Face Tracker for Head Position Detection

Swati P. Kale
Department of Eectronics Engineering
Shree Shankarprasad Agnihotri College of Engineering,
Wardha, Maharashtra, India.

Deepak Dandekar
Department of Electronics Engineering
Bapurao Deshmukh College of Engineering
Wardha, Maharashtra, India.

ABSTRACT
The driver fatigue detection is one of the most prospective commercial applications of facial expression recognition technology. Current facial features tracking techniques faces three challenges: 1) variety of light conditions and head orientation failure of some or all the facial features, 2) multiple and non rigid object tracking, and 3) facial feature occlusion. In this paper, we propose a new approach. First, the single camera (webcam) is used to detect face under various lighting conditions. The detected face is used to track facial features by using color model. Because color processing is very fast that mean time requirement is less. And from tracked facial features we predict the head motions in up-down and left-right direction. Furthermore, face movement are assumed to be smooth so that a facial features can be tracked with three point algorithm. Simultaneous use of YCbCr color mode, three point algorithms and the Geometric model greatly increases the prediction accuracy for each feature position. The experimental results shows validity of our approach to a real life facial tracking under various light condition, head orientations and facial expression.

General Terms
YCbcr color model, three point algorithm

Keywords
Face detection, Facial feature tracking, color model, geometric model.

1. INTRODUCTION
The trucking industry, highway safety advocates, and transportation researchers have all identified driver drowsiness as a high priority commercial vehicle safety issue. Drowsiness affects mental alertness, decreasing an individual's ability to operate a vehicle safely and increasing the risk of human error. Successfully addressing the issue of driver drowsiness is a multi-faced challenge. Operational factors such as work schedules, duty times, rest periods, recovery opportunities, and response to customer needs can vary widely. In addition, the interaction of the principal physiological factors that underlie the formation of sleepiness, namely the homeostatic drive for sleep is complex. Thus researches on detecting driver's state effectively and preventing accidents are of active meaning.

There are two methods to address fatigue of driver. One is to detect driver's behavior from vehicle side. It is nothing but the monitoring the steering, velocity of vehicle, latitudinal acceleration of vehicle etc. The other is to detect physiological signal and physical side of the driver. Physiological signal includes heart rate, body temperature brain wave etc. And physical analysis includes tracking of eye, yawing status, head orientation etc.

2. PROBLEM DEFINATION
From above method addressing of a driver from vehicle side and detection from physiological signal is expensive method as compared to detection of driver from physical side. And also complexity of the system is increases. Therefore, we address driver behavior from physical side by using a single camera.

In-vehicle technological i.e. addressing a driver from physical side approaches has effective tools to address fatigue. Sleepy drivers exhibit certain observable behaviors, including eye gaze, eyelid movement, pupil movement, head movement, and facial expression. The focus of this paper is on the last category of alertness monitoring technologies. This system continuously runs a face finding algorithm which ensures automatic subject calibration and re-acquisition. Then automatically generates a face model that takes into account unique facial features. This process takes less than a four second. Once this model is created, the system begins 3D head pose tracking and attention. Head pose is tracked to +/- 90 degrees of rotation. The system then finds status of the driver, monitoring the frequency head position and alarm system. All data is then output for prediction purposes analysis.

3. HEAD TRACKING SYSTEM

3.1 Problem Definition
1) Localization of face by using camera which is placed in front of the driver.
2) Detection of face by using color model because color processing is faster.
3) Tracking of facial features by using Hough Transform.
4) Develop Geometric face model by using tracked facial features.
5) Estimation of motion by using three point algorithms.
6) Addressing driver's status from output data analysis.

3.2 System Design
First video stream taken as an input. By applying color model driver's face is detected [1]. Then skin model images are binarized i.e. converted into 0 and 1 form. Afterwards the morphological processing and region growing operation are done on binary image to obtain coordinate data of every region. In this way face region picked out from driver's face. Then reduce face detection region to improve speed of the active facial tracking system. The implemented system then generates a face model that takes into account unique facial features. This process takes less than a second. Once face model is created, the implemented system begins head pose

tracking, providing three point algorithms on subject of facial features. The system design described below in figure 1.

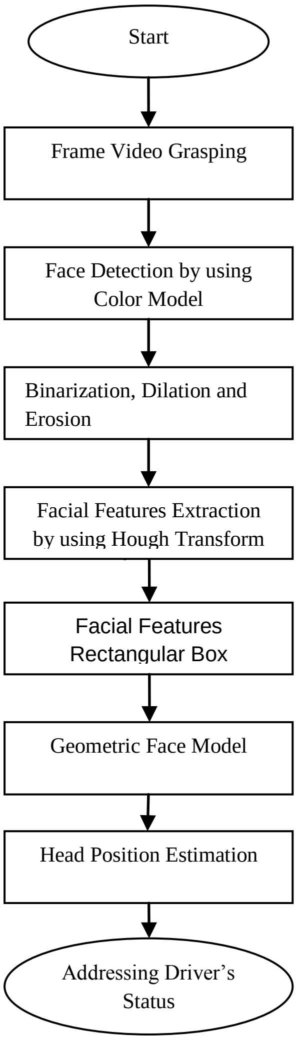

Fig 1: System Flow

3.3 Face Detection

Color is an important feature of human faces. A color model [1] is a specification of a 3-D coordinate system. Much of the research on skin color based face localization and detection is based on RGB, YCbCr and HSV [3] etc. color spaces model. Using skin-color as a feature for tracking a driver's face has several advantages. The second step to extract facial features [3] like eyes, nose tip etc. We tested the proposed algorithm over 1000 frames, which were comprised of 20 facial images. We obtained a 90 %success rate in the detection of eyes and nose as a facial feature. The purpose of a color model is to facilitate the specification of colors in some standard generally accepted way.

3.4 Ycbcr Model

YCbCr is a digital color system. These colors spaces separate RGB (Red-Green-Blue) into luminance and chrominance information. In this format, luminance information is stored as a single component (Y), and chrominance information is stored as two color-difference components (Cb and Cr). Cb represents the difference between the blue component and a reference value. Cr represents the difference between the red component and a reference value. YCbCr data can be double precision, but the color space is particularly well suited to uint8 data. For uint8 images, the data range for Y is [16, 235], and the range for Cb and Cr is [16, 240]. When R'G'B' are converted to Y, Cb and Cr then applying range on the input data. Due to this non skin pixels image is removed which is shown in the figure 2. The Intel IPP functions use the following basic equations to convert between R'G'B' in the range 0-255 and YCbCr:

$$Y = 0.257*R + 0.504*G + 0.098*B+ 16$$

$$Cb = 0.148*R - 0.291*G + 0.439*B + 128$$

$$Cr=0.439*R-0.368*G-0.071*B+128 \qquad \ldots\ldots (1)$$

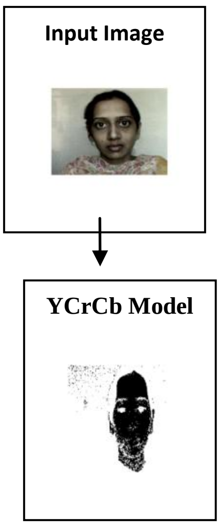

Fig 2: Color Model

3.5 Binarization

Binarization [4] is nothing but the conversion of skin model images into 0 and 1 format. The scheme diagram of the double-threshold method is shown in Fig. 5. The top-face boundary is too weak to be detected in the later thresholding procedure. The high threshold image is used to extract facial feature. These thresholding percentages are determined based on our data in order to achieve the best results. In our project the threshold value 0.99 is used for best result as shown below fig.3.Thresholding is performed to transform the output image of color model as explained above into gray-level images. And again transform into binary by using Matlab function as given below

im2bw (grayscale image, threshold value) converts the grayscale image to a binary image. The output image replaces all pixels in the input image with luminance greater than level (threshold value) with the value 1 (white) and replaces all other pixels with the value 0 (black). You specify level in the range [0, 1], regardless of the class of the input image.

imfill (Binary image,' holes') fills holes in the binary image. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

Threshold 0.40 **Threshold 0.60**

Threshold 0.90 **Threshold 0.99**

Fig 3: Binarization of image

3.6 Morphological Structuring and Features Extraction

To create morphological structuring element use morphological functions. Morphological functions are used as a tool for extracting image components. Their are two morphological operations, dilation and erosion. Dilation is used to constructs structuring elements with a variety of shapes and sizes. Erosion is used to shrinks interested object in a binary image.

seD = strel ('diamond', 1) % for structuring interested region.
imbw1 = imerode (imfl, seD) %to erode the image…. .(2)
After tracking features, the computation of the difference between the position of a facial feature [4] in one frame and its position in the next frame. These differences are used to change the wire frame face position according to the movement of the driver's face that is being tracked as shown in figure 4.

4. Geometrical Model

We apply the geometric face model [5] to locate head position. The model utilizes the configuration among eyes and nose tip [10] which can be extracted by using the circular Hough Transform [9] as shown in figure 5. Let the distance between the two centers of eyes be D. The geometric face model and related distances are described below by using three point algorithms. The equation of a line a x + b y + c = 0 in slope-intercept form is given by y = -a/b x – c/b. so the line has slope –a/b .Now consider the distance from a point (xo, yo) to the line. The line is specified by two points x1 = (x1, y1) and x2 = (x2, y2).

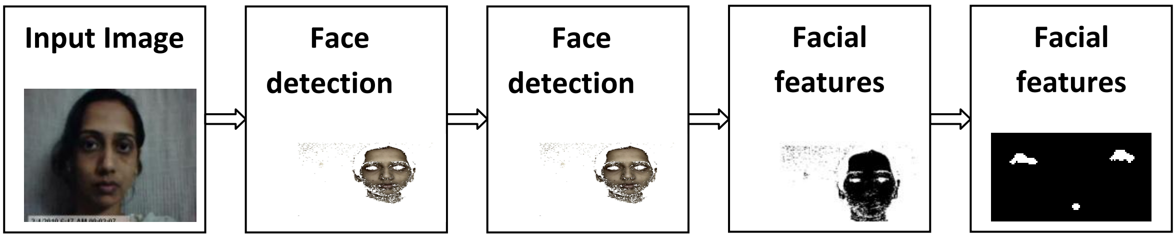

Fig 4: Face Detection System

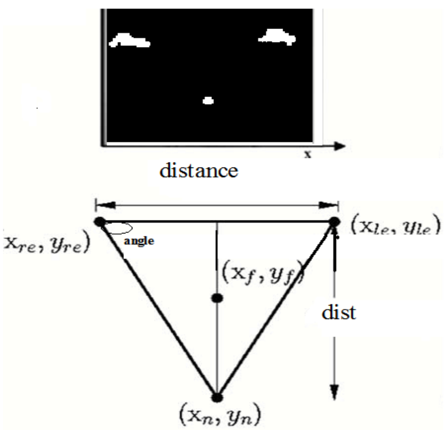

Fig 5: Three Point Model Design

As we show in Figure 5, with 2D eyes and nostril location and estimating the distance to head rotation point [6].We can recover 3D face direction using basic projections. If the calculated distance d is greater than 10 and less then or equal to 39 then driver is in 'AWAKE' condition. If the calculated distance d is greater than 39 and less then or equal to 42 then driver is in 'DROWSY' condition. And otherwise driver is in 'SLEEP' condition. Above measurement related to the distance d is calculated by observing more than 20 videos. Along with we can calculate angle of head rotation.

```
if dist > 10 &  dist <= 39
output ('Normal');
else if dist > 39 &  dist <= 42 || dist >= 50 & dist<=52
 output('Drowsy');
elsedist > 42
 Output('Sleeping mode');
end
```

Algorithm 1: Vertical Distance Measurement

```
if   distance > 41 & distance <= 56   || distance >= 73 &
distance < 75
Output (' down movement');
elseif distance > 76
Output (left-right movement');
elsedistance > 57 & distance < 71
Output (In front of camera');
```

Algorithm 2: Horizontal Distance Measurement

We are extracting centers of eyes by using Hough Transform [6] and with the help of these points calculate the distance. The horizontal distance between two eyes can be calculated by subtracting the both point coordinates and then squaring it. The head movement is calculated from following algorithm 2.

4.1 Experimental Result

The combined algorithms of measurement [9] of 'd', distance and angle measurement is used in Geometric Model [6] which is explained above. The output of the head position detection is given below. The correct rate of tracking is given by following formulae.

Correct Rate of Tracking = (Total no. of frames – Tracking frame failure)/ Total frame ……. (2)

Fig 6: Output of the system in medium Illumination

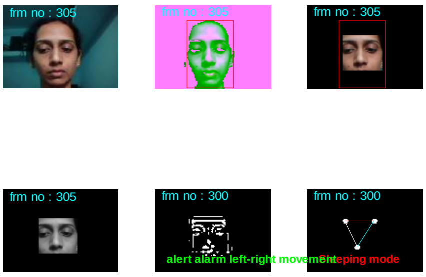

Fig 7: Output of the system in dark Illumination

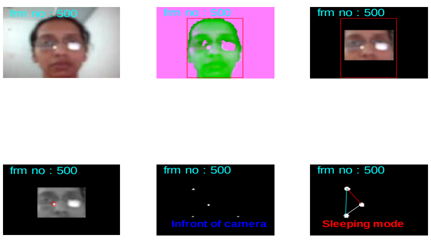

Fig 8: Output of the system with glasses

5. COMPARISON AND DISCUSSION

Implementation of a driver vigilance system to track facial features and estimate head pose is hard requirements on the algorithms utilized. To accomplish this demanding task, many different algorithms are applied. In this project, promising methods were searched from the literature. Based on the survey the implementation for a driver vigilance system is proposed. Skin detection has been a popular method for face region detection. It has been noticed to be highly suitable due to fast processing time and simple implementation of the algorithm. A shortcoming is that changing illumination conditions decrease the performance of the skin detection system. Because the initialization stage of the driver vigilance system depends heavily on skin detection, dramatically changing or bad illumination can cause facial features to be undetectable. Also if the skin areas are not solid and proper feature extraction is impossible using this method. Skin detection can be improved, for example, by using information from the previous frames. In the current implementation, skin detection is performed for every image individually. However, skin detection provides a good starting point for facial feature extraction method.

Morphological image processing techniques operations [9] by using Matlab function have proven to be efficient and powerful due to fast processing. In this study morphological methods are employed for facial feature extraction method. In the implemented face model, there are three features; eyes and nose tip. Based on the experiments, the expected initialization result for a user without glasses can be quite good. On the other hand, the result for a user with glasses can be quite poor as shown in table 1.

6. CONCLUSION

In future, the features to be extracted and tracked could be the corners of the eyes and the nose tip instead of the centers of the eyes and the nose tip. The corners are more permanent than the centers of these features due to possible deformations during tracking. One important issue is to choose an appropriate model for the head in the tracking method. In these thesis three facial features, the eyes and the nose tip, roughly model the head orientation. In the future, more accurate 3-D models could be applied that make it possible to achieve better head pose estimates. Furthermore, a general model for the head that applies to all the humans is almost impossible to generate. A difficult problem is how to select good features for tracking system. For a large number of feature points, it is crucial to perform pre-selection, for example, by eliminating all the points that are occluded with respect to the camera. There are some automatic pre algorithms for finding out optimal feature points, but it is nevertheless a very difficult task. The disadvantage of those selection algorithms is usually the computational complexity. The system was basically tested with both real image sequences and synthetic data. Although there are some problems, very good preliminary results were achieved. At this point of implementation, the conducted testing was enough to prove the correctness of the approach.

7. REFERENCES

[1] M. H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. **24**, no. 1,pp. 34–58, January 2002.

[2] B. Martinkauppi, M. Soriano, and M. Pietik¨ainen,"Detection of skin color under changing illumination:A comparative study," in Proceedings of the 12th International Conference on Image Analysis and Processing, September 2003, pp. 652–657, Mantova, Italy.

[3] K. W. Wong, K. M. Lam, and W. C. Siu, "A robust scheme for live detection of human faces in color images," Signal Processing: Image Communication, vol.18, no. 2, pp. 103–114, February 2003.

[4] Sandeep, K and A.N Rajagopalan. (2002)."Human Face Detection in Cluttered Color Images Using Skin Color and Edge Information". ICVGIP.

[5] S.-H. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, "Facial feature detection using geometrical face model: An efficient approach," Pattern Recognition system.

[6] Jari Hannuksela, Janne Heikkil¨a and Matti Pietik¨ainen" A Real Time Facial Feature Based Head Tracking" Machine Vision Group , Infotech Oulu , University of Oulu , Finland , Aug. 31-Sept. 3, 2004.

[7] Y. T. Liu, "Facial feature detection using geometrical face model: An efficient approach," Pattern Recognition, vol. 31, no. 3, pp. 273–282, March 1998.

[8] Shunji Katahara and Masayoshi Aoki" Motion Estimation of Driver's Head from Nostrils Detection" Faculty of Engineering, SEIKEI University, 23--25 January 2002.

[9] Zhiwei Zhu , Qiang Ji "Real Time and Non-intrusive Driver Fatigue Monitoring" Department of Electrical, Computer, and Systems Engineering Rensselaer Polytechnic Institute ,Troy, New York, USA , July,2004

[10] D. Dervinis "Head Orientation Estimation using Characteristic Points of Face"Department of Electronics, 2006

Table 1 : Comparison of Experimental Results

Sr. No.	Video Name	Total no. of frames	Illumination / with or without glasses	Correct tracked frames	Failure frames	Correct rate of tracking in percentage	Frame rate (fps)
1)	M2	97	Normal illumination and without glasses	96	1	98.96%	21
2)	M4	173	Normal illumination and without glasses	171	2	99.42%	21
3)	S22	500	Normal illumination and with glasses	475	35	92%	29
4)	S3	306	Dark illumination and with glasses	273	33	89.21%	30

*Frame rate set to15 fps.