplicity and explicit separation of luminance and chrominance components makes this color space attractive for skin color modeling. We have worked on several images of different persons and found that the Cr value should be in the range 125 to 165 and Cb should be in the range 76 to 126 for a pixel to be classified as a skin pixel. The reported facial feature detection rate is about 95% with near-frontal facial image. In addition to YCbCr model we have used HSV model also to get the better classification between skin colors with other. We have observed that in hue plane skin region takes the value between 0.01 to 0.1. These values are taken as the threshold values to separate the skin part in image. After

thresholding the image is converted to binary image to find the properties and to extract the face feature.

D. *Rejection of non Face Skin Region*

This step is the backbone of this algorithm. A binary image contains skin regions. This step decides that, which region is most probable human face region by using human face features methods and region properties of human face. These are area of human face, Euler number, bounding box properties, centroid, eccentricity like oval estimation and combination features of these methods. A binary image of skin color segmented image with skin regions passes through this step one by one, then these stages of this step remove non human face like regions by using their properties which are fixed according to the human face. Finally, when binary image is totally passed from this step, image contains only highly probable human faces like regions. We retrieve the locations of these human faces from a filtered or passed binary image from this step. We track the face part by plotting bounding box around these highly probable human faces in original image. Here, we describing following non human face rejections methods:

Small area: We calculated average area of human skin like regions of binary image, And compare this average area with each skin regions of binary image, if any skin regions is less than average value, then that skin region will be rejected [10]. Area of any skin region is calculated by counting no of skin region pixels. This method is helpful for removing very small skin regions from a binary image.

Euler number: Human faces contain some holes like eyes, eyebrows, a mouth, mustaches etc. the two eyes of human face are major holes for human faces. If any skin region doesn't contain any holes then, we can safely discard these type all skin region. The rejection is done by Euler number method [10], the number of holes in a region is computed using the Euler number of region, the formula for Euler number is:

$$E = C - H \qquad (7)$$

where *E* is Euler number of skin region is the total number of object in the image minus total number of holes in the object. *C* is the number of connected component, and, H is number of holes in a region. We can reject these type of skin regions which having Euler value greater than zero.

Eccentricity: After rejection of non human face region from above method, now the image comes in this stage. Generally, shape of human face is likely to oval shape, so whose region which have shape probable like to oval shape these are not rejected by this method, and those regions whose shape likely to line are rejected [11]. For finding shape of skin regions, we used region properties based eccentricities MATLAB function, for each region, the function gives its eccentricity value. An ellipse (skin region) whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment. The oval shape of a face can be approximated by an ellipse so we calculate Eccentricity of all skin connected regions and to discard all skin regions whose eccentricity greater than by 0.89905.

Bounding box properties: After passing from above methods, the binary image now passed from this stage, this stage reject non human face skin region based on height to width ratio [12]. Generally, height to width ratio of skin regions is measurable factor because it is also big factor for rejecting non human face like regions. If height to width ratio of skin region is less than by threshold value, then this skin region will be discarding from class of probable human faces. Here we decided 1.902 (obtained by number of trials and error) threshold values for height to width ratio. For determining height to width ratio of each skin region, we used region properties based Bounding Box MATLAB function.

Combination of bounding box properties and area: Above said method Bounding Box Properties work to maximum for rejecting non human face skin regions for those regions whom are not satisfy threshold value. In some typical cases, some region that don't belong probable human face class, but these regions satisfy height to width ration, and so these type of regions can't be rejected. For rejecting these types of regions, we introduced a method which is combination of Bounding Box method and area method. We calculate skin region area bounded by bounding box and also calculate this bounding box area. Bounding box area is the multiplication of height and width of bounding box. Skin area (skin region pixels bounded by bounding box) is less than rectangular or bounding box area by two times, then this skin region will be rejected.

Centroid: Generally, the human faces are evenly distributed in the centre, means, that the human faces are not present in the side of images. Therefore, the centroid of a face region should be found in a small window centered in the middle of the bounding box. We calculate y axis average centroid of a getting image after above methods. If centroids (Y coordinate) of skin regions are below and above by thresholds values, then these type of skin region will be rejected.

These steps are repeated for each frame of the captured video sequence to detect and track the human face in real time. The results we have obtained in the proposed method are presented in the next section.

V. EXPERIMENT RESULTS

We have implemented our face detection and tracking system using MATLAB software. First we tested our algorithm on images. After getting faithful results for images we tested our algorithm for video sequences. We got the better results from the system as we expected. The results of different steps involved in the proposed algorithm for image are shown in Fig. 5.
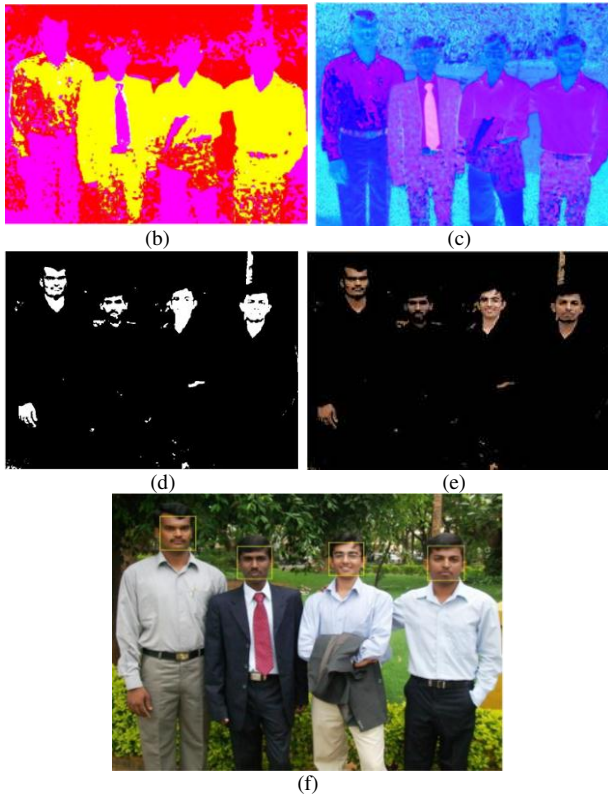

(a)

Figure 5. Results obtained in different stages of the proposed method.
(a) Input image. (b) Output of HSV model. (c) Output of YCbCr model.
(d) Skin Segmented Image after Thresholding. (e) Skin extracted image.
(f) Face Tracked Image

The results of the proposed algorithm for real time video sequence are shown in Fig. 6.
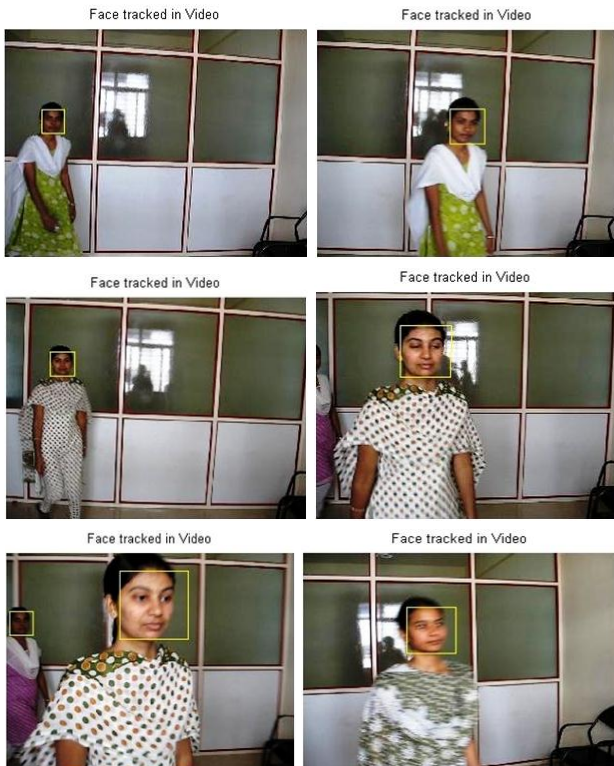

Figure 6. Real time face tracking results in a video sequence

VI. CONCLUSION

In this paper skin color based face detection and tracking system has been presented. The system is based on first segmenting the skin regions in an image. Then by using region properties face regions are extracted and tracked for further application. Our system gave the better results in good lighting condition. Under poor or low lighting condition performance of the proposed system is satisfactory. The system is intended for video surveillance applications and is able to tell whether or not a specific person is present in a video sequence shot. The system can be used in a range of applications like PC login security, robotics, attendance marking system, terrorists screening, passport authentication etc.

ACKNOWLEDGMENT

REFERENCES

[1]   J.-M. Kim, K.-H. Kim, and Ma.-K. Song, "Real time face detection and recognition using rectangular feature based classifier and modified matching algorithm," in *Proc. Fifth International Conference on Natural Computation*, 2009, pp. 171-175.

[2]   A. P. Sreedevi and B. S. S Nair, "Image processing based real time vehicle theft detection and prevention system," in *Proc. IEEE International Conference on Image Processing*, 2011, pp. 1-6.

[3]   W. L. Cheong, C. M. Char, Y. C. Lim, S. Lim, and S. W. Khor, "Building a computation savings real-time face detection and recognition system," in *Proc. 2nd International Conference on Signal Processing Systems*, 2010.

[4]   M. S. Devi and P. R. Baja, "Active facial tracking," in *Proc. Third International Conference on Emerging Trends in Engineering and Technology*, 2012.

[5]   M. Turk and A. Pentland., "Face recognition using eigenfaces," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Maui, Hawaii, Jun. 1991, pp. 586-591.

[6]   M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neurosicence*, vol. 3, no. 1, pp. 71-86, 1991.

[7]   A. Khaparde, S. Reddy, and Y. S. Ravipudi, "face detection using color based segmentation and morphological processing – A case study," in *Whitepaper Access Control, Technology sec watch, Whitepaper*, Sep. 2007.

[8]   K. H. B. Ghazali, "An innovative face detection based on skin color segmentation," *International Journal of Computer Applications*, vol. 34, no. 2, Nov. 2011.

[9]   B. Z. Wang, X. Y. Chang, and C. X. Liu, "Skin detection and segmentation of human face in color images," *International Journal of Intelligent Engineering and Systems*, vol. 4, no. 1, 2011.

[10] R. Vijayanandh and G. Balakrishnan, "Human face detection using color spaces and region property measures," in *Proc. 11th Int. Conf. Control, Automation, Robotics and Vision*, Singapore, Dec. 2010.

[11] C. N. R. Kumar and A. Bindhu, "An efficient skin illumination compensation model for efficient face detection", in *Proc. 32nd Annual Conference on IEEE Indstrial Electronics*, 2006, pp. 3444-3449.

[12] Y. H. Yi, D. K. Qu, and F. Xu, "Face detection method based on skin color segmentation and facial component localization," in *Proc. 2nd International Asia Conference on Informatics in Control, Automation and Robotics*, 2010.

Mr. Prashanth Kumar G. is working as an Asst. Professor in the Dept. Electronics and communication Engineering, Channabasaveshwara Institute of Technology, Gubbi, Tumkur, Karnataka, India. He obtained his B.E (Electronics and Communication) and M.Tech. (Signal Processing) from Siddaganga Institute of Technology, Tumkur, Karnataka, under VTU Belgaum, Karnataka, India. He has 5 years of teaching experience. He has published many papers in various international Journals and conferences. His Areas of interest include Image and video processing, Robotics, Artificial Intelligence and neural networks. Presently He is working on "Implementation and design of driverless vehicles".

Mr. Shashidara M. is working as an Asst. Professor in the Dept. Electronics and communication Engineering, Channabasaveshwara Institute of Technology, Gubbi, Tumkur, Karnataka, India. He obtained his B.E (Electronics and Communication) from Siddhartha Institute of Technology, Tumkur and M.Tech. (VLSI and Embedded Systems) from Dayanandsagar College of Engineering, Bangalore, Karnataka, under VTU Belgaum, Karnataka, India. He has 4.5 years of teaching experience. He has published many papers in various international Journals and conferences. His Areas of interest include Low power VLSI, Nanotechnology, Image processing.