**Chapter 1**

**INTRODUCTION**

Each year hundreds of people lose their lives due to traffic accidents around the world. Unfortunately Iran ranks first in the world in terms of road fatalities and each year approximately thirty thousands of fellow countrymen lose their lives in these events. The role of human factor in accidents cannot be ruled out. According to national statistics in 90 to 95 percent of car accidents in Iran, human factor plays a pivotal role. In general, the driver fatigue accounts for 25 percent of accidents and approximately 60 percent of road accidents result in death or serious injury. In a study by the National Transportation Research Institute (NTSRB) in which 107 random car accidents had been selected, fatigue accounted for 58% of the all accidents. A main cause of fatigue is sleeplessness or insomnia. Drivers' drowsiness is a major contributing factor in severe road accidents that claims thousands of lives every year. According to accident statistics presented by Oklahoma Transportation Institute, which showed 22 percent of all accidents were due to driver’s drowsiness and fatigue, Bittner at al. (2000) proposed an electronic system to alarm drivers.

In recent years, the use of intelligent systems in cars has developed considerably. These systems use wireless sensor networks to monitor and transmit the condition of the car and the driver. Smart cars which use software techniques to control engine speed, steering, transmission, brake etc. has improved the quality of driving drastically. Ad hoc networks were the first systems to develop the automatic navigation in cars. A noticeable weakness of these systems is that their response to environmental changes is not real time. It is especially important in driving where time is a critical factor in driver's decision. On the other hand, another method to check the driver fatigue is monitoring the physical condition and facial expressions of the drivers, which wireless sensor networks are unable to process and transmit these information with adequate precision. Driver fatigue is a significant factor in a large number of vehicle accidents. Recent statistics estimate that annually 1,200 deaths and 76,000 injuries can be attributed to fatigue related

Crashes.

**1.1 Existing system**

In the past few years, many researchers have been working on the development of safety systems using the different techniques. The most accurate techniques are based on physiological measures like brain waves, heart rate, pulse rate, respiration, etc. However, these techniques are intrusive since they require electrodes to be attached to the drivers, causing annoyance to them. A representative project in this line is the MIT Smart Car where several sensors (electrocardiogram, electromyogram, respiration, and skin conductance) are embedded in a car and visual information for sensor confirmation are used. In the advanced safety vehicle (ASV) project conducted by Toyota the driver must wear a wristband in order to measure his heart rate. Others techniques monitor eyes and gaze movements using a helmet or special contact lenses. These techniques, though less intrusive, are still not acceptable in practice. In existing system it includes driver fatigue detection procedure consists of different phases to properly analyze the following

1. Face Detection

2. Eye Detection

3. Yawning Detection

**1.2 Proposed system**

The proposed algorithm conducts the detection process by recording the video sequence of the drivers and image processing techniques. The system consists of four well-defined phases, namely the face detection, eye tracking, yawning detection and detection of head lowering.

The sequences of images from the camera are fed to the system. Initially, the system doesn’t know the initial position of the face. The system grabs the first image and tries to find the face region in the image using the skin color model. Due to unfavorable lighting conditions or initial head orientation of the driver, the localization might fail. So the system grabs another frame and repeats the same process until the face region is detected with certainty.

**1.3 Objectives**

This paper first will segment the face of driver using YCbCr color space and then detect the sign of driver’s fatigue which are head lowering, eye detection, yawning detection using HSV color space Sobel edge detection algorithm and K-means algorithm respectively.

**1.4 Organization of the report**

The rest of the report is organized as follow:

* Chapter 2- deals with the literature survey and the technologies used that are part of this project.
* Chapter 3- deals with the system specification which is a part of this project.
* Chapter 4- gives the complete system design with the help of the system architecture, data flow diagram, use case diagram, sequence diagram and class diagram.
* Chapter 5- deals with the system implementation
* Chapter 6- deals with the system testing, user interface testing and various test cases.
* Chapter 7- deals with the result analysis, final output.
* Chapter 8- this includes the conclusion and future enhancement part of the project. It briefly tells what the system is doing and what has to be accomplished.

**Chapter 2**

**LITERATURE SURVEY**

Literature survey is the most important step in software development process. Before developing the tool it is necessary to determine the time factor, economy and company strength. Once these things are satisfied, then next steps are to determine which operating system and programming language can be used for developing the tool. Once the programmer starts building the tools the programmer need lot of external support. This support can be obtained from senior programmer, from books and from website. Before building the system the above consideration are taken into account for developing the proposed system.

**2.1 Drowsiness Detection System for an Intelligent Vehicle**

The existing system implementation present’s the color-space separation was by far the most effective means of eliminating non-face regions from consideration. For the subsequent face-segmentation step, we found that the very simple method of looking for face-like shapes within the skin-probability image to be effective and computationally efficient. We did not have much success with morphological mouth. These more sophisticated approaches certainly have merit in a more general-purpose face detection program. However, with our very consistent and predictable set of training images, the simplest approach proves to be more than adequate, as is evidenced by the overall accuracy rating of 80%, and quick execution time of roughly one minute.

Although the presented skin color segmentation and face localization scheme is feasible under controlled lighting conditions, one major problem is changing and non-uniform lighting conditions. This is due to the still unsolved color constancy problem on image acquisition through video hardware. Eye localization is realized by matching a reconstructed image against the original input image. The reconstructed image is computed utilizing the eigen eye basis and therefore the reconstruction works best for eye like regions. The eye localization algorithm works robust and reliable in typical scenes. Once the face and eye positions are localized they are tracked by the algorithm. The center position and size of the facial area is tracked using a linear K-alman filter. The object motion model used for the Kalman filter assumes linear motion of the tracked object. The use of more complex motion models can further improve the quality of object tracking. Eye tracking is performed using a luminance-adapted block matching technique. In general this algorithm works robust and efficient.

**2.2 Real time detection of drowsiness of a driver**

In our proposed model we are trying to overcome the problems that we have experienced in existing system. We are using YCbCr algorithm for the segmentation of face. The main advantage of converting the image to the YCbCr domain is that influence of luminosity can be removed during our image processing. In the RGB domain, each component of the picture (red, green and blue) has a different brightness. However, in the YCbCr domain all information about the brightness is given by the Y-component, since the Cb (blue) and Cr (red) components are independent from the luminosity. The following conversions are used to segment the RGB image into Y, Cb, Cr components.

Our approach makes use of the Saturation value of a pixel to determine if the Hue or the Intensity of the pixel is more close to human perception of color that pixel represents. The K-means clustering of features combines pixels with similar color for segmentation of the image into the objects. We are also able to generate a histogram that enables us to perform a window-based smoothing of the vectors during retrieval of similar images.

1. A face detection algorithm for color images has been proposed using a skin-tone color model and facial texture features. It overcomes the difficulty of detecting the low-luma and high-luma skin tones by applying a nonlinear transform to the YCbCr color space. Proposed method detects skin regions over the entire image, and then generates face candidates based on the spatial arrangement of these skin patches, our goal is to design a system that detects faces and facial features, allows users to edit detected faces, and use these detected facial features as indices for identification and retrieval from image databases.
2. Face detection is accomplished by first performing a skin detection search of the input image based on color segmentation. Although skin colors differ from person to person, and race to race, it was found that the color remains distributed over a very small region in the chrominance plane.
3. It deals with color scale face images instead of the gray information. The gray images are avoided for two reasons: first humans easily locate faces in color images depending on skin color tone. A more pragmatic reason is that color would increase additional data source to differentiate between person to person and between the same persons in different light sources.
4. This work has demonstrated that face image preprocessing image resize is indeed a practical solution to speed up the image searching by carefully selecting/generating proper representation features. This is believed to be the first piece of work on face image preprocessing.
5. The three different color spaces RGB, HSI and YCbCr are briefly discussed. Moreover, the objectives of this paper have been accomplished as is evident from the results which show a higher accuracy of 92.69% as compared to the accuracy of all existing algorithms.

The Sobel edge detector uses a pair of 3X3 convolution masks, one estimating gradient in the *x* direction and the other estimating gradient in *y*–direction. It is easy to implement than the other operators. Transferring a 2-Dpixel array into statistically uncorrelated data set enhances the removal of redundant data, as a result, reduction of the amount of data required to represent a digital image. Considering data communication especially the internet, massive data transfer causes serious problems for interactive network users. Edge detection helps in optimizing network bandwidth and it is needed to keep track of data flowing in and out of the network. It helps to extract useful features for pattern recognition. Although the Sobel operator is slower to compute, it’s larger convolution kernel smoothes the input image to a greater extent and so makes the operator less sensitive to noise. The larger the width of the mask, the lower its sensitivity to noise and the operator also produces considerably higher output values for similar edges. Sobel operator effectively highlights noise found in real world pictures as edges though the detected edges could be thick. The Canny edge detector and similar algorithm solved these problems by first blurring the image slightly then applying an algorithm that effectively thins the edges to one-pixel. This may constitute a much slower process, hence, Sobel operator is highly recommended in massive data communication found in image data transfer. The Sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction and is therefore relatively inexpensive in terms of computations. On the other hand, the gradient approximation which it produces is relatively crude, in particular for high frequency variations in the image.

The proposed algorithm for face detection has a very good performance in detecting low quality faces and faces affected by environmental lightning conditions. However, it is sensitive to pose rotated faces and faces corrupted by other objects. Detection rates and false alarm rates on test sets are listed in Table, while the detection rate is acceptable. The implementation results were compared with other results and the performance of our proposed method is better than that of existing color image enhancement algorithms.

**Chapter 3**

**SYSTEM REQUIREMENTS AND SPECIFICATION**

**3.1 Hardware Requirements**

For running the real time application to detect driver fatigue, the hardware requirements are:

1. Processor : Intel Pentium IV Processor

2. RAM : 1 GB

3. HDD : 20 GB

4. Webcam : 15 Mega Pixels

**3.2 Software Requirements**

For running the real time application to detect driver fatigue, the software requirements are:

1. Operating System : Windows 95/XP/7

2. Programming Toolbox : MATLAB R2013a

**3.3 MATLAB**

MATLAB (matrix laboratory) is a [multi-paradigm](http://en.wikipedia.org/wiki/Multi-paradigm_programming_language) [numerical computing](http://en.wikipedia.org/wiki/Numerical_analysis) environment and [fourth-generation programming language](http://en.wikipedia.org/wiki/Fourth-generation_programming_language). Developed by [Math Works](http://en.wikipedia.org/wiki/MathWorks), MATLAB allows [matrix](http://en.wikipedia.org/wiki/Matrix_%28mathematics%29) manipulations, plotting of [functions](http://en.wikipedia.org/wiki/Function_%28mathematics%29) and data, implementation of [algorithms](http://en.wikipedia.org/wiki/Algorithm), creation of [user interfaces](http://en.wikipedia.org/wiki/User_interface), and interfacing with programs written in other languages, including [C](http://en.wikipedia.org/wiki/C_%28programming_language%29), [C++](http://en.wikipedia.org/wiki/C%2B%2B), [Java](http://en.wikipedia.org/wiki/Java_%28programming_language%29), and [Fortran](http://en.wikipedia.org/wiki/Fortran).

MATLAB is an interactive system for doing numerical computations. A numerical analyst called Cleve Moler wrote the first version of MATLAB in the 1970’s. It has since evolved into a successful commercial software package. MATLAB relieves you of a lot of the tasks associated with problems solving numerically. This allows you to spend more time thinking, and encourages you to experiment. MATLAB makes use of highly respected algorithms and hence you can be confident about your results. Powerful operations can be performed using just one or two commands. You can build up your own set of functions for a particular application. Excellent graphics facilities are available, and the pictures can be inserted into LATEX and Word document.

MATLAB is an interactive environment and programming language for numeric scientific computation. It will allow you to solve many technical computing problems, especially those with matrix and vector formulations. One of its distinguishing features is the use of matrices as the only data type.

MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problem and solutions are expressed familiar mathematical notation. Typical uses include,

1. Math and computation
2. Algorithm development
3. Data acquisition
4. Modeling, simulation, and prototyping
5. Data analysis, exploration, and visualization

**3.4 MATLAB Commands**

**1. imread()**

Read gray scale or color images from graphics file.

**2**. **imwrite ()**

Write gray scale or color images to graphics file.

**3. imshow ()**

Displays the image

**4. imaqfind ()**

It returns an array containing all the video input objects that exist in memory. If only a single video input object exists in memory, imaqfind displays a detailed summary of that object.

**5. imhist ()**

Display histogram of image data.

1. **subplot ()**

Subplot divides the current figure into rectangular panes that are numbered row wise. Each pane contains an axes object which you can manipulate using [axes Properties](jar:file:///C:/Program%20Files/MATLAB/R2009b/help/techdoc/help.jar%21/ref/axes_props.html). Subsequent plots are output to the current pane.

**3.5 System Components**

**3.5.1 Camera**

We use a video camera, which is suitable for vehicle- based applications to monitor the driver. The camera is positioned so that we can monitor the driver’s head and especially the eyes. We want to measure the movement of the head, eyes and eyelids with the aid of the camera. The camera is connected directly to the processor’s video input, and it must meet certain technical requirements including for example adequate resolution, frame rate.

**Chapter 4**

**DESIGN**

**4.1 System Architecture**

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures of the system.

System architecture can comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them. It can provide a plan from which products can be procured, and systems developed, that will work together to implement the overall system. There have been efforts to formalize languages to describe system architecture; collectively these are called architecture description languages (ADLs).

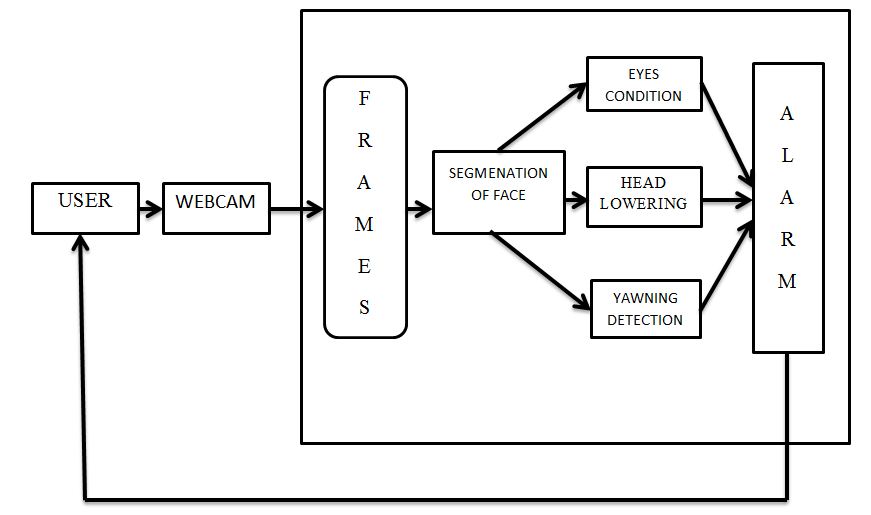
The general diagram of system has been shown in Fig.4.1. As it can be seen, the image received from camera is sent to central processor to be processed and then it will operate considering condition of drivers' face. The oxygen sensor in the car also constantly checks the density of oxygen in the air and sends alarm if there is lack of oxygen or sleep symptoms are detected in the face of the driver. 5v power supply transformer ensures working of system for long duration.

The following components are used in proposed system architecture:

Alarm: Used alert the driver when is goes to fatigue situation.

Camera: Captures the high quality image which is very much necessary for image processing.

Battery: 5v power supply is used so that the system can work for long duration.

****

**Fig 4.1 System Architecture**

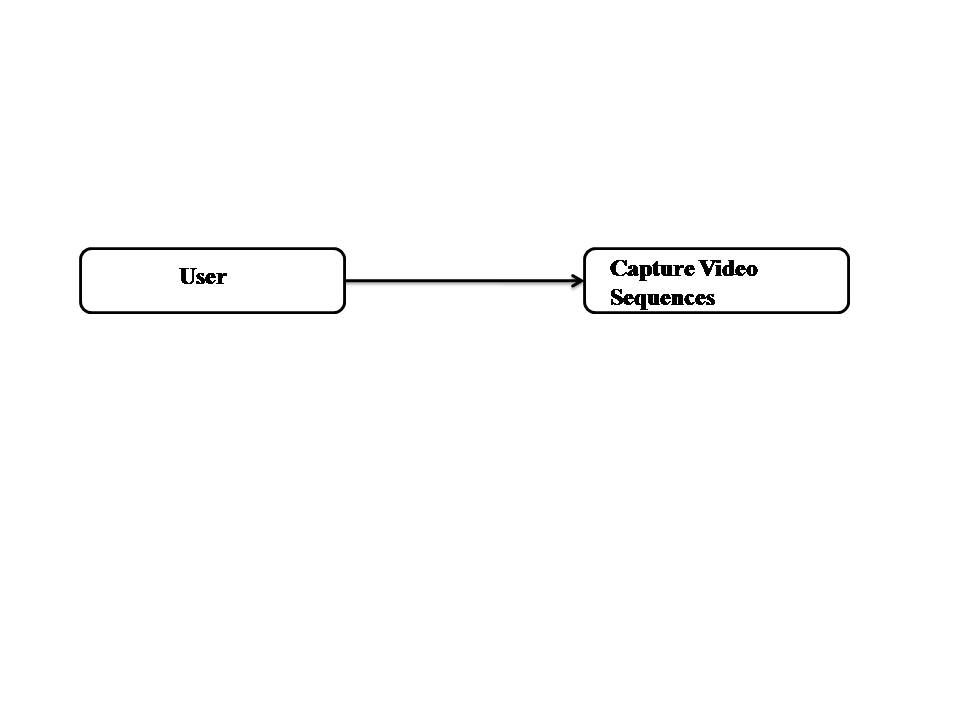
**4.2 Data Flow Diagram**

A data flow diagram (DFD) is a graphical representation of the flow of data through information system, modeling its process aspects. They can also be used for visualization of data processing.

A DFD shows what kinds of information will be input to and output from the system, where the data will come from and go to, and where the data will be stored. It does not show information about the timing of processes, or information about whether processes will operate in sequence or in parallel (which is shown on a [flowchart](http://en.wikipedia.org/wiki/Flowchart)).

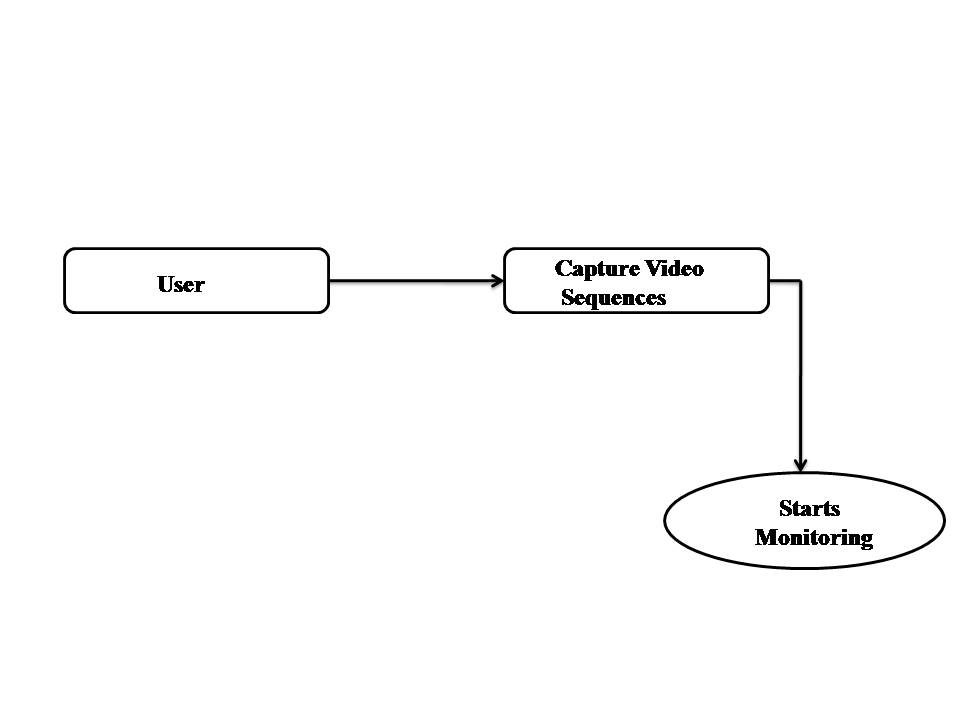
**Level 0:**

In this level, the person's face is filmed by a camera in the first step by receiving 15fps video sequence. Four different video sequences captures image frames.

****

**Fig 4.2 Capturing Video Sequences**

**Level 1:**

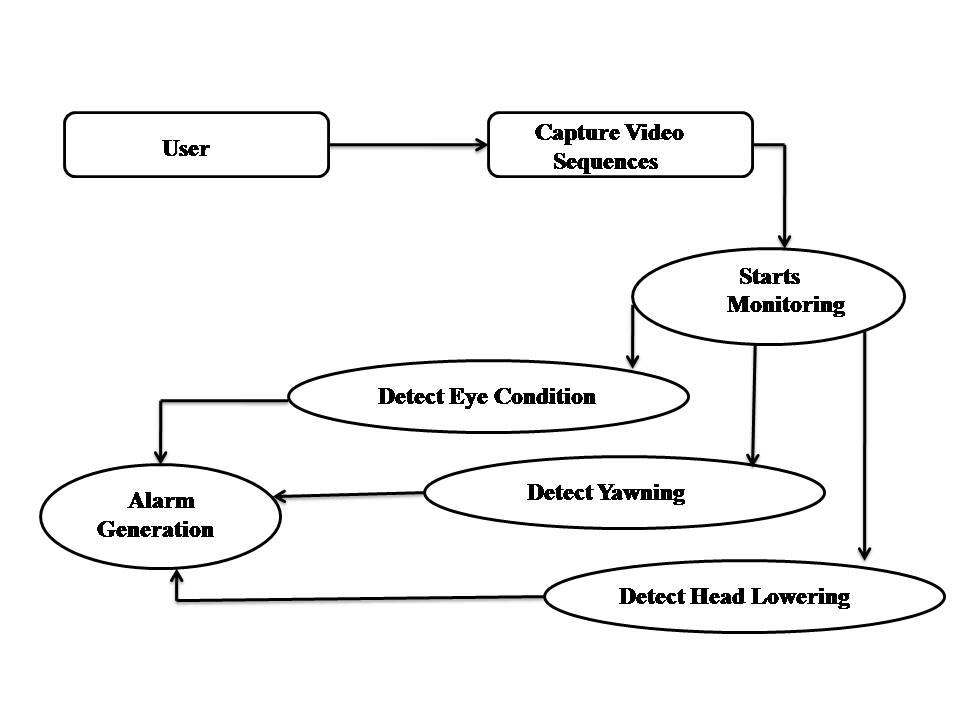
****

**Fig** **4.3 Monitoring Videos**

After the video is captured , the position of the eyes and mouth are detected. The eye movements and head movements are tracked.

**Level 2:**

Here, after the monitoring, if the eye condition of the driver is not proper or, he is yawning or his head is lowered due to drowsiness or fatigueness, an alarm is generated to alert the driver. The camera is positioned so that we can monitor the driver’s head and especially the eyes. We want to measure the movement of the head, eyes and eyelids with the aid of the camera. The camera is connected directly to the processor’s video input, and it must meet certain technical requirements including for example adequate resolution, frame rate.

****

**4.3 Sequence diagram**

A sequence diagram is an [interaction diagram](http://en.wikipedia.org/wiki/Interaction_diagram) that shows how processes operate with one another and in what order. It is a construct of a [Message Sequence Chart](http://en.wikipedia.org/wiki/Message_Sequence_Chart). A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario.

Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams, event scenarios. A sequence diagram shows, as parallel vertical lines (lifelines), different processes or objects that live simultaneously, and, as horizontal arrows, the messages exchanged between them, in the order in which they occur. This allows the specification of simple runtime scenarios in a graphical manner.



**Fig 4.4 sequence diagram for proposed system**

**4.4 Use case diagram**

Use case diagram is a representation of a user's interaction with the system and depicting the specifications of a [use case](http://en.wikipedia.org/wiki/Use_Case). A use case diagram can portray the different types of users of a system and the various ways that they interact with the system. This type of diagram is typically used in conjunction with the textual [use case](http://en.wikipedia.org/wiki/Use_Case) and will often be accompanied by other types of diagrams as well.

A use case defines the interactions between external actors and the system under consideration to accomplish a goal. Actors must be able to make decisions, but need not be human: "An actor might be a person, a company or organization, a computer program or computer system hardware, software, or both. Actors are always [stakeholders](http://en.wikipedia.org/wiki/Project_stakeholder), but not all stakeholders are actors, since they "never interact directly with the system, even though they have the right to care how the system behaves”. For example, "the owners of the system, the company's board of directors, and regulatory bodies such as the Internal Revenue Service and the Department of Insurance" could all be stakeholders but are unlikely to be actors.

In our proposed system sequence diagram there are two actors i.e. the user and system were the camera monitors users face and system records the video and creates images for segmentation of face. Segmentation of face is done in order to extract only the eye and mouth region and discard the surrounding region which we are not interested in. Then the conditions for fatigue and non fatigue are checked. If fatigue is detected then alarm is generated, if no fatigue is detected then no alarm is generated.



**Fig 4.5 Use case diagram for proposed system**

**4.5 Class diagram**

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure of a system by showing the system’s classes, their attributes, operations (or methods) and the relationships among objects

.

**Fig 4.6 class diagram for proposed system**

**Chapter 5**

**SYSTEM IMPLEMENTATION**

Implementation is the stage of the project when the theoretical design is turned into a working system. Thus it can be considered as the critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implement stage involves careful planning, investigation of the existing system and its constraints on implementation, designing of methods to achieve change over and evaluation of change over methods.

Here the proposed system includes four modules they are as follows

* Segmentation of face
* Detection of eyes condition
* Detection of head lowering
* Yawning detection

**5.1 Segmentation of face**

This is the very first module in which the face is segmented from the input image that is initially whatever the video that is recorded by the camera will be fragmented into the frames and then into the image, this image will be given as input for segmenting the face.

The partial segmentation of the image by selecting the appropriate threshold is based on dividing the image into the background and foreground classes. Thresholding is primarily concerned with selecting an appropriate threshold according to image histogram. That is, the value of thresholding or border as the brightness intensity is considered as the basis of the division and the brightness intensities greater and less than threshold is equal to 1 and zero respectively.

The purpose of face detection is to minimize the error rate in identifying facial expressions. The importance of this part is to measure the position of the eyes, the mouth and the head.

**5.1.1 Histogram**

A histogram is a graphical representation of the [distribution](http://en.wikipedia.org/wiki/Frequency_distribution) of data.

There are two types of histogram they are as follows

* Image histogram
* Color histogram

Image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. Image histograms are present on many modern digital cameras. Horizontal axis of the graph represents the tonal variations, while the vertical axis represents the number of pixels in that particular tone. In the field of computer vision image histograms can be useful tool for thresholding. This threshold value can be used for edge detection, image segmentation and co-occurrence matrix.

Representing an histogram of an image in matlab

A=imread(‘sample.jpg’);

hist(A);

Here imread() reads a grayscale or color image from the file specified by the string filename. If the file is not in the current folder, or in a folder on the MATLABpath, specify the full pathname. Basically here we are converting an image to histogram only for getting the threshold values in order to separate the foreground and background class.

|  |  |
| --- | --- |
| C:\Users\Mai\Documents\Bluetooth Folder\Capture.PNG | C:\Users\Mai\Documents\Bluetooth Folder\1.PNG |

Here the below figure shows the image in RGB color space and its histogram. Here we are finding histogram of the image in order to find the threshold value of the image.Using this threshold value we are classifying foreground and background image.

**Fig 5.1 Digital image Fig 5.2 Histogram of an image**

**5.1.2 YCbCr Color Space**

Initially the camera will record the video of a person who is driving so that all the images in that recorded video will be in the RGB color space that includes driver face along with surrounding area in the vehicle.

The purpose of face detection is to minimize the error rate in identifying facial expressions. The importance of this part is to measure the position of the eyes, the mouth and the head.

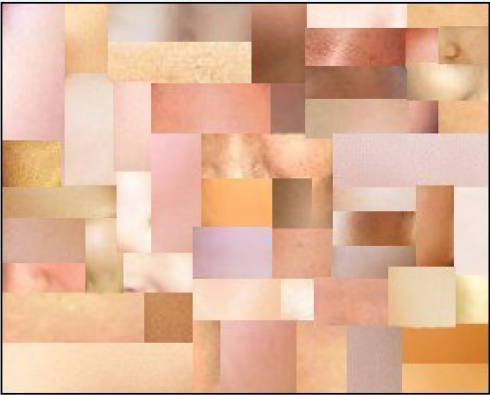
YCbCr, Y′CbCr, or Y Pb/Cb Pr/Cr, also written as YCBCR or Y′CBCR, is a family of [color space](http://en.wikipedia.org/wiki/Color_space)  used as a part of the [color image pipeline](http://en.wikipedia.org/wiki/Color_image_pipeline) in [video](http://en.wikipedia.org/wiki/Video) and [digital photography](http://en.wikipedia.org/wiki/Digital_photography) systems. Y′ is the [luma](http://en.wikipedia.org/wiki/Luma_(video)) component and CB and CR are the blue-difference and red-difference chroma components. Y′ (with prime) is distinguished from Y, which is [luminance](http://en.wikipedia.org/wiki/Luminance_(relative)), meaning that light intensity is nonlinearly encoded based on [gamma corrected](http://en.wikipedia.org/wiki/Gamma_correction) [RGB](http://en.wikipedia.org/wiki/RGB) primaries.

Y′CbCr is not an [absolute color space](http://en.wikipedia.org/wiki/Absolute_color_space); rather, it is a way of encoding [RGB](http://en.wikipedia.org/wiki/RGB) information. The actual color displayed depends on the actual [RGB](http://en.wikipedia.org/wiki/RGB) primaries used to display the signal. Therefore a value expressed as Y′CbCr is predictable only if standard RGB primary chromaticities are used.

The first step in the face detection algorithm is using skin segmentation to reject as much non- image based on skin color converting the RGB picture to YCbCr space or to HSV space. An YCbCr space segments the image into a luminosity component and color components. The main advantage of converting the image to the YCbCr domain is that influence of luminosity can be removed during our image processing. In the RGB domain, each component of the picture (red, green and blue) has a different brightness. However, in the YCbCr domain all information about the brightness is given by the Y-component, since the Cb (blue) and Cr (red) components are independent from the luminosity.

There are many ways of segmenting indication on whether a pixel is part of the skin or not. Background and faces can be distinguished by applying maximum and minimum threshold values for both Cb and Cr components.

Considering there is no standard color-image database which can assess the faces in images, we set up the color-image database with 500 skin images. There are some sample images that are shown in the following figure.



**Fig 5.3 Sample of skin images**

**5.1.3 Algorithm for converting an RGB image to YCbCr image**

Formula used for convert an RGB pixel to YCbCr pixel is as follows

Y=0.299R+0.5879G+0.114B

Cb=-0.169R-0.331G+0.5B

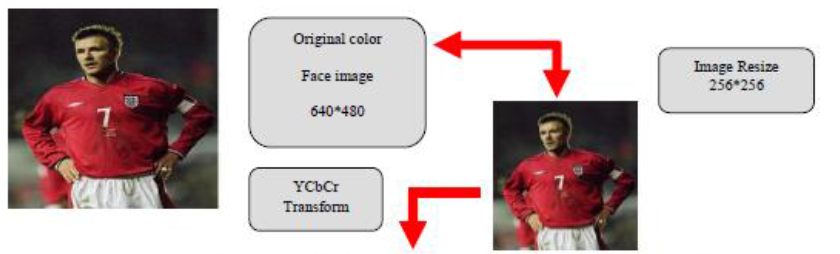
Cr=0.5R-0.419G-0.081B

Step 1: Reading an input image

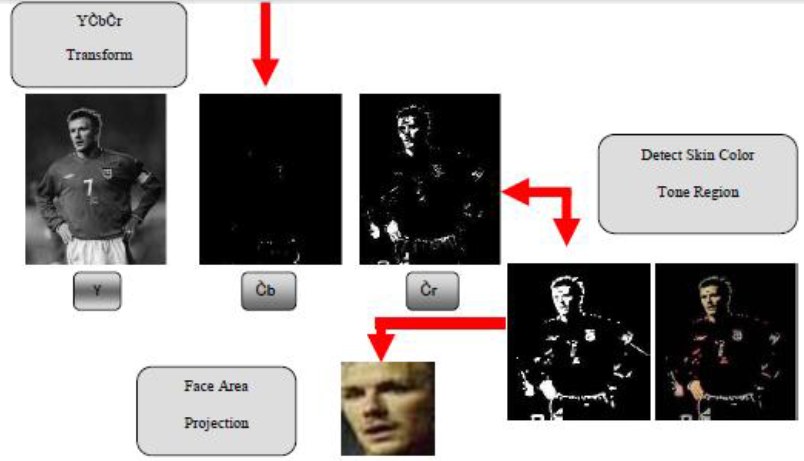
RGB= imread('sample.jpg');

Step 2: Converting an RGB image to YCbCr image

YCBCR = rgb2ycbcr(RGB)







**Fig 5.4 Face detection process**

**5.2 Detection of Eyes condition**

Important factor which helps detect driver fatigue is the state of eyes, i.e. whether they are open or closed. In the state of fatigue, eyelid muscles subconsciously attempt to accelerate the process of going to sleep. Using this property, determining whether eyes are open or closed is done by relying on the difference of brightness intensity of the pupil in the image and its symmetry. Locating the position of the eye in the frame taken from the drivers face is difficult. The position of the eye can be identified by drawing on geometry properties and symmetry. Edging is concerned with locating the position of areas or pixels where the brightness intensity has considerably increased. One of the effective operators for edge detection is the Sobel operator.

The position of the driver's eye is determined by using appropriate threshold. These two areas are separated using edge detection and in accordance with as well as the symmetrical properties of the eye, the gravity center of the eye is determined. Finally the pupil is identified. If eyes are open then it is treated as the normal state during which the alarm is not set off. If eyes are closed then it is treated as the fatigue state during which the alarm is set on.

Edge detection is the process of localizing pixel intensity transitions. The edge detection has been used by object recognition, target tracking, segmentation, and etc. Therefore, the edge detection is one of the most important parts of image processing. There mainly exist several edge detection methods like Sobel .These methods have been proposed for detecting transitions in images. Early methods determined the best gradient operator to detect sharp intensity variations. Commonly used method for detecting edges is to apply derivative operators on images. Computing the gradient in several directions and combining the result of each gradient. The value of the gradient magnitude and orientation is estimated using two differentiation masks.

In this work, Sobel which is an edge detection method is considered. Because of the simplicity and common uses, this method is preferred by the others methods in this work. The Sobel edge detector uses two masks, one vertical and one horizontal. These masks are generally used 3×3 matrices. Especially, the matrices which have 3×3 dimensions are used in matlab (edge.m). The masks of the Sobel edge detection are extended to 5×5 dimensions are constructed in this work. A matlab function, called as Sobel5×5 is developed by using these new matrices toolboxes.

**5.2.1 Algorithm for Sobel edge detection**

Step 1: Accept the input image.

imread(input image)

Reads a rgb color image from the file specified by the string filename. If the file is not in the current folder, or in a folder on the MATLAB path, specify the full pathname

Step 2: Specifies the Sobel method and the dimension of image

BW = edge(I,'sobel')

Sobel method takes input image as its input, and returns a binary image BW of the same size with 1's where the function finds edges in I and 0's elsewhere. And this syntax also specifies the dimension of the image to be 1 dimension.

Step 3: Reading the BW image

imread(BW)

It reads a grayscale image from the file specified by the string filename. If the file is not in the

current folder or in a folder on the MATLAB path, specify the full pathname.

Step 4: Specifies the threshold for the Sobel method

BW = edge(I,'s]’;obel',thresh)

It specifies the threshold for the Sobel method. Edge ignores all edges that are not stronger than thresh. If you do not specify thresh, or if thresh is empty ([]), edge chooses the value automatically. Threshold values ranging from 365 to 535

Step 5: Specifies the direction of detection for the Sobel method

BW = edge (I,'sobel', thresh, direction)

It specifies the direction of detection for the Sobel method. Direction is a string specifying whether to look for 'horizontal' or 'vertical' edges i.e.; gv and gh or 'both' (the default).

Step 6: Returns the threshold value.

[BW, thresh] = edge (I,'sobel',...)

Provides an optional string input. String 'no thinning' speeds up the operation of the algorithm by skipping the additional edge thinning stage. By default, or when 'thinning' string is specified, the algorithm applies edge thinning.

Step 7: Two masks are used for having the maximum edge at vertical and horizontal level i.e; gv and gh.

Mask along horizontal direction i.e. gh

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A11 | A12 | A13 | …… | A1K | + | M11 | M12 | M13 | **→** | B11 | B12 | B13 | ….. | B1K |
| A21 | A22 | A23 | …… | A2K | M21 | M22 | M23 | B21 | **B22** | B23 | ….. | B2K |
| A31 | A32 | A33 | …… | A3K | M31 | M32 | M33 | B31 | B32 | B33 | ….. | B3K |
| : | : | : | : | : |  | | | | | : | : | : | : | : |

Input image Mask gh output image

B22=(A11\*M11)+(A12\*M12)+(A13\*M13)+(A21\*M21)+(A22\*M22)+(A23\*M23)+(A31\*M3) +(A32\*M32)+(A33\*M33)

Mask along horizontal direction i.e. gv

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| A11 | A12 | A13 | …… | A1K | + | M11 | M21 | M31 | **→** | B11 | B12 | B13 | ….. | B1K |
| A21 | A22 | A23 | …… | A2K | M12 | M22 | M32 | B21 | **B22** | B23 | ….. | B2K |
| A31 | A32 | A33 | …… | A3K | M13 | M23 | M33 | B31 | B32 | B33 | ….. | B3K |
| : | : | : | : | : |  | | | | | : | : | : | : | : |

Input image Mask gv output image

Mask along vertical direction i.e. gv

B22=(A11\*M11)+(A12\*M21)+(A13\*M31)+(A21\*M12)+(A22\*M22)+(A23\*M32)+(A31\*M1)+(A32\*M23)+(A33\*M33)

Step 8: Returns vertical and horizontal edge responses to Sobel operators.

[BW,thresh,gv,gh] = edge(I,'sobel',...)

Returns the threshold value of eyes along vertical and horizontal value and the status of the eyes are determined

Step 9: Template matching is done based on status of eyes given from step8.

Step 10: Finally if the fatigue is detected then the alarm is generated to the user.

**5.2.2 Eye Template generation process**

After successful face and facial feature detection, eye state can be determined in every frame using the correlation coefficient template matching method. By considering the diversity to the nearby pixels and the similarity to the eye pixels sufficiently, the specific region of the eyes can be obtained. After this step, Sobel edge detection method is used again to detect the eyes’ precise boundaries. The method has been adopted frequently in the eye detection which mentioned above. Our method starts from left and right side, to find eyes, therefore we can detect the eyes separately. We segment the eyes from the image and use them to generate an eye template, by this means we obtain a rather stable eye template for the status analyzing and reduce the influence of light reflections. The eye template is generated as follows.

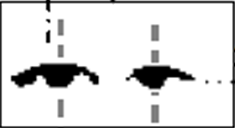
C:\Users\Mai\Desktop\facebook\12.PNG**BINARY PROCESSING**

**Eye template**

**And Operation**C:\Users\Mai\Desktop\facebook\17.PNG

**Or Operation**C:\Users\Mai\Desktop\facebook\16.PNG







**Image alignment**

**Fig 5.5 Eye Template generation process**

To distinguish the driver’s status the eyes’ states should be recognized ahead. There are two factors which can affect the size of the eyes in the frames. On the one hand, human eyes are always different in size. On the other hand, the distance between driver and the camera is the other reason. So we normalize the eye template to a fixed size of 12×30 before feature extraction. For each eye template, eye area, average height of pupil, width to height ratio are the most important features to judge eye’s status which is shown in below table.

**Table 5.1 Eye states and features**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Eye Region | Area(pixel) | Eye Template | Average Height | Ratio |
| Full open | G:\Final img\t1.PNG | 200 | G:\Final img\t4.PNG | 7.6 | 2.8750 |
| Half open | G:\Final img\t2.PNG | 155 | G:\Final img\t5.PNG | 6.8 | 3.0000 |
| Closed | G:\Final img\t3.PNG | 114 | G:\Final img\t6.PNG | 6.0 | 3.1667 |

The eye states can be divided into two types: full open, half open and closed. From the table above we can see that eyes of different states present different features. By analyzing the driver’s eyes states changes while driving we discovered two principles which can indicate driver drowsiness. Firstly, if a driver’s eyes keep closed over 4 consecutive frames it is believed that the driver is drowsy. Secondly, fatigue can be confirmed if a driver’s eyes only change between half open and closed over 8 consecutive frames. The different eye states of full open and half open sometimes cannot be well distinguished which has caused more false judgments and the fast movement of drivers head can result in the driver’s eyes tracking failure. Before the system is put into use we trained it in advance to get different states parameters for the driver aiming at improving the accuracy of the driver’s status analysis.

**5.3 Detection of Head Lowering**

The color space conversion which is useful for segmentation of the skin is the HSV space. Using the features of color space change in HSV, the lowering of the head can be detected by calculating the final area relative to the area of the original image. Using features of color space change in HSV space, the complexity between the level of image and undesired light intensity, which can produces error, is largely reduced. As hue varies from 0 to 1.0, the corresponding colors vary from red through yellow, green, cyan, blue, magenta, and back to red, so that there are actually red values both at 0 and 1.0. As saturation varies from 0 to 1.0, the corresponding colors (hues) vary from unsaturated (shades of gray) to fully saturated (no white component). As value, or brightness, vary from 0 to 1.0, the corresponding colors become increasingly brighter.

As the hue plane image in the preceding figure illustrates, hue values make a linear transition from high to low. If you compare the hue plane image against the original image, you can see that shades of deep blue have the highest values, and shades of deep red have the lowest values. (As stated previously, there are values of red on both ends of the hue scale. To avoid confusion, the sample image uses only the red values from the beginning of the hue range.)

Saturation can be thought of as the purity of a color. As the saturation plane image shows, the colors with the highest saturation have the highest values and are represented as white. In the center of the saturation image, notice the various shades of gray. These correspond to a mixture of colors; the cyan, greens, and yellow shades are mixtures of true colors. Value is roughly equivalent to brightness, and you will notice that the brightest areas of the value plane correspond to the brightest colors in the original image.

Formula for calculating head detection position

NPix Face

HDP =

NPix Face+NPix Black area

Using the features of color space change in HSV, the lowering of the head can be detected by calculating the final area relative to the area of the original image shown in above equation where NPixFace is the number of pixels is facial area and NPixBlackArea is the number of pixels in black area of the image where brightness is equal to zero. When HDP passes the threshold, the alarm relative to the type of risk (lowering of the head) is set off.

1 HDP Thr

0 HDP Thr

Alarm =

In this equation, Thr is the threshold of head's crossing the danger zone. If head detection position is less than threshold value then the alarm is generated.

**5.3.1 Algorithm for converting an RGB image to HSV image**

The steps for converting an RGB image to HSV image:

Step 1: Reading an input image

HSV= imread ('sample.jpeg');

It reads an rgb color image from the file specified by the string filename. If the file is not in the current folder, or in a folder on the MATLAB path, specify the full pathname.

Step 2: Convert image from RGB to HSV

im\_hsv = rgb2hsv(im);

Convert the given RGB color image into HSV color image. It converts a hue-saturation-value (HSV) color map to a red-green-blue (RGB) color map. H is an m-by-3 matrix, where mis the number of colors in the color map. The columns of H represent hue, saturation, and value, respectively. M is an m-by-3 matrix. Its columns are intensities of red, green, and blue, 0 respectively.

Step 3: Separate in channels (Hue, Saturation, Value)

H=HSV(:,:,1);

S=HSV(:,:,2);

V=HSV(:,:,3);

HSV is an m-by*-*n-by-3 image array whose three planes contain the hue, saturation, and value components for the image. For the obtained HSV color image separate each of the channels. The columns of H represent hue, saturation, and value, respectively. M is an m-by-3 matrix. Its columns are intensities of red, green, and blue, respectively.

Step 4: Display the images

subplot(2,2,1), imshow(H)

subplot(2,2,2), imshow(S)

subplot(2,2,3), imshow(V)

subplot(2,2,4), imshow(RGB)

After obtaining the separated channels plot the image for each Hue, Saturation and Value using subplot function and finally display the obtained images using imshow() function.

**5.4 Yawning Detection**

K-means clustering is a partitioning method that treats observations in your data as objects having locations and distances from each other. It partitions the objects into K mutually exclusive clusters, such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. Each cluster is characterized by its centroid, or center point. The function K-means performs K-Means clustering, using an iterative algorithm that assigns objects to clusters so that the sum of distances from each object to its cluster centroid, over all clusters, is a minimum.

With K-means clustering, you must specify the number of clusters that you want to create. We can choose from five different distance measures, depending on the kind of data you are clustering. Each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. K-means computes cluster centroid differently for each distance measure, to minimize the sum with respect to the measure that you specify.

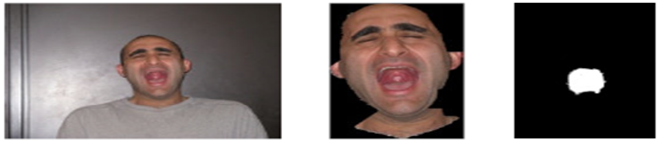
K-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible. You can control the details of the minimization using several optional input parameters to K-means, including ones for the initial values of the cluster centroid, and for the maximum number of iterations. First, load the data and call K-means with the desired number of clusters set to 2, and using squared Euclidean distance. To get an idea of how well-separated the resulting clusters are, you can make a silhouette plot. The silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters. The centroids of each cluster are plotted using circled X's. Three of the points from the lower cluster, plotted with triangles, are very close to points from the upper cluster, plotted with squares. But, in fact, because the upper cluster is so spread out, those three points are closer to the centroid of the lower cluster than to that of the upper cluster, even though they are separated from the bulk of the points in their own cluster by a gap. Because K-means clustering only considers distances, and not densities, this kind of result can occur.

In this paper the other sign of fatigue during driving, which is manifested in a person's face, is frequent yawning that is due to body reflexes when a person is exhausted and about to fall asleep. Various systems have been proposed for measuring yawning some of which are slow and time consuming while others are not very accurate in separating the mouth area at the time of yawning.

An efficient technique is needed that is able to display the changes in face configuration and detect the yawning. Among clustering methods used in segmentation of various parts of the image, the mean-based clustering or K-means was utilized. The objective function was to obtain then minimum distance between the classes, or basically between the image pixels.



**Fig 5.6 Normal Mouth Detection**



**Fig 5.7 Yawning Detection**

K-means uses a two-phase iterative algorithm to minimize the sum of point-to-centroid distances, summed over all K-clusters:

1. The first phase uses batch updates, where in each iteration consists of reassigning points to their nearest cluster centroid, all at once, followed by recalculation of cluster centroids. This phase occasionally does not converge to solution that is a local minimum, that is, a partition of the data where moving any single point to a different cluster increases the total sum of distances. This is more likely for small data sets. The batch phase is fast, but potentially only approximates a solution as a starting point for the second phase.
2. The second phase uses online updates, where points are individually reassigned if doing so will reduce the sum of distances, and cluster centroid are recomputed after each reassignment. For each of the iterations during the second phase consists of one pass though all the points. The second phase will converge to a local minimum, although there may be other local minima with lower total sum of distances. The problem of finding the global minimum can only be solved in general by an exhaustive (or clever, or lucky) choice of starting points, but using several replicates with random starting points typically results in a solution that is a global minimum.

Once the object gets partitioned into clusters each cluster in the partition is defined by its member objects and by its centroid, or center. The centroid for each cluster is the point to which the sum of distances from all objects in that cluster is minimized. K-means computes cluster centroids differently for each distance measure, to minimize the sum with respect to the measure that you specify. Classification of image pixels is based on the brightness intensity. The frequency of brightness intensity of images (histogram) has been used to select the appropriate number of secretions. Based on this classification, pixels can be divided into a maximum of 255 clusters. Here, the number of proposed clusters has been 5 to 10. Thus, classification of image by K-means algorithm can help select proper threshold and NOT brightness intensity acquired from clustering level of yawning detection in input image of frames. Finally, a large part of the segmented area in the image demonstrates the position of mouth and detects the yawning.

**5.4.1 Algorithm for K-means clustering**

The steps for K-means clustering as given below:

Step 1: Partitions the points in the n-by-p data matrix X into k clusters.

IDX = kmeans(X,k)

Input image is converted into n-by-p data matrix X. The X data matrix is converted into

K-clusters. K-means clustering is a method of vector quantization.

Step 2: Returns the k cluster centroid locations in the k-by-p matrix C.

[IDX,C] = kmeans(X,k)

Where each iteration consists of reassigning points to their nearest cluster centroid, all at once, followed by recalculation of cluster centroid.

Step 3: returns the within-cluster sums of point-to-centroid distances in the 1-by-k vector sumd.

[IDX,C,sumd] = kmeans(X,k)

This phase occasionally does not converge to solution that is a local minimum, that is, a partition of the data where moving any single point to a different cluster increases the total sum of distances. This is more likely for small data sets.

Step 4: Returns distances from each point to every centroid in the n-by-k matrix D.

[IDX,C,sumd,D] = kmeans(X,k)

Points are individually reassigned if doing so will reduce the sum of distances, and cluster centroids are recomputed after each reassignment. Each iteration during the second phase consists of one pass though all the points. The second phase will converge to a local minimum, although there may be other local minima with lower total sum of distances.

Step 5: Enables you to specify optional parameter/value pairs to control the iterative algorithm used by K-means.

[...] =kmeans(...,param1,val1,param2,val2,...)

Finally an exact centroid to the k-cluster is found. And additional options are specified with their parameters and its values.

**5.5 CODING**

**FACE SEGMENTATION**

%%

vid=videoinput('winvideo',1);

triggerconfig(vid,'manual');

set(vid,'FramesPerTrigger',1 );

set(vid,'TriggerRepeat', Inf);

% View the default color space used for the data — The value of the ReturnedColorSpace property indicates the color space of the image data.

color\_spec=vid.ReturnedColorSpace;

% Modify the color space used for the data — To change the color space of the returned image data, set the value of the ReturnedColorSpace property.

if ~strcmp(color\_spec,'rgb')

set(vid,'ReturnedColorSpace','rgb');

end

start(vid)

% Create a detector object

faceDetector = vision.CascadeObjectDetector;

faceDetectorLeye = vision.CascadeObjectDetector('EyePairBig');

faceDetectorM = vision.CascadeObjectDetector('Mouth');

for ii = 1:500

trigger(vid);

im=getdata(vid,1); % Get the frame in im

imshow(im)

subplot(3,4,[1 2 5 6 9 10]);

imshow(im)

% Detect faces

bbox = step(faceDetector, im);

if ~isempty(bbox);

bbox = bbox(1,:);

% Plot box

rectangle('Position',bbox,'edgecolor','r');

Ic = imcrop(im,bbox);

subplot(3,4,[3 4]);

imshow(Ic)

bboxeye = step(faceDetectorLeye, Ic);

if ~isempty(bboxeye);

bboxeye = bboxeye(1,:);

Eeye = imcrop(Ic,bboxeye);

% Plot box

rectangle('Position',bboxeye,'edgecolor','y');

else

disp('Eyes not detected')

end

if isempty(bboxeye)

continue;

end

Ic(1:bboxeye(2)+2\*bboxeye(4),:,:) = 0;

% Detect Mouth

bboxM = step(faceDetectorM, Ic);

if ~isempty(bboxM);

bboxMtemp = bboxM;

if ~isempty(bboxMtemp)

bboxM = bboxMtemp(1,:);

Emouth = imcrop(Ic,bboxM);

% Plot box

rectangle('Position',bboxM,'edgecolor','y');

else

disp('Mouth not detected')

end

else

disp('Mouth not detected')

end

[nre nce k ] = size(Eeye);

% Divide into two parts

Leye = Eeye(:,1:round(nce/2),:);

Reye = Eeye(:,round(nce/2+1):end,:);

% Emouth

subplot(3,4,7)

imshow(Leye);

subplot(3,4,8)

imshow(Reye);

subplot(3,4,[11,12]);

imshow(Emouth);

pause(0.00005)

end

end

**EYE DETECTION**

bboxeye = step(faceDetectorLeye, Ic);

if ~isempty(bboxeye);

bboxeye = bboxeye(1,:);

Eeye = imcrop(Ic,bboxeye);

% Plot box

rectangle('Position',bboxeye,'edgecolor','y');

else

disp('Eyes not detected')

end

if isempty(bboxeye)

continue;

end

Ic(1:bboxeye(2)+2\*bboxeye(4),:,:) = 0;

**Sobel Edge Detection**

% Divide into two parts

Leye = Eeye(:,1:round(nce/2),:);

Reye = Eeye(:,round(nce/2+1):end,:);

subplot(3,4,7)

imshow(edge(rgb2gray(Leye),'sobel'));

subplot(3,4,8)

imshow(edge(rgb2gray(Reye),'sobel'));

**FEATURE EXTRACTION**

load DB

cl = {'open','close'};

dim = [30 60;

30 60

40 65];

delete(imaqfind)

vid=videoinput('winvideo',1);

triggerconfig(vid,'manual');

set(vid,'FramesPerTrigger',1 );

set(vid,'TriggerRepeat', Inf);

% View the default color space used for the data — The value of the ReturnedColorSpace property indicates the color space of the image data.

color\_spec=vid.ReturnedColorSpace;

% Modify the color space used for the data — To change the color space of the returned image data, set the value of the ReturnedColorSpace property.

if ~strcmp(color\_spec,'rgb')

set(vid,'ReturnedColorSpace','rgb');

end

start(vid)

% Create a detector object

faceDetector = vision.CascadeObjectDetector;

faceDetectorLeye = vision.CascadeObjectDetector('EyePairBig');

faceDetectorM = vision.CascadeObjectDetector('Mouth');

tic

% Initialise vector

LC = 0; % Left eye closer

RC = 0; % Right eye closer

MC = 0; % Mouth closer

TF = 0; % Total frames

TC = 0; % Total closure

Feature = [];

c1p = 1;

for ii = 1:100

trigger(vid);

im=getdata(vid,1); % Get the frame in im

imshow(im)

subplot(3,4,[1 2 5 6 9 10]);

imshow(im)

% Detect faces

bbox = step(faceDetector, im);

if ~isempty(bbox);

bbox = bbox(1,:);

% Plot box

rectangle('Position',bbox,'edgecolor','r');

Ic = imcrop(im,bbox);

subplot(3,4,[3 4]);

imshow(Ic)

bboxeye = step(faceDetectorLeye, Ic);

if ~isempty(bboxeye);

bboxeye = bboxeye(1,:);

Eeye = imcrop(Ic,bboxeye);

% Plot box

rectangle('Position',bboxeye,'edgecolor','y');

else

disp('Eyes not detected')

end

if isempty(bboxeye)

continue;

end

Ic(1:bboxeye(2)+2\*bboxeye(4),:,:) = 0;

% Detect Mouth

bboxM = step(faceDetectorM, Ic);

if ~isempty(bboxM);

bboxMtemp = bboxM;

if ~isempty(bboxMtemp)

bboxM = bboxMtemp(1,:);

Emouth = imcrop(Ic,bboxM);

% Plot box

rectangle('Position',bboxM,'edgecolor','y');

else

disp('Mouth not detected')

continue;

end

else

disp('Mouth not detected')

continue;

end

[nre nce k ] = size(Eeye);

% Divide into two parts

Leye = Eeye(:,1:round(nce/2),:);

Reye = Eeye(:,round(nce/2+1):end,:);

subplot(3,4,7)

imshow(Leye);

subplot(3,4,8)

imshow(Reye);

subplot(3,4,[11,12]);

imshow(Emouth);

Leye = rgb2gray(Leye);

Reye = rgb2gray(Reye);

Emouth = rgb2gray(Emouth);

% Template matching using correlation coefficient

% Left eye

% Resize to standard size

Leye = imresize(Leye,[dim(1,1) dim(1,2)]);

c1 =match\_DB(Leye,DBL);

subplot(3,4,7)

title(cl{c1})

% Right eye

% Resize to standard size

Reye = imresize(Reye,[dim(2,1) dim(2,2)]);

c2 = match\_DB(Reye,DBR);

subplot(3,4,8)

title(cl{c2})

% Mouth

% Resize to standard size

Emouth = imresize(Emouth,[dim(3,1) dim(3,2)]);

c3 = match\_DB(Emouth,DBM);

subplot(3,4,[11,12]);

title(cl{c3})

if c1 == 2

LC = LC+1;

if c1p == 1

TC = TC+1;

end

end

if c2==2

RC = RC+1;

end

if c3 == 1

MC = MC + 1;

end

TF = TF + 1; % Total frames

if toc>10

Feature = [Feature;LC/TF RC/TF MC/TF TC]

tic

% Initialise vector

LC = 0; % Left eye closer

RC = 0; % Right eye closer

MC = 0; % Mouth closer

TF = 0; % Total frames

TC = 0; % Total closure

end

c1p = c1;

pause(0.00005)

end

end

save FA Feature

**Yawning Detection**

% K means clustering

X = Emouth(:);

[nr1 nc1 ] = size(Emouth);

cid = kmeans(double(X),2,'emptyaction','drop');

kout = reshape(cid,nr1,nc1);

subplot(3,4,[11,12]);

% Segment

Ism = zeros(nr1,nc1,3);

% Ism(:,:,3) = 255;

% Ism(:,:,3) = 125;

Ism(:,:,3) = 255;

bwm = kout-1;

bwm3 = cat(3,bwm,bwm,bwm);

Ism(logical(bwm3)) = Emouth3(logical(bwm3));

imshow(uint8(Ism));

**Head Lowering:**

function segment = skin\_seg2(I1)

% Convert image to double precision

I=double(I1);

[hue,s,v]=rgb2hsv(I);

% Ycbcr = rgb2ycbcr(I);

% % cb=Ycbcr(:,:,2)+128;

% % cr=Ycbcr(:,:,3)+128;

cb = 0.148\* I(:,:,1) - 0.291\* I(:,:,2) + 0.439 \* I(:,:,3) + 128;

cr = 0.439 \* I(:,:,1) - 0.368 \* I(:,:,2) -0.071 \* I(:,:,3) + 128;

[w h]=size(I(:,:,1));

segment = 140<=cr & cr<=165 & 140<=cb & cb<=195 & 0.01<=hue & hue<=0.1;

segment=imfill(segment,'holes');

segment=bwmorph(segment,'dilate');

segment=bwmorph(segment,'majority')

**Chapter 6**

**TESTING AND IMPLEMENTATION**

The purpose of testing is to discover errors. Testing is the process of trying to discover faults or defects in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product it is the process of exercising software with the intent of ensuring that the software system meets its requirements and user exceptions and does not fail in an unacceptable manner.

A good test case is the one that has a high probability of finding an as yet undiscovered error. A successful test is the one that uncovers a undiscovered error. Testing may be carried out during the implementation phase to verify that the software behaves as intended by its designers and after implementation is complete.

* 1. **Software Testing**

A primary purpose of testing is to detect software failures so that defects may be discovered and corrected. Testing cannot establish that a product functions properly under all conditions but can only establish that it does not function properly under specific conditions. The scope of software testing often includes examination of code as well as execution of that code in various environments and conditions as well as examining the aspects of code: does it do what it is supposed to do and do what it needs to do. In the current culture of software development, a testing organization may be separate from the development team. There are various roles for testing team members. Information derived from software testing may be used to correct the process by which software is developed.

* 1. **Software Testing Types**

Software testing life cycle is the process that explains the flow of the tests that are to be carried on each step of software testing of the product. The V-model i.e. Verification and validation model is a perfect model which is used in the improvement of the software project. This model contains software development life cycle on one side and software testing life cycle on the other hand side. A checklist for software tester sets a baseline that guides him to carry on the day-to-day activities.

**6.2.1 Black Box Testing**

It explains the process of giving the input to the system and checking the output, without considering how the system generates the output. It is also called as Behaviour Testing.

**Functional Testing**

In this type of testing, the software is tested for the functional requirements. This checks whether the application is behaving according to the specification.

**Performance Testing**

This type of testing checks whether the system is performing properly, according to the user’s requirements. Performance testing depends upon the Load and Stress Testing that is internally or externally applied to the system.

**Integration Testing**

Integration Testing is the phase in [software testing](http://en.wikipedia.org/wiki/Software_testing) in which individual software modules are combined and tested as a group. This mostly focuses in the design and construction of the software architecture. The purpose of integration testing is to verify functional, performance, and reliability [requirements](http://en.wikipedia.org/wiki/Requirement) placed on major design items.

Integration testing is further classified into Bottom-up Integration and Top-Down Integration testing.

**Bottom-up Integration Testing**

It is an approach to integrated testing where the lowest level components are tested first, then used to facilitate the testing of higher level components. The process is repeated until the component at the top of the hierarchy is tested. All the bottom or low-level modules, procedures or functions are integrated and then tested. After the integration testing of lower level integrated modules, the next level of modules will be formed and can be used for integration testing.

**Top-Down Integration testing**

It is an approach to integrated testing where the top down integrated modules are tested and the branch of the module is tested step by step until the end of the related module.

**System Testing**

System testing is the testing conducted on a complete, integrated system, to evaluate the system’s compliance with the specified requirements. This type of software testing validates that the system meets its functional and non-functional requirements and also intended to test beyond the bounds defined in the software/hardware requirement specifications.

**6.2.2 White Box Testing**

It is the process of giving the input to the system and checking, how the system processes the input, to generate the output. It is mandatory for a tester to have the knowledge of the source code.

**Unit Testing**

This type of testing is done at the developer’s site to check whether a particular piece/unit of code is working fine. Unit testing deals with testing the unit as a whole.

**6.3 A list of sanity Test Cases**

Prior to any testing it’s mandatory to write test cases, so that it helps in

1. Giving an approach, description, pre-conditions to achieve the expected result.
2. Test cases are reusable, helps in regression testing while releasing various versions.

**Table 6.1 Test cases for the applications**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Test case ID** | **Test Case** | **Input** | **Expected output** | **Obtained Output** | **Result** |
| TID1 | Detecting face | Input Frame | Face detected | Segmentation of face | Pass |
| TID2 | Detecting eyes | Segmentation of face | Eyes detected | Sobel edge eyes | Pass |
| TID3 | Detecting Head lowering | Input frames | Head lowered detection | Change position | Pass |
| TID4 | Detection of mouth | Segmentation of face | Detected mouth | Clustered mouth with large hole | Pass |
| TID5 | Detection of face | Closed eyes | Fatigue | Alarm generated | Pass |
| TID6 | Detection of face | Mouth opened | Fatigue | Alarm generated | Pass |
| TID7 | Detection of face | Mouth closed | Non Fatigue | No Alarm | Pass |
| TID8 | Detection of face | Eyes opened | Non Fatigue | No Alarm | Pass |
| TID9 | Detection of face | Mouth closed | Non Fatigue | No Alarm | Pass |
| TID10 | Detection of face | Eyes closed and Mouth closed | Fatigue | Alarm generated | Pass |
| TID11 | Detection of face | Eyes opened and Mouth opened | Fatigue | Alarm generated | Pass |

**Chapter 7**

**RESULT ANALYSIS**

The average accuracy (AAC), the detection rate (DR) and false alarm rate (FAR) in accordance with equation has been calculated. These three factors, which have been proposed for assessing the detection accuracy of the video sequence, indicate the acceptable performance of the proposed system in detecting the signs of fatigue in driver's face at the time of driving.

NTP + NTN

AAC =

NTP+NTN+NFP+NFN

Number of true positive

DR =

Number of drowsy video frames

FAR =

Number of drowsy video frames

In these equations,

NTP is the number of frames in which fatigue has been detected by the algorithm

NFN is the number of frames in which the algorithm has failed to detect fatigue

NTN is the number of frames in which there is no sign of fatigue and the algorithm has not detected any by mistake.

NFP is the number of frames in which there is no sign of fatigue but the algorithm has detected some by mistake.

**7.1 Result Analysis Table**

**Result analysis for applications**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Video clip | Number of frames | Number of frames with fatigue in drivers face | | Number of frames without fatigue in drivers face | | Average  Accuracy  (AAC) | Detection Rate  (DR) | False Alarm Rate  (FAR) |
| NTP | NFN | NTN | NFP |
| Movie1 | 90 | 31 | 2 | 46 | 6 | 90.58% | 93.93% | 11.53% |
| Movie2 | 90 | 5 | 0 | 16 | 1 | 95.45% | 100% | 5.88% |
| Movie3 | 90 | 11 | 2 | 39 | 2 | 92.59% | 84.61% | 4.87% |
| Total | 270 | 47 | 4 | 101 | 9 | 92.993% | 92.8% | 7.426% |

The above table is a result analysis of the application that we have developed. This analysis is based on driver’s fatigue and non-fatigue detection and This detection gives the average accuracy (AAC), detection rate (DR) ,false alarm rate(FAR).

To evaluate the performance of our proposed application, several experiments were performed. Table shows the convolution matrix for detecting the fatigue to determine the Average Accuracy and Detection Rate (DR), False Alarm Rate (FAR).

**7.2 Final outcome**

When the person’s face is captured by the camera first it will extract the background and foreground classes then for the extracted face part segmentation is done. By observing the eye and mouth state it will check for driver fatigue.

**Snapshot for driver Non-Fatigue**

Here we can observe that eye is opened and mouth is closed so there is no sign of fatigue detected. Hence Alarm is not generated.

**Snapshot for closing eyes**

Here the states of eyes are closed so there is a sign of fatigue which is detected. Hence Alarm is generated.

**Snapshot for opening of mouth**

Here the state of mouth is opened so there is a sign of fatigue which is detected. Hence Alarm is generated.

**Eyes are not detected**

**Chapter 8**

**CONCLUSION AND FUTURE ENHANCEMENT**

**8.1 Conclusion**

The system proposed in this paper is acceptable level of performance and an average accuracy of 93.18%. The high fatalities of road accidents, which is primarily due to human errors committed out of fatigue, justifies the use of this system to alarm drivers at the time of driving. High-speed data processing and great accuracy distinguish this system from the similar ones. The development and improvement of this system can save the lives of millions of people annually.

The processing rate or framing of this camera is 15 fps and the first video sequence is related to the state where the head is in a lowered position which includes 85 sample frames. The second video sequence deals with the recording of the open or closed state of the eyes in which 48 image frames indicate that eyes are closed in a 6 seconds period while 65 frames show that the eyes are normally open. The third video sequence shows the yawning or the frequent opening of the driver's mouth. And finally the fourth video sequence is a combination of all three modes and its recon ling takes a longer time. The average accuracy (AAC), the detection rate (DR) and false alarm rate (FAR) has been calculated. These three factors, which have been proposed for assessing the detection accuracy of the video sequence, indicate the acceptable performance of the proposed system in detecting the signs of fatigue in driver's face at the time of driving.

**8.2 Future Enhancement**

In future works, a driver’s distraction identification system will be developed. With its complex and ever-changing nature, including the effect of the light and the condition of shooting environment, it makes the skin segmentation of human faces in color images severely affect face detection, and also makes it an important research topic. A method of face-region segmentation based on skin detection has been proposed in this paper, which partly comes from other studies. Compared with the conventional method of segmentation, we put these methods into this article, such as adjudging the images with the light interference, enhancing the images and improved threshold segmentation. Determination of the light interference not only improves the accuracy of image segmentation in the follow-up processing, but also expands the scope of application with skin segmentation in color images. Image enhancement mainly deals with the skin-likelihood image which is transformed through Gaussian model, aiming at getting the gray images with better and higher contrast. In this article, we use the method that combines the histogram with Otsu, which is the initial use of histogram threshold method to determine the threshold, and then we set it as the threshold of Otsu, and respectively, find the best.

On detecting the drowsiness the system generates the alarm to wake up the driver. Ultrasonic sensors are connected on the left and behind the car to detect the distance between the car and the road side and to detect any other car behind. Based on the situation the system tries to reduce the speed and to stop the car.When the driver is feeling drowsy the future technology would be enhanced so that the sensors are applied in cars and the car gives the indication to the neighboring vehicle and just moves towards the lane and park the vehicle.

**BIBLIOGRAPHY**

[1] C. C. Liu, S. G. Hosking, "Predicting driver drowsiness using vehicle measures, recent insights and future challenges," J. safety Res, 40 (4) , pp 239 -245.2009 .

[2] G. Hosseini, H. Hossein-Zadeh, A "Display driver drowsiness warrning system", International Conference of the road and traffic accidents, Tehran University, 2006 .

[3] L. M. Bergasa, J. Nuevo, "Real-Time system for monitoring driver vigilance," IEEE Transactions on Intelligent Transportation Systems, 7(1) , pp 63-77. 2006.

[4] M. J. Flores, 1 M. Anningll and A Escalera "Real-Time Warning System for Driver Drowsiness Detection Using Visual Information", J Intell Robot System, DOT 10.100 7 /sI0846 -009 -9391 -1 , pp 10 -33. 18 N. 2009.

[5] M. Rizon and T. Kawaguchi, "Automatic eye detection using intensity and edge information”, Proceedings IEEE TENCON, vol. 2, pp. 415- 420, 2000.

[6] P. Gejgus, M. Sparka, "Face Tracking in Color Video Sequences", The Association for Computing Machinery Inc., New York, 2003.

[7] R G Gonzales, and RE. Woods, " Digital Image Processing," Second Edition, Prentice Hall, 2002.

[8] T. Azim, M.A. Jaffar, AM. Mirza. "Automatic Fatigue Detection of Drivers through Pupil Detection and Yawning Analysis." In: Proc. Fourth International Conference on Innovative Computing, Information and Control, 2009, pp. 441 -445.