

A Review on Driver Face Monitoring Systems for Fatigue and Distraction Detection

Mohamad-Hoseyn Sigari¹, Muhammad-Reza Pourshahabi²
Mohsen Soryani³ and Mahmood Fathy³

¹ Control and Intelligent Processing Center of Excellence (CIPCE)
Electrical and Computer Engineering Department, College of Engineering
University of Tehran, Tehran, Iran

² Machine Vision Research Lab., Computer Engineering Department, College of Eng.
Ferdowsi University of Mashhad, Mashhad, Iran

³ Computer Engineering Department, Iran University of Science and Technology,
Tehran, Iran

*hoseyn_sigari@ieee.org, reza.pourshahabi@gmail.com, soryani@iust.ac.ir,
mahfathy@iust.ac.ir*

Abstract

Every year, many car accidents due to driver fatigue and distraction occur around the world and cause many casualties and injuries. Driver face monitoring systems is one of the main approaches for driver fatigue or distraction detection and accident prevention. Driver face monitoring systems capture the images from driver face and extract the symptoms of fatigue and distraction from eyes, mouth and head. These symptoms are usually percentage of eyelid closure over time (PERCLOS), eyelid distance, eye blink rate, blink speed, gaze direction, eye saccadic movement, yawning, head nodding and head orientation. The system estimates driver alertness based on extracted symptoms and alarms if needed. In this paper, after an introduction to driver face monitoring systems, the general structure of these systems is discussed. Then a comprehensive review on driver face monitoring systems for fatigue and distraction detection is presented.

Keywords: *driver face monitoring system, fatigue detection, drowsiness detection, distraction detection*

1. Introduction

Improvement of public safety and the reduction of accidents is one of the important goals of the intelligent transportation systems (ITS). One of the most important factors in accidents, especially on rural roads, is the driver fatigue and monotony. Fatigue reduces driver perception and decision-making capability to control the car. Research shows that usually after 2-3 hours of continuous driving, driver is fatigued and steering performance deteriorated [1]. In the early afternoon hours, after eating lunch and at midnight, driver drowsiness is much more than other times. In addition, drinking alcohol, drug addiction and using hypnotic medicines can lead to loss of consciousness [1, 2, 3].

In different countries, different statistics was reported about accidents that happened due to driver fatigue and distraction. Generally, the reason of about 20% of the crashes and 30% of fatal crashes is the driver drowsiness and lack of concentration. In single-vehicle crashes

(accidents that only one car is damaged) or crashes involving heavy vehicles, the percentage of accidents was reported up to 50% [2, 4, 5, 6, 7, 8, 9, 10]. According to the current studies, it is expected that the amount of crashes will be reduced by 10% to 20% using driver face monitoring systems [11].

Now, the concepts of fatigue, drowsiness and distraction from the viewpoint of physiology and psychology are explained. However from the viewpoints of physiology and psychology, the concepts of fatigue and drowsiness are different, but in the literatures found in the field of ITS, fatigue and drowsiness are synonymous terms. Hypo-vigilance is another term means lack of consciousness and may include drowsiness, distraction or both.

A precise and scientific definition for fatigue has not been presented yet; therefore, there is not any quantitative criterion to measure it. Fatigue occurs in three different types: sensory fatigue, muscle fatigue and cognitive fatigue. Sensory fatigue and muscular fatigue are only measurable and there is not any way to measure cognitive fatigue [7, 9]. Although a precise definition for fatigue is not presented yet, but there is a relation between the fatigue/drowsiness and body temperature, electrical skin resistance, eye movement, breathing rate, heart rate and brain activity [3, 9, 12, 13]. However, the best tool for measuring fatigue and drowsiness is brain activity monitoring, but in this approach, brain signals must be received from the electrodes that connected to the driver head which make it as an intrusive approach. After monitoring of brain activity, the most significant symptom of fatigue is appeared in eye. According to the researches, the latency between the visual stimulus and its response is one of the main measures to determine the consciousness. This latency is known by a parameter called Psychomotor Vigilance Task (PVT) that shows the response speed of a person to his/her visual stimulation. Researches show that there is a very close relation between PVT and the percentage of closed eyelids in a period of time. The percentage of eyelid closure over time called PERCLOS [8]. Therefore, there is a close relation between fatigue and percentage of eye closure. Driver face monitoring systems use this relation to estimate driver fatigue/drowsiness.

Monotony of a certain task can reduce the concentration of person and may cause distraction. Monotony is caused by three main reasons: (1) lack of personal interest, (2) doing a repetitive task for long time and (3) external factors (like talking with mobile phone). Monotony in driving usually is caused by the second and third reasons. Prolonged driving on highways with flowing traffic has a negative effect on driver concentration. In this case, driver is not fatigued, but due to the monotony of driving, his/her concentration will gradually be decreased and the driver will not have a careful control on the vehicle. Driver distraction can also be caused by talking to people or mobile phone and listening to music [7, 9]. Driver distraction can be estimated by head and gaze direction determination. The main problem for distraction detection is that if head is forward and eyes look toward the road, the driver does not necessarily pay attention to the road. In other words, looking toward the road is not paying attention to it [9].

Due to the importance of early detection of driver fatigue and drowsiness to avoid accidents, many researches were done on this subject in the past decade. The researches on the methods for driver distraction detection are also being done, but are less developed than the methods of driver fatigue detection. However, fatigue and distraction can be considered as two separate concepts, since both of these factors reduce driver alertness; both categories are investigated in some studies. We can divide the most important approaches for fatigue/distraction detection into three categories: (1) approaches based on bioelectric signals (*e.g.*, EEG and ECG), (2) approaches based on steering motion, and (3) approaches based on driver face monitoring. These approaches can be investigated from different viewpoints

including ability of fatigue detection, ability of distraction detection, accuracy, simplicity and detection speed (Table 1).

Distraction detection is more difficult than fatigue detection, but the approaches based on steering motion and the approaches based on driver face monitoring can estimate the lack of the driver concentration in limited circumstances. Comparison between the main approaches is briefly listed in Table 1. The approaches based on bioelectric signals have a very good accuracy and speed at detecting fatigue, but they are usually intrusive. Additionally, the driver distraction cannot be identified from his/her bioelectric signals using current technology by now. The approaches based on driver face monitoring have lower accuracy than the approaches based on steering motion, but they can detect driver fatigue and distraction earlier. Three main approaches for driver fatigue/distraction detection are compared subjectively in Table 1.

Table 1. Comparison between the main approaches for driver fatigue and distraction detection

	Approaches based on Bioelectric Signals	Approaches based on Steering Motion	Approaches based on Driver Face Monitoring
Fatigue Detection	Yes	Yes	Yes
Distraction Detection	No	Yes	Yes
Accuracy	Very Good	Good	Moderate
Simplicity	Difficult	Relatively Easy	Easy
Detection Speed	Very Fast	Slow	Fast

This paper organized in ten sections. In this section, a brief introduction to the concepts *fatigue* and *distraction* was presented. Also, different approaches for driver fatigue and distraction detection were introduced and compared from different aspects. In section 2, a general block diagram of driver face monitoring systems are investigated and the main challenges of such systems are explained. In Sections 3, 4 and 5, different methods for face detection, eye detection and detection of other components of face will be reviewed. Tracking methods and their application are explained in Section 6. In Section 7, different methods for symptom excretion are investigated. The methods for fatigue and distraction estimation based on extracted symptoms are explained in Section 8. In Section 9, a summary of article which are reviewed in this paper will be presented. Finally, the last section is related to conclusions and future works.

2. Driver Face Monitoring System

The driver face monitoring system is a real-time system that investigates driver physical and mental condition based on processing of driver face images. Driver status can be detected from eyelids closure, blinking, gaze direction, yawning and head movement. This system will alarm in hypo-vigilance states such as drowsiness, fatigue and distraction. The systems based on the driver face monitoring can be divided into two general categories. In first category, driver fatigue and distraction is detected only by processing of eye region. There are many researches based on this approach. The main reason of this large amount of researches is that main symptoms of fatigue and distraction appear in driver eyes. Moreover, processing of eye region instead of total face region has less computational complexity.

In the other category, symptoms of fatigue and distraction are detected not only from eyes, but also from other regions of face and head. In these approaches, not only activities of eyes are considered, but also other symptoms such as yawning and head orientation is extracted.

Major parts of driver face monitoring system are:

- Imaging
- Hardware platform and the processor
- Intelligent Software

2.1. Imaging

The part for imaging includes lighting and camera and may include imaging controller if necessary. Since driver face monitoring system should work in all light conditions, lighting and camera selection is one of the most important stage of system design. Lighting devices not only should provide enough light in environment, but also should not hurt driver eyes. Thus, near infrared (IR) spectrum is usually used in lighting. Camera should be chosen based on the lighting spectrum as well. Visible spectrum is also used in the driver face monitoring systems, but face lighting in night is troubling for driver vision system. Thus, it seems that visible spectrum is not very useful in real conditions.

Another approach for imaging is pulsed lighting/imaging. In pulsed imaging, a controller is used to synchronize lighting and imaging. In this case, the controller adjusts light source(s) turning on/off and camera aperture opening. The pulsed imaging approach usually used in near IR lighting spectrum. The main advantages of pulsed imaging method are reduction of environment noise impact on the image and reduction of light source power consumption [14].

2.2. Hardware Platform and the Processor

Hardware platform may include main board, one or more processors and Human Machine Interface (HMI). In real systems, hardware platform should be in the form of embedded system and should be inexpensive as possible. Processor of hardware can be one of the types of conventional microprocessors, Digital Signal Processors (DSP), Field Programmable Gate Array (FPGA) and Application-Specific Integrated Circuit (ASIC) or a combination of them. Image processing algorithms for analysis and interpretation of the facial images usually has a large computational complexity. Although the conventional microprocessors have much flexibility to execute different software, they are not suitable for this purpose, because in this case the system usually will not be real-time. On the other hand, ASICs have high efficiency to process data in real-time, but they are expensive. Moreover, ASIC has no flexibility for programming. DSP and FPGA are faster than conventional microprocessors but they are usually slower than ASIC, but their main advantage compared to ASIC, is their flexibility in programming. DSP and FPGA are flexible to programming via software and hardware respectively. Therefore, in the driver face monitoring systems, DSP or FPGA is often used as processor.

The system proposed in [17] is based on DSP. The proposed systems that are based on FPGA technology can be found in [15] and [16]. In addition, there are a few designed driver face monitoring systems using conventional microprocessors for real environment, except for system simulations. For example, the proposed system in [14] uses a PC 104+ computer. These systems usually use conventional microprocessors in the form of embedded computers.

2.3. Intelligent Software

Intelligent software is the most important part of driver face monitoring system and is divided into two main parts: image processing algorithms and decision-making algorithms.

The main goals of image processing algorithms include preprocessing, detection and tracking of face, eyes and other facial components, and extraction of appropriate symptom from facial images. Indeed, image processing algorithm is the main part of driver face monitoring system. These algorithms usually require large amounts of memory and impose a high computational load on the processor and hardware. To design an accurate and real-time system, we should focus on these algorithms.

After extraction of appropriate symptom from images, decision-making algorithms determine the level of driver alertness based on extracted symptoms. Finally, an appropriate output is generated for the system. Decision-making algorithms should be able to detect driver fatigue and distraction and make an appropriate decision; hence, they are important. The shorter duration that decision-making algorithm can detect driver drowsiness or distraction result in higher performance of the system. Meanwhile, after driver hypo-vigilance detection, warning method is important too.

Figure 1 shows a general flowchart of intelligent software related to a driver face monitoring system. Although in different systems this flowchart is somewhat variable, but the overall flowchart is plotted in Figure 1.

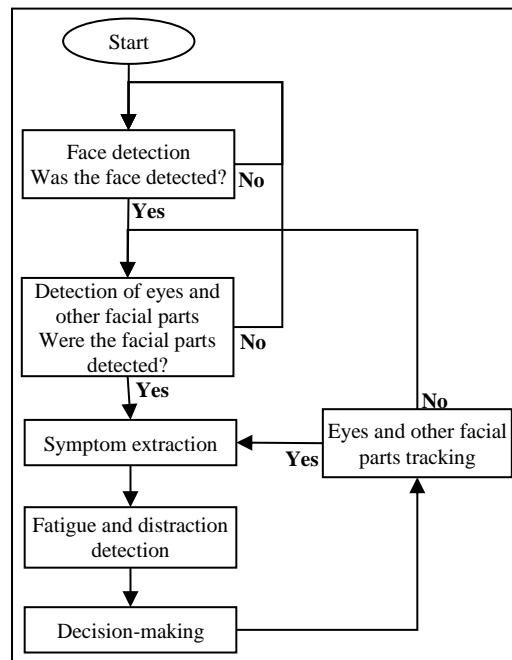


Figure 1. General flowchart of driver face monitoring systems

Alarm can be transmitted to humans through three main senses: visual, auditory and tactility. In most of designed systems, the only system output is an alarming sound. This is a simple output and the variable parameters of sound are amplitude, frequency and tone of sound [18]. Visual warning can be used as a warning display or warning light. This method was used in AWAKE project [18]. However this method is suitable for warning about driver distraction but it has little effect for warning about drowsiness. Another way of warning is

alarming via the sense of touch. This warning is usually in the form of the steering vibration [19] or vibration of the driver safety belt [18].

2.4. Main Challenges

In driver face monitoring systems, two main challenges can be considered: (1) "how to measure fatigue?" and (2) "how to measure concentration?". These problems are the main challenges of a driver face monitoring system.

The first challenge is how to define fatigue exactly and how to measure it. Despite the progress of science in physiology and psychology, there is still no precise definition for fatigue. Certainly, due to the lack of precise definition of fatigue, there is not any measurable criterion to measure it [9]. Although a precise definition for fatigue is not defined yet, but there is a relation between fatigue and body temperature, electrical resistance of skin, eye movement, breathing rate, heart rate and brain activity [3, 9, 12, 13]. One of the first and most important signs of fatigue appears in eyes. As previously mentioned, there is a very close relation between PVT and the percentage of eyelid closure over time (PERCLOS). Therefore, almost in all driver face monitoring systems, processing of eye region and investigation of closed eyelids is the most important criterion used to estimate fatigue.

The second challenge is measurement of driver attention to road. Driver attention can be estimated from driver head and gaze direction. The main problem is that if driver head is forward and driver looks toward road, the driver does not necessarily pay attention to the road. In other words, looking toward road is not paying attention to it [9].

Beside the mentioned main challenges of the driver face monitoring systems, developing a real-time system on conventional hardware platforms, reducing error of the system in detection of face and its components, reducing error of face tracking and increasing system accuracy in detection of fatigue and distraction are considered as other problems of such systems.

2.5. Evaluation Criteria

Complete and exact evaluation of driver face monitoring system is performed based on False Positive Rate (FPR) and False Negative Rate (FNR) [72]. For face and eye detection algorithms, these criteria are also used to determine the accuracy of the algorithm [20]. Generally, FPR and FNR are used for complete evaluation of a detection algorithm.

FPR, also known as false alarm ratio or type I error, is the system error rate in the false detection of samples; while FNR, also known as type II error, is the system error rate related to samples that falsely not detected. According to Table 2, FPR and FNR are calculated by equation (1) and (2). In an ideal system, both of FPR and FNR are zero.

$$FPR = \frac{FP}{FP+TN} \quad (1)$$

$$FNR = \frac{FN}{FN+TP} \quad (2)$$

Table 2. Definition of FPR and FNR for a detection system

		Test Result	
		0	1
Ground Truth	0	TN (True Negative)	FP (False Positive)
	1	FN (False Negative)	TP (True Positive)

3. Face Detection

In most of the driver face monitoring systems, face detection is the first part of image processing operations. In systems that fatigue and distraction is detected based on processing of facial region, face detection is considered as very important part of the system. Also in most of methods based on processing of eye region, due to difficulty of eye detection directly, face is detected at first and then eyes are detected. The most important problems of face detection are [20]:

- In-plane face rotation
- Out-of-plane face rotation
- The presence or the absence of makeup, beard and glasses
- Mental conditions (happiness, crying, and etc.)
- Illumination conditions
- Covering part of the face with an object and
- Real-time processing.

Face detection methods can be divided into two general categories [20]: (1) feature-based and (2) learning-based methods. Learning-based methods usually more robust than feature-based methods, but they often take more computational resources. However, these methods can achieve a detection rate about 80-90% or higher in laboratory conditions, but both of them usually fail in real conditions, especially in night light, In this section, the most popular face detection methods are investigated.

3.1. Feature-based Methods

In feature-based methods, the main assumption is that face in the image can be detected based on some simple features, independent of ambient light, face rotation and pose. These methods are usually used for detection of one face in image. In noisy image or the environment with low illuminations, these algorithms have low accuracy [20]. One of the simplest methods for face detection is projection. Projection is sum of pixel values in a specified direction. Projection is usually chosen along a vertical or horizontal direction. This method can be applied on binary or gray-level images. Projection along the horizontal and vertical dimensions of image I with dimensions $M \times N$ is computed by equations (3) and (4) respectively.

$$P_H(j) = \sum_{i=1}^M I(i, j) \quad (3)$$

$$P_V(i) = \sum_{j=1}^N I(i, j) \quad (4)$$

The main problem in this method is that the background is assumed uniform. If the background is complex or cluttered, face detection is failed. In the system was proposed by Wang *et al.* [16], the position of face was determined by the vertical projection of the gray level image. In this paper, the driver background was assumed non-cluttered while the face has lighter pixels than background.

Another feature-based face detection approach is based on skin color. Skin color model is usually determined based on the probability distribution in a color space. Then, facial region is detected in image by applying a threshold on the modeled distribution. For skin color modeling, RGB [17, 21], YCbCr [22] or HIS [23, 24] color space may be used. These methods are usually appropriate for face detection in images with simple background, but its error is high when the face background is cluttered. Methods based on the color model are suitable for face detection in the visible spectrum and are not applicable in the IR spectrum.

3.2. Learning-based Methods

Learning-based methods deal with face detection using a number of training samples. These methods benefit from statistical models and machine learning algorithms. Generally, learning-base methods have less error rates in face detection, but these methods usually have more computational complexity [20]. Viola *et al.* [25] presented an algorithm for object detection, which uses very simple features named Haar-like features. In this algorithm, many numbers of Haar-like features are extracted from the image, and a number of effective features are selected using AdaBoost algorithm, and then these features are processed in a hierarchical structure similar to the decision tree. Due to the simple extracted features and selection of the best features, this algorithm is relatively fast and robust. This algorithm was used in [7, 26, 27, 28, 29, 44, 51, 59, 74] for face detection.

In addition to the above method which is very efficient and fast, other learning-based methods were also used in some researches, but as they are less efficient than previous method, they are not popular. For example, Hamada *et al.* [30] used neural networks for face detection. In this method, an edge detection method was applied on the image, and then the result image was scanned by a window and is evaluated by neural network to detect face.

4. Eye Detection

In all driver face monitoring systems, eye region is always processed for symptom extraction, because the most important symptoms are related to the eyes activity. Therefore, eye detection is required before processing of eye region. Eye detection methods can be divided into three general categories: (1) methods based on imaging in IR spectrum, (2) feature-based methods and (3) other methods.

4.1. Methods based on Imaging in IR Spectrum

One of the quick and relatively accurate methods for eye detection is the methods based on imaging in IR spectrum. In these methods, physiological and optical properties of eye in IR spectrum are used.

Retina reflects about 90% and 40% of the radiations with a wavelength of 850 nm and 950 nm respectively. Thus, eye can be detected using two separate cameras in front of driver face. This method was used by Grace *et al.* [14] for eye tracking, but positions of eyes are determined manually at the first time. In this method, two different band-pass optical filters were used. One of them passes IR light with wavelength of 850 nm and the other one passes IR light with wavelength of 950 nm. Each of these optical filters was mounted on a separated camera. Direction of both cameras is toward the driver head. In this system, it was assumed that images captured from both of the cameras are similar and only differ in retina area. The main disadvantage of this system is that it uses two cameras. This property of the system causes increasing total price of system and the complexity of imaging.

In [6, 31, 32, 33, 58], IR lighting was provided by two LED sets which play a very important role in these systems. According to Figure 2, LEDs are divided into two concentric circles with different radius. At any moment, only one of these two LED rings is on and the other ring is off. When the inner LED ring is turned on, light beams are reflected from the pupil to the camera, and eye is brighter than other parts of face, but when the outer LED ring is turned on, pupil is dark. These two different eye states are called bright pupil and dark pupil respectively (Figure 3). Synchronization between the camera and LEDs can be controlled using a micro-controller. In this system, bright pupil appears in even frames and dark pupil appears in odd frames.

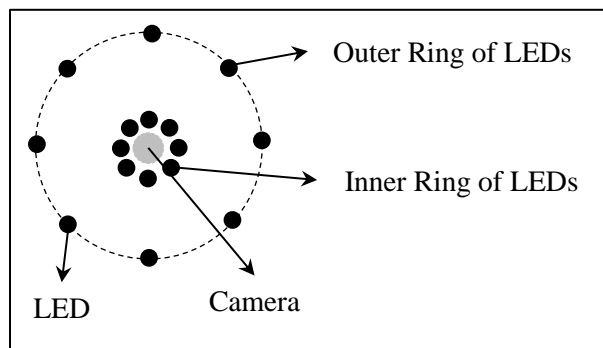


Figure 2. IR lighting system introduced in [6] including two rings of LEDs

In eye detection stage, the property of dark and bright pupil was used. At first, the difference between even and odd frames was obtained. Then, the pupils were detected by applying a threshold. To ensure that the detected points are the center of eyes, size, shape and relative distance of points from each other were evaluated. This evaluation prevents to detect other bright points (such as reflection of glasses) as eye. The proposed method is very efficient for detection of eyes with or without glasses or even with sunglasses.

Zhu *et al.* [34] proposed a similar method for eye detection, except that after initial eye detection, used Support Vector Machine (SVM) for increasing accuracy of eye detection. In this paper, various SVM kernels are investigated and it is shown that Gaussian kernel has the best accuracy. In [35, 36, 37], similar methods were also used for eye detection.

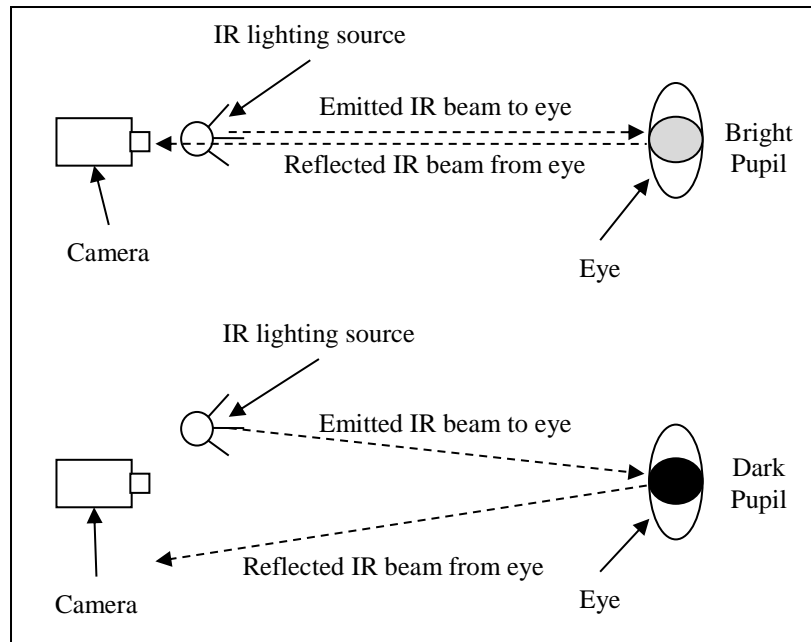


Figure 3. The bright and dark pupil in the IR lighting system introduced in [6]

Zhao *et al.* [38] used only one IR light source for eye detection which was located along the optical axis of the camera. According to position of the light source with respect to driver, pupil is always seen brightly in the image. Therefore, eye locations were detected using opening morphological operation and subtracting the resulted image from the original image. Then, some candidate points were extracted by applying a threshold on difference image and then using connected component analysis. In addition, SVM and Generalized Symmetry Transformation (GST) were used to increase the accuracy of eye detection. The eye locations were determined by combining the results of SVM and GST.

Methods based on lighting in IR spectrum are able to detect eyes without face detection, and therefore these methods have less computational complexity than other methods. These methods are more suitable for the systems which can detect the driver fatigue and distraction only based on the processing of eye region. The main disadvantage of this method is that if eyes are closed, system is not able to detect eyes.

4.2. Feature-based Methods

4.2.1. Methods based on Binarization

Image binarization is one of the simplest methods in image segmentation which can be used in eye detection. Usually more complicated processing is needed to detect the proper location of eye, because results of binarization usually have high error rate.

In [7], with assumption that eye is the darkest points in face, eye location was determined. For this purpose, after binarization of face image, large contours are detected. The first central moment of two largest contours of face image is determined as eye center.

Smith *et al.* [39, 40] obtained binary image of face region after face detection based on the skin color. Binarization of face region was performed based on the skin color and causes eyes

appear black, while other parts of the face appear white. Then, the connected component analysis was used to increase the accuracy of eye detection.

Roshani Tabrizi *et al.* [41] proposed an eye detection algorithm that detects eye in the HSV color space. They also proposed another method in [42] that is based on the method proposed in [43] for the precise eye detection. In this method, a linear transform was applied on image in YCbCr color space, and then the transformed image was converted to binary image. This method has a very good accuracy for eye detection in color images, but it fails when illumination of environment is low.

4.2.2.Methods based on Projection

As previously mentioned, projection is one of the simplest methods for face detection that can also be used for eye detection. In this method we assume eye is darker than skin surface. Therefore, horizontal projection of face image in eye location has smaller values. In [16, 26, 27, 45, 51], projection is the main method for eye detection and in [21, 46], projection was used to determine initial location of eye. To increase the accuracy of projection based methods for eye detection, detection is usually performed with one or more prior assumptions. For example, because eye is always in the upper half of face, projection for the bottom half of face is not calculated and beard or mustache will not reduce the accuracy of eye detection. Furthermore, usually after eye detection, some methods such as connected component analysis and checking the geometric relations between two detected areas were performed to increase the accuracy of eye detection [22].

Although projection can accurately detect eye in simple conditions, but projection is not efficient for black people or those who have eyeglasses. In [24], projection was not applied directly on image (either gray-level or binary image), but projection of the edge detected image was used. This method is less sensitive to face color, but it fails for detection of eyes with sunglasses.

Another method which is based on the projection and is insensitive to face color is variance projection. In this method, variance of pixels in a given direction is calculated instead of calculating projection of pixels in image. Due to the large difference between the values of pixels corresponding to pupil (dark) and eye (white), variance projection can be used for eye detection. In [21], this method was used for eye detection.

4.3.Other Methods

There are a few methods for eye detection based on other approaches that have been used in driver face monitoring systems. These methods are usually time-consuming; therefore they are not used in real-time systems.

Batista [21] detected eye region based on a face model. For this purpose, the estimated eyebrow area was separated from face image and was processed with Sobel edge detection operator. Then, the eyebrows area was detected by calculating projection of edges of image. Because eyes region is always under eyebrows area, eyes region was detected. It was assumed that pupil is the darkest area in eye. With this assumption, after intensity improvement, image binarization was performed and after edge detection, pupil was detected.

In [44], two different methods were proposed for eye detection in day and night. The proposed method for eye detection in day is based on searching an elliptical gray-level model in the top half of face region. In night mode, an IR lighting and imaging system was used which can detect pupil directly. It seems that the combination of these two methods make the system robust and efficient.

5. Detection of Other Components of the Face

In some driver face monitoring systems, other component of driver face such as mouth or nose were detected. However there are a few researches in this field, the methods presented for detection of other components of driver face is reviewed in this section.

5.1. Mouth Detection

In some driver hypo-vigilance detection systems such as [21, 28, 39, 40, 77], open or closed mouth was used as a measure of driver fatigue. Most of these systems detect the mouth based on red color features of lips. The most important disadvantage of these methods is that they can only work properly with color image in visible spectrum and suitable light conditions.

5.2. Nose Detection

Nose tip location with respect to head and eyes is an appropriate criterion for determination of the direction of the head. Three-dimensional model of face can be estimated by only one camera and using location of nose tip with respect to eyes. After estimating 3D model of face, determination of head direction is easily possible. Bergasa *et al.* [31, 32, 33], extracted the nostrils for determination of head direction. In this method, camera was located slightly lower than driver face while nostrils are often visible. The most important property of nostrils is blackness with respect to face skin. Thus, by applying a threshold on nose area, nostrils are detected. This method is failed for nostrils detection of black people and people who have mustache.

5.3. Salient Points Detection

In some recent researches such as [73, 74, 75], the salient points of face are detected after face detection. In these researches, the salient points are tracked over time. Thus, according to the relative position of salient points, hypo-vigilance symptoms are extracted. The most common method which is used for salient point detection is Shi-Tomasi method [76]. This method detects salient corners of a gray-level image as salient points. Tracking and analysis of salient points of face make the system more robust against occlusion.

6. Tracking of Face and Its Components

Usually, entire image and entire face region is considered for detection of face and eye respectively. Searching entire image for object detection for all frames of a sequence increases the computational complexity of the system. Therefore, usually after detection of the face or eyes at the first frame, face or eye tracking is performed in the next frames. With assumption of limited movement of face in two consecutive frames, tracking searches only a limited area of the image instead of the entire image. Thus, tracking usually increase the performance of system.

At first, an appropriate definition for tracking seems to be necessary. One definition for tracking was proposed as: tracking is the following of moving parts of an image in a sequence of frames [47]. Another definition is estimation of motion path of a moving object in successive frames and determination of the object boundary during movement [48].

Two main steps in tracking are motion and matching. Motion is estimation of object position in the next frame. Then, matching phase should be performed in the next step. Matching is determination of actual location of the target in the frame based on the estimated location [47].

The main challenges of object tracking are [47, 48, 49]:

- Some information is omitted due to the mapping of objects from three-dimensional to two-dimensional space
- Target objects may have complex shape or motion.
- Occluding part of the target by other objects.
- Ambient light changing.
- Tracking objects in real-time.

Almost in all methods of tracking, some limitations on target object tracking are assumed for performing more accurately and quickly tracking. One of the most common restrictions is smooth motion of target object, so it is assumed that the object will never move suddenly. Other limitations which are assumed for object tracking includes constant speed or acceleration of the object motion and the constant geometry or appearance of the target object during tracking [48, 49].

6.1. Search Window

The simplest method for motion estimation is based on defining a search window around the current position of target object. The larger search window size results in more accurately estimation. Increasing the window size increases the computational complexity and reduces the processing speed of tracking. Therefore, window size should be selected appropriately with respect to size and motion of the object [47].

In [14], search window was used for eye tracking. This system was developed on a PC104+ and could achieve a processing speed of 10 fps. The main disadvantage of this system is that in the cases of rapid movement of head, tracking is disrupted and detection should be done again manually.

Rang-ben *et al.* [50] used a template matching method based on Hausdroff distance for eye tracking. Accordingly, the similarity between template and search window was calculated, and the point with the lowest distance was determined as eye location. Tracking speed of 10 fps was reported using this method.

In [17, 22], eye tracking was performed using template matching. In this method, the driver eye template in both open and closed status for the left and right eyes was determined by subtracting two images during blinking. Then, entire face region was searched using correlation method for eye detection. Correlation method is independent to intensity changes of input image or template image.

In [24], detected eye region was chosen as eye template to use in the next stage of eye tracking. Eye tracking was done based on calculation of Sum of Absolute Difference (SAD). In this method, size of search area is 10 pixels larger than size of the eye template.

6.2. Adaptive filters

Search window is a very simple method that has large computational complexity and low accuracy rate. Another solution for object tracking is based on adaptive filters. The first type of adaptive filters is Kalman filter (KF) which is firstly introduced in 1960 by Kalman [52]. Later, other methods such as Extended Kalman Filter (EKF) and Particle Filter (PF) were presented which overcome some problems of simple KF.

6.2.1. Kalman Filter

Kalman filter as one of the most well-known methods of object tracking is based on control theory. This method is able to predict the status of a linear system with Gaussian distribution using a recursive algorithm. Kalman filter determines the object position in the next frame based on the motion type of the moving object in the previous frames so that the presence probability of the target object at that point is maximum [47, 48].

The basic Kalman filter was designed for the state estimation in the linear systems, but other types of the Kalman filter such as extended Kalman filter (EKF) [52, 53] and unscented Kalman filter (UKF) [54] were proposed for nonlinear systems. The advantage of the UKF over the EKF is that UKF has lower computational complexity and has more accuracy in tracking phase [54].

In the method proposed by Ji et. al. [6], eye tracking was performed using Kalman filter on a Sun Ultra 30 (300 MHz) computer with a speed of 20 frames per second.

In [46], face tracking was done using Kalman filter based on a number of scale independent features. This method assumes that face components are on a plane, and the changes of face image in each frame can be estimated using an affine transform.

Because Kalman filter is able to track only one moving object, in [31, 32, 33] two Kalman filters were used for tracking two eyes. In the proposed method, each eye was tracked independently using a separate Kalman filter.

One way to determine head direction is based on face pose tracking. In [21, 37], it was assumed that face components are on a plane. At first, face must be placed directly in front of camera for system initialization. The system detects frontal face pose based on distance between two eyes and then started face tracking.

Zhang *et al.* [26, 27] used UKF filter for eye tracking. The main advantage of UKF filter is accurate estimation of eye location with low computational complexity. Tracking error of this method was reported as 0.5%.

6.2.2. Particle Filter

Particle filter is similar to the Kalman filter for predicting the system status. Particle filter is not only able to predict a non-linear system, but also it is able to predict the next status of the system with any arbitrary distribution. The arbitrary distribution is estimated based on previous distribution of particles. Particle filter is able to track multiple moving objects simultaneously [55].

One of the important problems of particle filter is that the number of particles required for estimating the system status grows exponentially by increasing the system dimension. Thus, many particles are required to estimate face motion which includes a large number of features. McCall *et al.* [56] used a Hierarchical Particle Filter (HPF) for tracking of face and its components. In this method, facial image was divided into smaller areas by a Dynamic Bayesian Network (DBN) in a hierarchical structure, and then each area was tracked using a particle filter. In this tracking method, Haar-like features and facial template were used as local and global features respectively.

6.3. Other Tracking Methods

Sometimes, a combination of different tracking methods is used hierarchically to increase tracking accuracy. In [35, 36, 37], a hybrid method was proposed for pupil tracking. In these methods, tracking was performed based on Kalman filtering at first. If Kalman filter could not calculate a good estimation for eye location, tracking was performed using mean shift method.

In another research, Smith *et al.* [39, 40] used a hybrid of four tracking methods hierarchically for lip tracking. In this method, if tracking error is larger than a threshold, the other methods will be used.

In [44], face tracking method was developed using condensation algorithm in conjunction with neural networks, but eye tracking method was based on search window and template matching. Condensation algorithm was proposed by Isard *et al.* [71] to track active contours using stochastic approach.

In [75], Lucas–Kanade method was used for computation of optical flow and tracking of salient points. Lucas–Kanade method is a method for optical flow estimation which assumes that the optical flow is constant in a local neighborhood of the pixel. This method was used in [75] for tracking of each salient points independently.

7. Symptom Extraction Related to Fatigue and Distraction

In driver face monitoring systems, useful symptoms for fatigue and distraction detection can be divided into four general categories:

- Symptoms related to eye region
- Symptoms related to mouth region
- Symptoms related to head
- Symptoms related to face

These symptoms appear in almost all persons, however quality and pattern of these symptoms is different in each person. Moreover, pattern of a symptom may change in a period of time [57]. Therefore, these challenges have to be considered for symptom extraction.

7.1.Symptoms Related to Eye Region

Eye is the most important area where the symptoms of drowsiness and distraction appear in it. Therefore, many of driver face monitoring systems detect driver fatigue and distraction only based on the symptoms extracted from eyes. In this section, the symptoms extracted from eye is explained.

7.1.1.Eye Closure

The simplest symptom to detect drowsiness is eye closure. Despite the simplicity of this symptom, it provides very useful information about drowsiness and even driver anesthesia. Eye closure can be used in two different forms for detection of driver drowsiness: (1) continuous eye closure and (2) percentage of the eye closure in a certain period of time.

In continuous eye closure method, if eyes are closed continuously for a certain period of time, drowsiness will be detected. This method is not efficient and not very reliable because driver drowsiness will be detected only when his/her eyes are closed completely for a certain period of time. In this case, an accident may occur during this time. This symptom was used in [15, 17, 26, 27, 29]. Continuous eye closure is usually associated with some other symptoms for increasing system accuracy.

Another method is based on percentage of eye closure (PERCLOS). PERCLOS was used as an important symptom for detection of drowsiness in [6, 9, 21, 37, 40, 44, 58, 59].

In [21, 45], closed and open eyes were detected based on calculation of image projection of eye region. This method is relatively simple and has good accuracy.

In [16, 50], a method was presented which detect closed and open eyes based on edges of eyelids and eye corners. Computational complexity of edge detection is slightly more than projection but it seems that its accuracy is not much different.

In [7, 46], optical flow was used for detection of eye closure. In these methods, opening and closing state of eyelids was detected based on calculation of optical flow.

In [14, 31, 32, 33], eye location was detected based on lighting and imaging in the IR spectrum considering the property of reflected IR beam from pupil. Closed or open eye state was detected using the same way. Then PERCLOS was calculated and was used for driver drowsiness detection.

In [44], Gabor features and SVM classifier were used to determine eye status. This method determined whether eye is open or closed only in day light condition. For night, another method based on lighting and imaging in the IR spectrum was used.

Hamada *et al.* [30] investigated relation between time of eye closure during blinking and drowsiness. In this study, they tried to represent an adaptive algorithm to detect eye closure and drowsiness based on period of eye closure. After eye detection, eye region is divided into several columns. The first derivative of gray level of each column is calculated and the maximum and minimum values of derivation are considered as parts of the upper and lower eyelids. The average distance between the upper and lower eyelids can be considered as a measurement of eye closure. By determining distance between eyelids, waveform of distance between two eyelids during time can be obtained. Eye closure time is determined using intersection of this waveform and the second derivation of the waveform. Average accuracy of this method for detection of closed eyes is approximately 97%, while average accuracy of closed eye detection using a fixed threshold is about 83%. After detection of closed and open eye, time distribution of eye closure was plotted that is similar to a Gaussian distribution (Figure 4). Due to differences between people in blinking pattern, average and standard deviation of this distribution is different for different people. Therefore, an adaptive threshold should be used for detection of drowsiness in different people. The adaptive threshold which is shown by T_s was obtained according to equation (5).

$$T_s = T_c + 0.7A \tag{5}$$

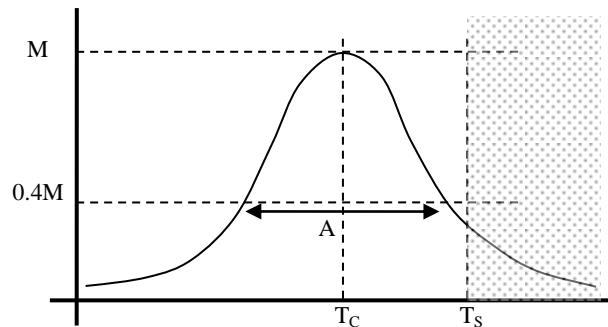


Figure 4. The algorithm proposed in [30] for determination of an adaptive threshold based on time distribution of eye closure

7.1.2.Distance between the Eyelids

In [22, 51, 59], eyelid distance was used for drowsiness detection. Distance between eyelids is decreased when driver is fatigued. In this method, eyelid distance is calculated based on the horizontal projection of eye region.

7.1.3. Eye Blink Speed

Another symptom of fatigue is slow eye blink speed. Eye blink speed is the time between opening and closing eyelids during one blink event. In [6, 21, 37], this symptom was used to detect fatigue. If speed of blinking is larger than a threshold (about 0.5 to 0.8 seconds) driver drowsiness is detected.

7.1.4. Eye Blink Rate

Eye blink rate is the number of blinking event in a certain period of time. This symptom can be used to detect fatigue and distraction. Eye blink rate varies in different people. If eye blink rate is much less than the normal state, it shows distraction and if it is much higher than the normal state, it shows driver drowsiness. This symptom was used in [6, 31, 32, 33, 40, 59] for fatigue and distraction detection.

Another symptom for driver fatigue detection that was introduced in [60] is the rate of sequential blinking. In this method, the number of sequential blinking with very close intervals is counted and then sequential eye blink rate will be used to measure fatigue.

7.1.5. Gaze Direction

Gaze direction is an effective symptom to determine driver distraction. Using this symptom, it is possible to determine lack of driver attention to road. Moreover, driver decision to overtake or lane changing can also be predicted using this symptom. Based on previous researches, it is shown that driver changes his/her looking permanently from the current lane to the destination lane and vice versa during overtaking [61, 62].

In [6], gaze direction was estimated based on center of pupil and the reflected light from cornea. The light reflected from cornea is called glint which is weaker than the reflected light from pupil and its diameter is smaller. The reflected light from pupil is only appear in the even frames, while the reflected light from the cornea is appear in all frames. Direction of the line which connects center of pupil to center of glint determines the gaze.

The method proposed in [37] is very similar to the aforementioned method for detection of the gaze direction, but a Generalized Regression Neural Network (GRNN) was used to detect gaze direction. Accuracy of this method for detection of gaze direction is about 96%.

7.1.6. Eye Saccadic Movements

One of the symptoms which can be extracted from gaze direction is the rate of eye saccadic and irregular movements. When driver has normal consciousness, he/she often has a quick review of surroundings. This quick review is occurred with irregular and rapid saccadic movements. Amount of eye saccadic movements should be in a certain field of view. Excessive eye saccadic movement is a symptom of distraction while reduction of this feature is a symptom of fatigue. Ji *et al.* [37] used this symptom as *eye saccadic movements* and Bergasa *et al.* [31, 32, 33] used the same symptom as the *fixed gaze* to determine level of driver consciousness.

7.2. Symptoms Related to Mouth Region

Yawning is a symptom of fatigue that is occurred with mouth opening. Driver speaking while driving can be considered as a symptom of distraction. Therefore, the symptoms related to mouth region can be used to detect driver drowsiness and distraction.

In [28, 77], open mouth was detected based on the ratio of width to height of mouth. If mouth is closed, this ratio is low and it is higher when mouth is open.

Ji *et al.* [37] computed Gabor wavelet coefficients at 8 fixed points around mouth for detection of open mouth. Gabor wavelet coefficients have different values in a closed mouth state compared to open mouth state. Thus, by comparing these coefficients, driver yawning was recognized. Number of yawning in a certain period of time was used as a symptom for drowsiness detection.

The main disadvantage of aforementioned methods is that the system recognizes driver yawning wrongly when mouth is open for whatever reason (*e.g.*, talking). The main reason of such error is that the selected model for yawning is inappropriate. Yawning includes spatio-temporal information. Therefore, a spatio-temporal model should be used for yawning. Li [63] used a Hidden Markov Model (HMM) to model yawning. In this method, circular HMM with four states was used as spatio-temporal structure for yawning modeling.

7.3.Symptoms Related to Head

7.3.1.Head Nodding

One of symptoms related to drowsiness is head nodding resulted from dozing. During drowsiness, head is gradually bent. This symptom is one of the symptoms related to drowsiness or anesthesia. The systems presented in [7, 15, 31, 32, 33, 37] used these symptom to detect drowsiness.

7.3.2.Head Orientation

Determination of head orientation is an effective symptom for detection of driver distraction and prediction of driver intent to lane change. In general, a three-dimensional model of head is required for determination of head orientation.

Ji *et al.* [6] used one camera to determine head orientation. For this purpose, three-dimensional model of head was extracted using pupil properties. The pupil properties including shape, size and its gray level and the distance between two pupils will change when head is rotated. Therefore, direction of head rotation was determined with an error of 15 degrees.

In [21, 40, 44], an oval face model was defined which there are some important facial points inside the model. In this model, after face detection, locations of the facial points are determined. Face rotation results in changing distances between these points. Accordingly, angle and orientation of head rotation are determined and the rate of head rotation was calculated in last 60 seconds. This symptom was used to determine the driver distraction.

In [31, 32, 33, 37, 64], a three-dimensional model of head was used to determine head orientation. In this model, it was assumed that face components are on a plane. Thus, using the distance between two eyes and also the distance between eyes and nose tip, head orientation in 9 directions was determined: straight, top, bottom, left, right, top right, top left, bottom right, bottom left.

Facial tracking is one of the methods to determine head orientation. In [37], assuming that face components are on a plane, head orientation was computed using Kalman filter and location of one eye.

7.3.3.Fixed Head

Driver fixed head over a long period of time is a symptom of driver distraction. When driver focuses on something except driving task in his/her mind, his/her head will be fixed. This symptom was used in [7] as one of the symptoms for driver distraction detection.

7.4. Symptoms Related to Face

Facial expression results from emotions, physiology activities, verbal and non-verbal communications. Changes in facial muscles results in changes in face image features in a period of time. Face features is extracted and processed using a spatio-temporal algorithms and data structures. Some methods for feelings and emotions recognition based on facial expression have been performed till now. Facial expressions are usually divided into several basic expressions: normal, happy, sad, surprised, angry, but one or more specific expressions can be defined for a given application, *e.g.*, detection of unexpected change in human physical condition by facial expression monitoring [65]. Driver feeling recognition can be used for determination of the driver distraction.

In [65], a method for unusual facial expression detection was proposed using Gabor features and SVM. This method monitored human status and detects unexpected change in human physical condition *e.g.* heart attack.

Gu *et al.* [66] determined the amount of driver consciousness based on his/her napping, yawning and neglecting. In this method, napping, yawning and neglecting as facial expressions are extracted using mid-level image features. The mid-level features including eye movement, head movement and changes of top half and bottom half of the face. The mid-level features was extracted from low-level features related to pupil, three-dimensional face position, and tracking of facial feature points. Combination of the extracted low-level features for determination of the facial expressions has performed using Dynamic Bayesian Network (DBN). DBN provides a spatio-temporal model for determination of the facial expressions. In DBN, face pose changes are modeled with changing the network status from one state to another one. In this model, facial expressions determine by sequence of several status changes.

In [67], a system was proposed to recognize key facial feature points at varying pre-accident intervals and use them to predict minor and major accidents. This system used vehicle status and driver face images simultaneously and combined them to produce an accident prediction system.

8. Fatigue and Distraction Detection

After symptom extraction, detection of driver fatigue and distraction based on extracted symptoms are discussed. Determination of driver state is considered as a classification problem. The simplest form of this classification consists of two classes: consciousness and unconsciousness. Unconsciousness class can include fatigue and distraction states. At a higher stage, classification may consist of three classes: consciousness, drowsiness and distraction. Similarly it is possible to increase the number of recognizable states. In the highest level of classification, system can evaluate driver drowsiness or distraction by assigning a number in range of [0,1].

8.1. Methods based on Threshold

The simplest method for driver fatigue or distraction detection is based on applying a threshold on each extracted symptom. In the systems presented in [41, 42], driver drowsiness was detected by applying a constant threshold on PERCLOS. In [68], the first stage was driver face identification and then an appropriate threshold was chosen for the system based on physical and psychological characteristics of the identified driver. This method for determination of the adaptive threshold is more efficient for hypo-vigilance detection.

8.2. Knowledge-based Approaches

In a knowledge-based approach, decision-making about driver fatigue and distraction is based on knowledge of an expert. In these methods, knowledge usually appears in form of *if-then* rules. The most common of these methods are fuzzy expert systems. In researches such as [31, 32, 33, 59], knowledge of an expert was used in form of fuzzy rules for determination of driver fatigue by a value in range of [0,1].

In [40], a Finite State Machine (FSM) was proposed for hypo-vigilance detection. Structures and relations between states in the proposed FSM were defined by an expert.

8.3. Methods based on Probability Theory

Driver fatigue and distraction detection based on extracted symptoms is a difficult problem. In fact, extracted symptoms from driver face have uncertainty. In addition, the relations between extracted features and level of consciousness have uncertainty too. One of the methods for representation of uncertainty is based on probability theory. The methods based on probability theory are able to determine probability of the driver fatigue or distraction by a value in the range of [0,1].

8.3.1. Bayesian Network

In [6, 37, 66], a Bayesian network was used for driver fatigue detection. It is possible to represent an uncertain knowledge in form of an acyclic directed graph using Bayesian networks. Bayesian network is able to predict and inference future states based on past states and current observed information.

In [66], facial expressions including yawning, dozing and neglecting were recognized based on mid and low-level features using DBN and then level of driver consciousness was determined. In this method, DBN was used for presenting a spatio-temporal model for the recognition of facial expressions and moreover it provides a tool for determination of level of consciousness.

Bayesian network has a large computational complexity. In some problems, naive Bayesian network is used when we can assume that observations are independent. As computational complexity and required initial information in naive Bayesian network are lower, processing speed of this model is higher. In [56], a naive DBN was used for facial expression recognition.

8.3.2. Dempster-Shafer Theory

Prior probability and conditional probability of classes must be determined before using Bayesian network. Determining prior probability and conditional probability of different classes are difficult and sometimes impossible in real problems. Therefore, instead of using Bayesian networks, Dempster-Shafer theory can be used. Dempster-Shafer theory is a combination of probability theory and fuzzy theory. The main disadvantage of Dempster-Shafer theory is that computational complexity of this model is high.

Yang *et al.* [4] used a combination of several symptoms for driver fatigue detection. In this method, a general model was defined for the system, and then prior probability of each symptom was determined based on some heuristic rules. After initial decision-making based on different symptoms, the final decision about the driver consciousness was made using Dempster-Shafer theory.

8.4. Statistical Methods

Neural networks (NN) are one of the main methods in statistical pattern recognition. Eskandarian *et al.* [69, 70] detected driver drowsiness based on combination of two categories of symptom using NN. One symptom is based on steering motion and another one is based on driver face monitoring. In this system, drowsiness level was determined by a combinational NN. This NN consists of two sections: unsupervised and supervised. The unsupervised section includes input layer and hidden layer which clusters input data. The supervised section only contains output layer which performs classification process. The proposed NN is simpler than supervised NNs and more accurate than unsupervised NNs.

9. Summary

We reviewed driver face monitoring systems from different viewpoint in the previous sections. In this section, a coarse comparison of some of the reviewed systems is presented from the discussed viewpoints. Also, we explain the critical problems in the current systems.

Table 3 shows a comparison of the driver face monitoring systems. In Table 3, different parts of the reported systems are explained briefly: detection of face, eye and other components of face, tracking method, symptom extraction and fatigue/distraction detection. Finally, robustness of mentioned systems is compared subjectively.

According to results shown in Table 3, however the method for fatigue/distraction detection is effective to achieve a robust system, but current driver face monitoring systems suffer from two main problems: (1) robust and precise detection and tracking of face and facial components, (2) robust and precise symptom extraction. These challenges are:

- Developing illumination invariant algorithms for detection and tracking of face and facial components
- Developing algorithms for detection and tracking of face and facial components with robustness against skin color and partial occlusion (*e.g.*, glasses, sun-glasses and whisker)
- Developing robust algorithms for driver face tracking with respect to head rotations in different directions
- Developing algorithms for symptom extraction with robustness against different races (*e.g.*, eyelid distance of East Asian people is low on normal state)
- Developing algorithms for symptom extraction from eye region with robustness against wearing glasses and sun-glasses
- Developing color invariant algorithms for symptom extraction from mouth region with robustness against skin-color and occlusion (*e.g.*, whisker)
- Fast processing to achieve real-time systems

Table 3. Summary of the driver face monitoring systems for detection and tracking of face, eye and other components of face and comparison between them

Ref.	Face Detection	Eye Detection	Detection of Other Components	Tracking Method	Symptom Extraction	Fatigue or Distraction Detection	Robustness
[6]		Imaging in the IR Spectrum		Adaptive filters (Kalman filter)	PERCLOS Eye blink speed Gaze direction Head rotation	Probability theory (Bayesian network)	Very Good
[33]		Imaging in the IR Spectrum	Nose detection (using dark color of nostrils and face model)	Adaptive filters (Kalman filter)	PERCLOS Eye blink rate Eye saccadic movement Head nodding Head orientation	Knowledge-based (Fuzzy expert system)	Very Good
[37]		Imaging in the IR Spectrum and verification by SVM	Mouth detection (edge based) Nose detection (edge based)	Other methods (combination of Kalman filter and mean shift)	PERCLOS Eye blink speed Gaze direction Eye saccadic movement Yawning Head nodding Head orientation	Probability theory (Bayesian network)	Very Good
[44]	Learning-based (using Viola-Jones method)	Other methods (elliptical model in daylight and IR imaging in nightlight)		Other methods (condensation algorithm in conjunction with neural networks)	PERCLOS Eye blink rate Head orientation	Thresholding	Good
[51]	Learning-based (using Viola-Jones method)	Explicitly by Feature-based (projection)		Search window (based on face template matching)	PERCLOS Distance between eyelids	Thresholding	Good
[59]	Learning-based (using Viola-Jones method)	Explicitly by Feature-based (projection)		Search window (based on face template matching)	PERCLOS Distance between eyelids Eye blink rate Head orientation	Knowledge-based (Fuzzy expert System)	Good
[27]	Learning-based (using Viola-Jones method)	Feature-based (projection)		Adaptive filters (UKF)	Continuous eye closure	Thresholding	Average
[40]	Feature-based (in an unknown color space)	Feature-based (binarization)	Mouth detection (color based)	Other methods (combination of 4 hierarchical tracking method)	PERCLOS Eye blink rate Gaze direction Yawning Head orientation	Knowledge-based (Finite State Machine)	Average
[46]		Feature-based (projection)		Adaptive filters (Kalman filter)	Eye blink rate		Poor
[21]	Feature-based (in RGB color space)	Feature-based (variance projection and face model)	Mouth detection (color based)	Adaptive filters (Kalman filter)	PERCLOS Eye blink speed Head rotation		Poor
[22]	Feature-based (in YCbCr color space)	Feature-based (projection and connected component analysis)		Search window (eye template matching)	Eyelid distance	Thresholding	Very Poor

It seems that the best current systems are working in IR spectrum which combats to some above challenges. For example, the system proposed in [6, 33, 37] are almost illumination invariant and robust against wearing glasses.

10. Conclusion and Future Works

According to the statistics, 20% of the crashes are due to the driver fatigue and distraction. Therefore, design and development of systems for driver fatigue and distraction detection is not only reasonable but also it seems very necessary. One of the approaches for fatigue and distraction detection is driver face monitoring. Commercial use of these systems has not been common yet, but the volume of researches which is done or supported by big automobile companies in this field shows that the driver face monitoring system will appear soon in cars and heavy vehicles.

The current research to produce driver face monitoring systems is successful, but it seems more researches are necessary to develop commercial systems with enough accuracy. Using IR spectrum for lighting and imaging is common. Despite of lighting in visible spectrum, IR lighting is non-intrusive for human vision system. A lot of current methods are just based on processing of eye region, because of real-time processing. Only some methods were presented for processing of other face components, which usually have higher computational complexity.

In the methods based on the processing of eye region, symptoms such as eye closure, percentage of eyelid closure over time (PERCLOS), eyelid distance, eye blink rate and gaze direction are used for fatigue and distraction detection. In the methods based on the processing of the face and head region, not only the eye symptoms are used but also some symptoms such as yawning, head nodding (resulted from dozing), head orientation and facial expression are also used. However using more symptoms for fatigue and distraction detection may increase the accuracy of the system, it causes more computational complexity.

There are many researches for face and eye detection. The algorithms that use color model for detection are unused for the systems that are based on the IR imaging. Among other methods, the methods that can be implemented as real-time or near real-time are only appropriate for real systems. Moreover, face detection algorithm should be able to detect face in different light conditions (day and night), environments with sudden light changes (entrance to or exit from a tunnel) and different skin colors (especially for black skins). Current real-time face monitoring systems have not been able to combat some of above problems yet. Usually accurate and robust methods have high computational complexity. On the other hand, simple methods with low computational complexity have low accuracy in real conditions. Therefore, in driver face monitoring systems there is always a tradeoff between accuracy and processing speed.

In a lot of current systems, there are still some problems in the driver eye detection. The main problems are related to detection of eye for black skin people and the driver who has glasses (especially sunglasses). The best current methods to combat such problems are based on IR imaging.

In driver face monitoring systems, in order to decrease computational complexity, tracking methods are used instead of face and eye detection in each frame. In other words, face or eye is detected only for one frame and in the next frames, the system will track the face or eye, except in the case of tracking failure. The main methods used for tracking are template matching in the search window and Kalman filter. Tracking method usually fails in the case of fast movements of head or out-of-plane head rotation. It seems the most current systems

need a more robust tracking method to decrease the errors in symptom extraction phase. The best current face tracking systems such as [37, 46, 56] divide the face to different components and track each component with respect to an overall tracking.

Briefly, it seems that the most challenging problems of driver face monitoring systems are related to face and eye detection, tracking methods and symptom extraction, especially in night light. If scientists improve these parts of driver face monitoring systems, fatigue and distraction detection will be easily performed by current classification methods. In other words, current driver face monitoring systems work well in laboratory conditions, but they usually fail in night light and sudden light changes.

In different approaches for fatigue and distraction detection, the methods based on steering motion are the main rival of the driver face monitoring systems. The methods based on the steering motion is older than driver face monitoring systems, therefore they are more accurate. But their main drawback is that the system is not able to detect the driver drowsiness and distraction till it happened. For example the driver may drowse without any symptoms of unbalanced controlling of the car; therefore the system will not detect the driver drowsiness. The driver face monitoring systems can make an alarm as soon as the symptoms of fatigue or distraction in the driver face are occurred. Future driver face monitoring systems can recognize driver feeling and emotion by extracting the driver facial expression. Thus, such system will alarm the driver if his/her mental state is not good. Also, next driver face monitoring systems can detect unusual and abnormal state of driver such as heart attack or epilepsy and send S.O.S message to police.

References

- [1] Fatigue, SafetyNet 2009, European Commission Road Safety, http://ec.europa.eu/transport/road_safety/specialist/knowledge/pdf/fatigue.pdf, (2009).
- [2] N. L. Haworth, T. J. Triggs and E. M. Grey, "Driver Fatigue: Concepts, Measurement and Crash Countermeasures", Human Factors Group, Department of Psychology, Monash University, (1988) June.
- [3] C. T. Lin, L. W. Ko, I. F. Chung, T. Y. Huang, Y. C. Chen, T. P. Jung and S. F. Liang, "Adaptive EEG-Based Alertness Estimation System by Using ICA-Based Fuzzy Neural Networks", IEEE Transactions on Circuits and Systems, vol. 53, no. 11, (2006).
- [4] G. Yang, Y. Lin and P. Bhattacharya, "A Driver Fatigue Recognition Model Using Fusion of Multiple Features", Proceedings of IEEE International Conference on Systems, Man and Cybernetics (SMC), (2005) October, Hawaii, USA.
- [5] C. Papadelis, C. Kourtidou-Papadeli, P. D. Bamidis, I. Chouvarda, D. Koufogiannis, E. Bekiaris and N. Maglaveras, "Indicators of Sleepiness in an Ambulatory EEG Study of Night Driving", Proceedings of 28th IEEE Annual International Conference of the Engineering in Medicine and Biology Society (EMBS), (2006) September, New York, USA.
- [6] Q. Ji and X. Yang, "Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance", Real-Time Imaging, vol. 8, (2002).
- [7] T. Brandt, R. Stemmer, B. Mertsching and A. Rakotonirainy, "Affordable Visual Driver Monitoring System for Fatigue and Monotony", Proceedings of IEEE International Conference on Systems, Man and Cybernetics, (2004) October, Hague, Netherlands.
- [8] P. S. Rau, "Drowsy Driver Detection and Warning System for Commercial Vehicle Drivers: Field Operational Test Design, Data Analyses and Progress", National Highway Traffic Safety Administration of USA (NHTSA), (2005).
- [9] T. Von Jan, T. Karnahl, K. Seifert, J. Hilgenstock and R. Zobel, "Don't Sleep and Drive-VW's Fatigue Detection Technology", Cen. for Automotive Safety Res., Adelaide University, Australia, (2005).
- [10] S. Boverie, D. Daurenjou, D. Esteve, H. Poulard and J. Thomas, "Driver Vigilance Monitoring - New Developments", Proceedings of 15th IFAC World Congress on Automatic Control, (2002), Barcelona, Spain.
- [11] M. Bayly, B. Fildes, M. Regan and K. Young, "Review of Crash Effectiveness of Intelligent Transport Systems", TRAFFIC Accident Causation in Europe (TRACE), (2007).
- [12] H. Cai and Y. Lin, "An Experiment to Non-Intrusively Collect Physiological Parameters Towards Driver State Detection", Proceedings of SAE 2007 World Congress, (2007) April, Detroit, Michigan, USA.

- [13] T. Nakagawa, T. Kawachi, S. Arimitsu, M. Kanno, K. Sasaki and H. Hosaka, "Drowsiness Detection Using Spectrum Analysis of Eye Movement and Effective Stimuli to Keep Driver Awake", *DENSO Technical Review*, vol. 12, no. 1, (2006).
- [14] R. Grace, V. E. Byrne, D. M. Bierman, J. M. Legrand, D. Gricourt, R. K. Davis, J. J. Staszewski and B. Carnahan, "A Drowsy Driver Detection System for Heavy Vehicles", *Proceedings of 17th AIAA/IEEE/SAE Digital Avionics Systems Conference (DASC)*, (1998) November, Washington, USA.
- [15] F. Moreno, F. Aparicio, W. Hernandez and J. Paez, "A Low-cost Real-Time FPGA Solution for Driver Drowsiness Detection", *Proceedings of 29th IEEE Annual Conference of the Industrial Electronics Society*, (2003) November, Virginia, USA.
- [16] F. Wang and H. Qin, "A FPGA based Driver Drowsiness Detecting System", *Proceedings of IEEE International Conference on Vehicular Electronics and Safety*, (2005) October, Xian, China.
- [17] H. Veeraraghavan and N. Papanikolopoulos, "Detecting Driver Fatigue Through the Use of Advanced Face Monitoring Techniques", *Intelligent Transportation System Institute, Department of Computer Science and Engineering, University of Minnesota*, (2001).
- [18] AWAKE System, <http://www.awake-eu.org/system.html>.
- [19] M. Montiglio, S. Martini and V. Murdocco, "Development of a Lane Keeping Support System for Heavy-Trucks", *Proceeding of 13th World Congress & Exhibition on Intelligent Transport Systems and Services*, (2006) October, London, United Kingdom.
- [20] M. H. Yang, D. J. Kriegman and N. Ahuja, "Detecting Faces in Images: A Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 1, (2002).
- [21] J. Batista, "A Drowsiness and Point of Attention Monitoring System for Driver Vigilance", *Proceeding of IEEE Intelligent Transportation Systems Conference*, (2007) October, Seattle, USA.
- [22] W. Dong and X. Wu, "Driver Fatigue Detection Based on the Distance of Eyelid", *Proceeding of IEEE International Workshop VLSI Design & Video Technology*, (2005) May, Suzhou, China.
- [23] K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation Facial Feature Extraction and Tracking", *Signal Processing: Image Communication*, vol. 12, no. 3, (1998).
- [24] W. B. Horng, C. Y. Chen, Y. Chang and C. H. Fan, "Driver Fatigue Detection Based on Eye Tracking and Dynamic Template Matching", *Proceeding of IEEE International Conference on Networking, Sensing & Control*, (2004) March, Taipei, Taiwan.
- [25] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", *Proceeding of International Conference on Computer Vision and Pattern Recognition (CVPR)*, (2001), Kauai, HI, USA.
- [26] Z. Zhang and J. Zhang, "A New Real-Time Eye Tracking for Driver Fatigue Detection", *Proceeding of 6th IEEE International Conference on ITS Telecommunications*, (2006). Chengdu, China.
- [27] Z. Zhang and J. Zhang, "Driver Fatigue Detection Based Intelligent Vehicle Control", *Proceeding of 18th International Conference on Pattern Recognition (ICPR)*, (2006) September, Hong Kong, China.
- [28] T. Wang and P. Shi, "Yawning Detection for Determining Driver Drowsiness", *Proceeding of IEEE International Workshop on VLSI Design & Video Technology*, (2005) May, Suzhou, China.
- [29] A. Liu, Z. Li, L. Wang and Y. Zhao, "A Practical Driver Fatigue Detection Algorithm Based on Eye State", *Proceeding of Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics*, (2010) September, Shanghai, China.
- [30] T. Hamada, T. Ito, K. Adachi, T. Nakano and S. Yamamoto, "Detecting Method for Drivers' Drowsiness Applicable to Individual Features", *Proceeding of IEEE Intelligent Transportation Systems*, (2003) October, Shanghai, China.
- [31] L. M. Bergasa, J. Nuevo, M. A. Sotelo and M. Vhquez, "Weal-Time System for Monitoring Driver Vigilance", *Proceeding of IEEE Intelligent Vehicles Symposium*, (2004) June, Parma, Italy.
- [32] L. M. Bergasa and J. Nuevo, "Real-Time System for Monitoring Driver Vigilance", *IEEE International Symposium on Industrial Electronics*, (2005) June, Dubrovnik, Croatia.
- [33] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea and M. E. Lopez, "Real-Time System for Monitoring Driver Vigilance", *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, (2006).
- [34] Z. Zhu, K. Fujimura and Q. Ji, "Real-Time Eye Detection and Tracking Under Various Light Conditions", *ACM Eye Tracking Research & Application symposium*, (2002), New Odeans, Louisiana, USA.
- [35] X. Liu, F. Xu and K. Fujimura, "Real-Time Eye Detection and Tracking for Driver Observation Under Various Light Conditions", *Proceeding of IEEE Intelligent Vehicle Symposium*, (2002) June, Versailles, France.
- [36] Z. Zhu, Q. Ji, K. Fujimura and K. Lee, "Combining Kalman Filtering and Mean Shift for Real Time Eye Tracking Under Active IR Illumination", *Proceeding of 16th International Conference on Pattern Recognition*, (2002) August, Quebec, Canada.
- [37] Q. Ji, Z. Zhu and P. Lan, "Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue", *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, (2004).

- [38] S. Zhao and R. R. Grigat, "Robust Eye Detection under Active Infrared Illumination", Proceeding of 18th IEEE International Conference on Pattern Recognition (ICPR), (2006) September, Hong Kong, China.
- [39] P. Smith, M. Shah and N. V. Lobo, "Monitoring Head/Eye Motion for Driver Alertness with One Camera", Proceeding of 15th IEEE International Conference on Pattern Recognition, (2000), Barcelona, Spain.
- [40] P. Smith, M. Shah and N. V. Lobo, "Determining Driver Visual Attention With One Camera", IEEE Transactions on Intelligent Transportation Systems, vol. 4, no. 4, (2003).
- [41] P. R. Tabrizi and R. A. Zoroofi, "Open/Closed Eye Analysis for Drowsiness Detection", Proceeding of 1st Workshops on Image Processing Theory, Tools and Applications, (2008) November, Sousse, Tunisia.
- [42] P. R. Tabrizi and R. A. Zoroofi, "Drowsiness Detection Based on Brightness and Numeral Features of Eye Image", Proceeding of 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing, (2009) September, Kyoto, Japan.
- [43] R. L. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, (2002).
- [44] M. J. Flores, J. M. Armingol and A. Escalera, "Driver Drowsiness Warning System Using Visual Information for Both Diurnal and Nocturnal Illumination Conditions", EURASIP Journal on Advances in Signal Processing, (2010).
- [45] L. Gan, B. Cui and W. Wang, "Driver Fatigue Detection Based on Eye Tracking", Proceeding of 6th World Congress on Intelligent Control and Automation, (2006) June, Dalian, China.
- [46] M. Lalonde, D. Byrns, L. Gagnon, N. Teasdale and D. Laurendeau, "Real-time Eye Blink Detection with GPU-based SIFT Tracking", Proceeding of 4th Canadian Conference on Computer and Robot Vision, (2007) May, Montreal, Canada.
- [47] E. Trucco and K. Plakas, "Video Tracking: A Concise Survey", IEEE Journal of Oceanic Engineering, vol. 31, no. 2, (2006).
- [48] A. Yilmaz, O. Javed and M. Shah, "Object Tracking: A Survey", ACM Computing Surveys, vol. 38, no. 4, (2006).
- [49] H. Yang, L. Shao, F. Zheng, L. Wang and Z. Song, "Recent Advances and Trends in Visual Tracking: A Review", Neurocomputing, vol. 74, no. 18, (2011).
- [50] W. Rang-ben, G. Ke-you, S. Shu-ming and C. Jiang-wei, "A Monitoring Method of Driver Fatigue Behavior Based on Machine Vision", Proceeding of IEEE Intelligent Vehicles Symposium, (2003) June, Columbus, Ohio, USA.
- [51] M. H. Sigari, "Driver Hypo-Vigilance Detection based on Eyelid Behavior", Proceeding of 7th International Conference on Advances in Pattern Recognition (ICAPR), (2009) February, Kolkata, India.
- [52] G. Welch and G. Bishop, "An Introduction to the Kalman Filter", Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, USA, (2004) April.
- [53] D. A. Forsyth and J. Ponce, "Tracking, in Computer Vision, A Modern Approach", Prentice Hall, (2003), pp. 492-504.
- [54] S. J. Julier and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems", Proceeding of 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, (1997), Orlando, Florida, USA.
- [55] D. A. Forsyth and J. Ponce, "Tracking, in Computer Vision, A Modern Approach", Prentice Hall, (2003), pp. 511-522.
- [56] J. C. McCall and M. M. Trivedi, "Facial Action Coding Using Multiple Visual Cues and a Hierarchy of Particle Filters", Proceeding of IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPRW), (2006), New York, USA.
- [57] A. Amditis, A. Polychronopoulos, E. Bekiari and P. C. Antonello, "System Architecture of a Driver's Monitoring and Hypovigilance Warning System", Proceeding of IEEE Intelligent Vehicle Symposium, (2002) June, Versailles, France.
- [58] I. Park, J. H. Ahn and H. Byun, "Efficient Measurement of Eye Blinking under Various Illumination Conditions for Drowsiness Detection Systems", Proceeding of 18th International Conference on Pattern Recognition, (2006) September, Hong Kong, China.
- [59] M. H. Sigari, M. Fathy and M. Soryani, "A Driver Face Monitoring System for Fatigue and Distraction Detection", International Journal of Vehicular Technology, (2013).
- [60] D. Liu, P. Sun, Y. Q. Xiao and Y. Yin, "Drowsiness Detection Based on Eyelid Movement", Proceeding of 2nd International Workshop on Education Technology and Computer Science, (2010) March, Wuhan, China.
- [61] D. D. Salvucci and A. Liu, "The Time Course of a Lane Change: Driver Control and Eye-Movement Behavior", Transportation Research, Part F: Traffic Psychology and Behaviour, vol. 5, no. 2, (2002) June.
- [62] H. Zhou, M. Itoh and T. Inagakit, "Detection of Driver Intention of a Lane Change Through Monitoring Eye and Head Movements", Shisutemu, Joho Bumon Gakujutsu Koenkai Koen Ronbunshu, (2006).
- [63] H. Li, "Computer Recognition of Human Emotions", Proceeding of IEEE International Symposium on Intelligent Multimedia, Video and Speech Processing, (2001) May, Hong Kong, China.

- [64] M. P. Da Silva, V. Courboulay, A. Prigent and P. Estrailier, "Fast, Low Resource, Head Detection and Tracking for Interactive Applications", *PsychNology Journal*, vol. 7, no. 3, (2009).
- [65] K. Suzuki, J. Kobayashi, T. Takeshima and K. Yamada, "Detection of Unusual Facial Expression for Human Support Systems", *Proceeding of Conference of IEEE Industrial Electronics*, (2008) November, Orlando, USA.
- [66] H. Gu and Q. Ji, "Facial Event Classification with Task Oriented Dynamic Bayesian Network", *Proceeding of Conference on Computer Vision and Pattern Recognition (CVPR)*, (2004) July, Washington, USA.
- [67] M. E. Jabon, J. N. Bailenson, E. Pontikakis, L. Takayama and C. Nass, "Facial Expression Analysis for Predicting Unsafe Driving Behavior", *Pervasive Computing*, vol. 10, no. 4, (2011).
- [68] J. Zeng, Y. Sun and L. Jiang, "Driver Distraction Detection and Identity Recognition in Real-time", *Proceeding of WRI Global Congress on Intelligent Systems*, (2010) December, Wuhan, China.
- [69] A. Eskandarian and R. A. Sayed, "Detecting Driver Fatigue by Monitoring Eye and Steering Activity", *Proceeding of Annual Intelligent Vehicles Systems Symposium*, (2003) June, Michigan. USA.
- [70] A. Eskandarian and R. A. Sayed, "Analysis of Driver Impairment, Fatigue, and Drowsiness and an Unobtrusive Vehicle-Based Detection Scheme", *Proceeding of International Conference on Traffic Accidents*, (2005) December, Tehran, Iran.
- [71] M. Isard and A. Blake, "Condensation: Conditional Density Propagation for Visual Tracking", *International Journal of Computer Vision*, vol. 29, no. 1, (1998).
- [72] Type I and Type II Errors, http://en.wikipedia.org/wiki/Type_I_and_type_II_errors.
- [73] J. Jimenez-Pinto and M. Torres-Torriti, "Driver Alert State and Fatigue Detection by Salient Points Analysis", *Proceedings of the IEEE international conference on Systems, Man and Cybernetics (SMC)*, (2009) October, San Antonio, Texas, USA.
- [74] J. Jimenez-Pinto and M. Torres-Torriti, "Face Salient Points and Eyes Tracking for Robust Drowsiness Detection", *Robotica*, vol. 30, no. 5, (2012).
- [75] J. Jimenez-Pinto and M. Torres-Torriti, "Optical Flow and Driver's Kinematics Analysis for State of Alert Sensing", *Sensors*, vol. 13, (2013).
- [76] J. Shi and C. Tomasi, "Good Features to Track", *Proceedings of the 9th IEEE Conference on Computer Vision and Pattern Recognition*, (1994) June, Seattle, WA, USA.
- [77] S. Abtahi , B. Hariri and S. Shirmohammadi, "Driver Drowsiness Monitoring Based on Yawning Detection", *Proceedings of Instrumentation and Measurement Technology Conference*, (2011) May, Binjiang, China.

Authors



Mohamad-Hoseyn Sigari received his B.Sc. and M.Sc. degrees in computer engineering from Ferdowsi University of Mashhad and Iran University of Science and Technology, in 2006 and 2008 respectively with first rank honor. At present, he is a Ph.D. candidate in the department of electrical and computer engineering, University of Tehran, Tehran, Iran. His research interests are machine vision, biometric, semantic video analysis and intelligent transportation systems.



Muhammad-Reza Pourshahabi received his B.Sc. and M.Sc. degrees in computer engineering from Ferdowsi University of Mashhad, Iran, in 2006 and 2009 respectively. Now, he is a faculty member of computer engineering department, Salman institute of higher education, Mashhad, Iran. Also, he is a research assistant in machine vision research laboratory, Ferdowsi University of Mashhad. His research interests are image processing, medical image processing and biometric.



Mohsen Soryani received his Ph.D. degree in electronics engineering from Heriot-Watt University, Edinburgh, Scotland, in 1986. At present, he is an assistant professor in department of computer engineering, Iran University of Science and Technology, Tehran, Iran. His research interests are image processing, machine vision, computer networks and computer architecture.



Mahmood Fathy received his Ph.D. degree in image processing and processor design from University of Manchester Institute of Science and Technology, Manchester, UK, in 1991. Now, he is an associate professor in department of computer engineering, Iran University of Science and Technology, Tehran, Iran. His research interests are image processing, machine vision, computer network, computer architecture and intelligent transportation system.