**NOVATEUR PUBLICATIONS**
**International Journal of Research Publications in Engineering and Technology [IJRPET]**
**ISSN: 2454-7875**
**VOLUME 3, ISSUE 6, Jun.-2017**

# A REVIEW ON: MULTIPLE FACE DETECTION AND TRACKING USING SKIN TONE SEGMENTATION TECHNIQUE

MISS KRASHNA V. PANPALIYA
Electronics and Telecommunication Raisoni College of Engineering Amravati, India, krish25panpaliya@gmail.com

PROF. SAGAR S. BADNERKAR
Electronics and Telecommunication Raisoni College of Engineering Amravati, India

**ABSTRACT:**

**Recent years have witnessed renewed interest in developing skin segmentation approaches. Skin feature segmentation has been widely employed in different aspects of computer vision applications including face detection and tracking systems. This is mostly due to the attractive characteristics of skin color and its effectiveness to object segmentation. On the contrary, there are certain challenges in using human skin color as a feature to segment dynamic face features, due to various illumination conditions, complicated environment, and computation time or real-time method. These challenges have led to the insufficiency of many of the skin color segmentation approaches. Therefore, to produce simple, effective, and cost efficient skin segmentation, a localized approach for multiple face detection and tracking in input video are proposed which is based on skin tone segmentation algorithm and dynamic facial features (eyes, nose, and mouth) extraction method which increase accuracy of detection and tracking multiple faces in video. In this proposed work, segmentation of skin regions from a video frames is done by using RGB color model which help to remove non skin like pixel from a video frames. Each segmented skin regions are tested to know whether region contain human face or not, by extracting facial features. Once the face is detected then they are tracked in video. The experimental results showed that the proposed scheme achieved good performance in terms of accuracy and computation time.**

**INDEX TERMS: Face detection, Face tracking, facial features extraction (eyes, nose, mouth), skin color segmentation**

## INTRODUCTION:

In recent years there has been a growing interest in the problem of skin segmentation, which aims to detect human skin regions in a video. The human face is central to our identity. It plays an essential role in everyday interaction, communication and other routine activities. Detection of the human face is an essential step in computer vision and many biometric applications such as automatic face recognition, video surveillance, human computer communication and large scale face image retrieval system. The first and foremost important step in any of this system is the accurate detection of the presence human faces in an image or video.

## FACE DETECTION AND TRACKING

Multiple face detection and tracking are the attraction of many research groups because of its application in many fields as a security system and surveillance, as face tagging, content-based image and video indexing, human – computer interface and many others. Face detection is the main step for facial analysis algorithms i.e. face features extraction whose motto is to detect wheather or not faces are present in an image. We have studied a multiple algorithm approaches for face detection, which is in effect a series of simple rejection blocks. In designing the final algorithm many different methods have been tried. The first step is skin segmentation, which has the best result to reject most of the data. For reducing the complication of data, we use more specific rejection classifiers. Fisher Linear Discriminate and Template matching are used for eigenface method. So in the final version, we used eigenface projection method.

Skin tone segmentation is commonly used in algorithm for face detection. It is an often used cue in human motion tracking, especially in face tracking. Skin color detection is orientation invariant and fast to process. It is very effective because it usually involves a small amount of computation and can be done regardless of pose. Color is known to be a useful cue to extract skin regions. This allows easy face localization of potential facial regions without any consideration of its texture and geometrical properties. Most technique up to date is pixel based skin detection method which determines each pixel as skin or non-skin. Skin tone segmentation is representing the RGB image into a new 3-D transformed space such that the various skin colors lie close to each other and this space is a small, constrained space. Various colors spaces provide us various discriminately between skin pixels and non skin pixels over various illumination conditions.

An image can be presented in a number of different color space models such as RGB, HSV, YCrCb. These are some, but certainly not all of the color space models available in image processing. Therefore, it is important to

choose the appropriate color space for modeling human skin color. The factors that need to be considered are application and effectiveness. The intended purpose of the face segmentation will usually determine which color space to use; at the same time, it is essential that an effective and robust skin color model can be derived from the given color space. For instance, in this paper, we propose the use of the RGB color space, and the reason is twofold. First, an effective use of the color information for modeling human skin color can be achieved in this color space. Second, this format is typically used in video coding, and therefore the use of the same, instead of another, the format for segmentation will avoid the extra computation required in conversion. The RGB color model is given in below fig. 1

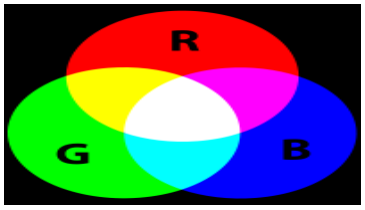


Fig. 1 RGB color model

The RGB color model is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue. The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors. The skin color region segmentation by RGB color space was performed by them logical OR operation.

**OBJECTIVE:**

In this thesis our aim is to improve the performance of multiple face detection and tracking for input video. The detection and tracking of multiple faces proceed simultaneously with image segmentation in a competitive and cooperative manner. Problem occurred in variety of dynamic environments, it has a strong adaptability, but it is generally difficult to obtain complete outline of moving object, responsible to appear the empty phenomenon, as a result the detection of moving face is not accurate. So need to improve the technique.

To achieve this, the following specific objectives are

- To study the algorithm about face detection and tracking.
- To study regarding implementation of skin tone segmentation algorithm and facial features extraction.
- To propose face detection techniques that to effectively identify face of interest area in video sequence.
- To propose enhanced face tracking techniques based on dynamic face features extraction.
- To analyze the performance parameters of system
- To detect and track multiple human faces in input video.

**REVIEW OF LITERATURE:**

There are various approaches proposed by various researchers for face detection and tracking. In this paper a brief description of these approaches are discuss.

L. Silva K. Aizawa and M. Hatori (1995) proposed a method, which they called edge pixel counting, to detect and track facial features in video sequences. This method is based on the fact that a higher edge concentration is observed in the vicinity of facial features, while slightly outside of such features a less edge concentration is observed. In spite of its simplicity, the method fails in the presence of cluttered backgrounds, glasses, hat, etc.[1]

Jie Yang (1996), presents a real-time face tracker. It can track a person's face while the person moves freely (e.g. walks, jumps, sits down and stand up) in a room. They employ three types of model for developing the system. First, they present a stochastic model to characterize skin color distributions of human faces which are use for tracking a human face in various poses and views. Second, a motion model is used to estimate image motion and to predict the search window. Third, a camera model (canon VC-CL camera) is used to predict and compensate for camera motion. The principle in developing this system can be extended to other tracking problems such as tracking the human hand [2]

R. Fraud, 0. J. Bernier (2001) proposed a fast and accurate face detector based on neural networks model: the constrained generative model (CGM). To detect the side view faces and to decrease the number of false alarms, a conditional mixture of networks is used but this system has more processing time and less accuracy.[3]

Other face detection algorithms based on eyes and nose detection were proposed by Zhu et al. (2002) and Verma et al. (2003). Zhu et al proposed a method in which a geometric function is used to detect pupils. The candidate pupils are trained by a Support Vector Machine (SVM) classifier but the detection method is not satisfactory for closed eyes.[4]

Verma et al In proposed a method in which the camera used for eyes and nose detection can be mounted to any place with different environmental conditions. The detection is based on two detectors which are frontal and profile view detectors. Also, the method can detect multiple faces as well as change scales and poses.[5]

AdaBoost is another well-known and widely used algorithm for face detection proposed by Laptev and Ivan (2006). The main idea behind AdaBoost algorithm is to obtain a powerful classifier using a large number of weak classifiers. This method can be used for the detection of different types of objects such as motorbikes, bicycles, people, and cars. The two main drawbacks of AdaBoost are the requirement of large databases for object and non-object samples and the detection method is not satisfactorily powerful for the similar objects. [6] Binu Muraleedharan Nair and Jacob Foytika (2011), proposed a system which combined the multi-pose face detection, face recognition, and face tracking. The system was tuned to work at real-time and have good accuracy in recognizing and tracking faces. But, some issues such as the lighting variations and small background inclusion into face recognition affect the performance. Moreover, the only three poses are included in the system, i.e. the frontal, right and left. They should include a pose estimation module which will help them to determine the pose and then select a suitable eigenspace rather than using the cascades. [7]

Anima Majumder, L. Behera, Venkatesh K. Subramanian (2011) presented a method for automatic detection of facial features using viola-jones object detector along with haar like cascaded features are used to detect eyes, nose and face. An algorithm using the H-plane of the HSV color space is proposed for detecting eye pupil from the eye detected region. In this paper they only tested two different facial expressions i.e. neutral face and smiling face.[8]

Prashanth Kumar G. and Shashidhara M (2014), present 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. They 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 the region is a human face or not, by using human face features based on knowledge of geometrical properties of a human face**.** But this system gives better result in good lighting conditions. Under poor or low lighting condition performance of the proposed system is satisfactory. [9]

Manjunath Hire math, P. S. Hiremath (2015) proposed a novel method for the detection and tracking of a face in a video. The detection of the human face is achieved by using fuzzy geometric face model and face tracking is done by using a sparse representation based classification algorithm. In this paper, they considered single frontal face in the video frames with different motions, head tilts, lighting conditions, expressions, and backgrounds. The proposed approach yields better average detection and tracking, which is robust and runs in almost real time. [10]

Tania Akter Setu & Dr. Md. Mijanur Rahman (2016) presented human faces detection within images and segmenting the face into numbered regions which are the face, mouth, eyes and nose regions respectively. For face detection, they used the Viola–Jones object detection framework. Sometimes the VJOD make a false frame of object detection. So they try to detect the problem of identification and improve the detection quality by changing the threshold value. It detects only the frontal 2D faces of the human. If we use side view of an image it cannot detect correctly. [11]

**PROPOSED WORK:**

The detection of multiple faces and tracking them in a video is a significant research area, which has many important applications. The human face detection in a video frame and its feature extraction are considered in this work. The human faces are non-rigid and have a high degree of variability in color, aliveness and pose, several features of the human face are not common to other pattern detection issues. Hence, human face detection is a more complex task in the computer vision. In addition, face variability occurs in the inter-spatial relationship among facial features, namely eyes, nose, and mouth. An effective face detection system shall detect the presence of all human faces in the input video. The proposed face detection and tracking in input video are given in fig. 2

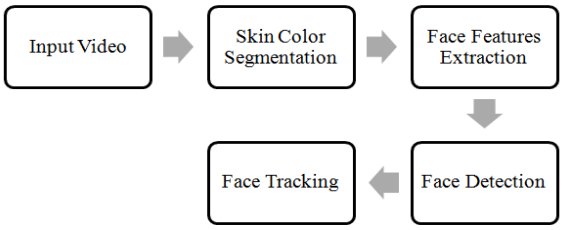


Fig. 2 Face detection and tracking

This system is describe in step wise as below

**STEP 1: INPUT VIDEO:**

The video is an electronic medium for the recording, copying, playback, broadcasting and display of moving visual media. Video systems vary in display resolution, aspect ratio, color capabilities and other qualities. There are some characteristics of video streams.

- **FRAME RATE:**

Frame rate (frames per second or FPS) is the rate at which consecutive images called frames are displayed in an animated display. Frame rate is the number of still

pictures per unit of time of video, ranges from six or eight frames per second (frame/s) Video is the combination of images or frames and audio. Video frame is one of the many still images which compose the complete moving pictures. When the moving pictures are displayed each frame is flashed on a screen for a short time (usually 1/24, 1/25, or 1/30 of a second) and then immediately replaced by the next one.

- **ASPECT RATIO:**

Aspect ratio is the dimension of video screen and video pictures elements. It is the ratio between width and height of video screen. The aspect ratio is expressed as two numbers separated by a colon (x: y). The values x and y do not represent actual widths and heights but rather the relationship between width and height. As an example, 8:5, 16:10 and 1.6:1 are three ways of representing the same aspect ratio. Pixels on computer monitors are usually square, but pixels used in digital video often have non square aspect ratios.

- **VIDEO DIMENSION:**

One of the most important visual characteristic of any video is it's width and height. The width and height of video are usually measured in pixels and are collectively termed as the "Dimensions" of the video. Thus if a video is 320 pixel wide and 240 pixel in height it is said to have dimension of 320 x 240 pixels.

- **COLOR SPACE AND BITS PER PIXEL:**

Color model name describes the video color representation. The number of different colors in an image or frames is depend on the depth of color or bits per pixel (bpp). The formula for the calculation of total number of combination that can be made from bit, it would be like this

Total No. of Color Combination = (2) bpp

Where bpp denotes bits per pixel.

There are different types of video such as AVI (Audio Video Interleave), MP4, MPEG, MPG, 3GP etc. In this work, we take any type of video as input video. Digital video is a representation of moving visual images in the form of encoded digital data. Digital video comprises a series of orthogonal bitmap digital images displayed in rapid succession at a constant rate. In the context of the video, these images are called frames. Video length depends on frame rate i.e. video is how many second was given to play.

**STEP 2: SKIN COLOR SEGMENTATION:**

Skin color is one of the most significant features of human face. Color processing is also faster and robust in nature compared to other features like edge, shape and texture etc. The skin tone segmentation is the inspiration to use skin color analysis 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. The color of human skin is different from the color of most other natural objects in the world. One important factor that should be considered while building a statistical model for color is the choice of a Color. In this paper, skin tone segmentation is done by using RGB color model.

Skin detection technique is the classification of image into skin and non –skin categories. The skin color is often used as a cue for detecting, localization and tracking targets containing skin, like faces in an image. Detecting human skin tone is of at most importance in numerous applications such as, video surveillance, face and gesture recognition, human computer interaction, human pose modeling ,image and video indexing and retrieval, image editing, vehicle drivers, drowsiness detection, controlling users, browsing behavior (e.g., surfing indecent sites) and steganography. Many applications are based on the face detection. Everyone knows very well that face color is skin color. It is very important to detect only skin parts in the human body. Methods of Detection skin is classified in two methods pixel based method and region based method. In Pixel-Based Methods, each pixel is classified individually as skin or non-skin, independently from its neighbors in contrast. Region Based Methods are try to take the spatial arrangement of skin pixels into account during the detection stage to enhance the methods performance. In this thesis, we used the pixel based method. An image is a collection of pixels and each pixel has some value. This value of the pixel is set according to color. By using this concept, classify the image in skin and non-skin pixels. The non- skin pixels of the RGB image is in the range of R>95 & G >40 & B>20 except this range, all pixels are skin pixels. we take a frame and find skin color using skin tone segmentation technique. It is a technique of discrimination between the skin and non-skin pixels of an image. Skin segmentation aims to remove all non-skin pixels in order to acquire good results in skin classification. In this article, a skin segmentation scheme based on RGB (Red, Green, Blue) pixels' color is developed. RGB color space is the most commonly used color space in digital images. It encodes colors as an additive combination of three primary colors: red(R), green (G) and blue (B). RGB color-space is often visualized as a 3D cube where R, G, and B are the three perpendicular axes shown in below fig. 3

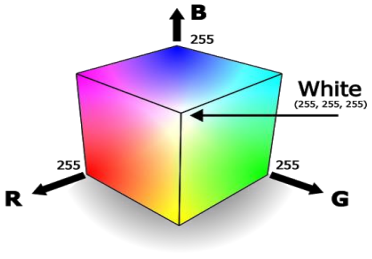


Fig. 3 RGB color space

One main advantage of the RGB space is its simplicity. In other words, the location of a given skin patch in the RGB color cube will change based on the intensity of the illumination under which such patch was imaged. This results in a much-stretched skin color cluster in the RGB color cube. If skin color present then algorithm will proceed for face detection and tracking using face features extraction.

**STEP 3:**

In this step face features are located in skin segmented area. We locate eyes, nose, and mouth. If we locate these features then we can find face easily. This is the most important stage for face detection and tracking in video sequence. In this step a novel technique using the basic concept of facial geometry are proposed to locate the eyes position, nose position and mouth position. Once the faces are detected then using template matching method we track the faces and then rectangle box will be drawn on them.

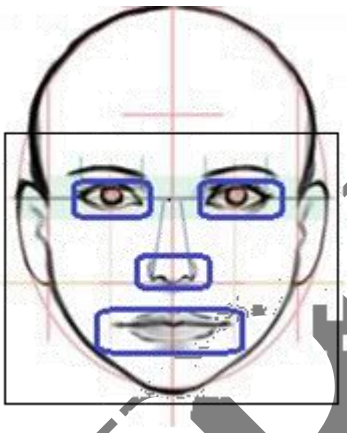

Fig 4: Detect Face

**STEP 5: FACE TRACKING**

Face tracking is important in many computer vision applications including activity, recognition, automotive safety and surveillance. In many face detection systems, the input is a video sequence consisting of one or more faces. It is necessary to track each face over this video sequence so as to extract the information that will be processed by the recognition system. Face tracking can be divided along different lines depending upon the method used, e.g., head tracking, feature tracking, image-based tracking, model-based tracking. In this thesis, face tracking is depend on face features detection. Once a face is detected, then by tracking facial features i.e. eyes, nose and mouth, the faces are tracked in the video sequence The major challenges encountered by face tracking systems are robustness to pose changes, lighting variations, and facial deformations due to changes of expression, occlusions of the face to be tracked and clutter in the scene that makes it difficult to distinguish the face from the other objects.

**CONCLUSION:**

In this paper, a new Multiple Face Detection and Tracking system in video was proposed using skin tone segmentation algorithm and facial features extraction method. By utilizing beneficial features of skin tone segmentation algorithm using RGB color model has been implemented for detection of face. The facial features extraction method was used for detection and tracking of human face. On comparing with previous work, the project design found more accuracy for multiple human face detection and tracking in input video.

**REFERENCES**

1) L. Silva, K. Aizawa and M. Hatori, "*Detection and Tracking of Facial Features"*, Proc. of SPIE Visual Communications and Image Processing, Taiwan, May 1995
2) Jie Yang and Alex Waibel, "*A real-time face tracker*", 3rd IEEE workshop on Application of computer vision, proceedings, 1996.
3) R. Feraud, 0. J. Bernier, J. E. Viallet, and M. Collobert, "*A fast and accurate face detector based on neural networks,*" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 1, pp. 42-53, January 2001
4) Zhu Z., Ji, Q., Fujimura, K., & Lee, K., "*Combining Kalman filtering and mean shift for real time eye tracking under active IR illumination*", In Pattern Recognition, 16th International Conference, Vol. 4, pp. 318-321, 2002, IEEE.
5) Verma Ragini Choudhury, Cordelia Schmid and Krystian Mikolajczyk, "*Face detection and tracking in a video by propagating detection probabilities."* Pattern Analysis and Machine Intelligence, IEEE Transactions on 25.10 (2003): 1215-1228.
6) Laptev, Ivan. "*Improvement of Object Detection Using Boosted Histograms*." BMVC. Vol. 6. 2006
7) Binu Muraleedharan Nair and Jacob Foytika, "*Multi – pose face recognition and tracking system",* Procedia Computer Science 6, 381–386, 2011
8) Anima Majumder, L. Behera, Venkatesh K. Subramanian, "*Automatic and Robust Detection of Facial Features in Frontal Face Images",* Computer Modelling and Simulation (UKSim), 2011 UKSim 13th International Conference on 30 march- 1 april 2011. IEEE.
9) Prashanth Kumar G. and Shashidhara M, "*Real Time Detection and Tracking of Human Face Using Skin Color Segmentation and Region Properties*", International Journal of Signal Processing Systems Vol. 2, No. 2, December 2014.
10) Manjunath Hiremath, P. S. Hiremath, "*Enhanced Face Detection and Tracking In Video Sequence Using Fuzzy*

*Face Model and Sparse Representation Technique"*, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS),Volume 4, Issue 6, ISSN 2278-6856 ,November - December 2015

11) Tania Akter Setu & Dr. Md. Mijanur Rahman, "*Human Face Detection and Segmentation of Facial Feature Region",* Global Journal of Computer Science and Technology: G Interdisciplinary Volume 16 Issue 1 ISSN: 0975-4172 & Print ISSN: 0975-4350, 2016.