Real Time Non-intrusive Monitoring and Prediction of Driver Fatigue

Zhiwei Zhu

E-mail: zhuz@rpi.edu

Department of Electrical, Computer and System Engineering,

Rensselaer Polytechnic Institute, Troy, NY, 12180, USA.

Qiang Ji

E-mail: qji@ecse.rpi.edu

Department of Electrical, Computer and System Engineering

Rensselaer Polytechnic Institute, Troy, NY, 12180, USA

Peilin Lan

E-mail: plan@cs.unr.edu

Dept. of Computer Science, University of Nevada, Reno

Manuscript correspondence:

Dr. Qiang Ji

E-mail: qji@ecse.rpi.edu,

Telephone: 1-518-276-6440

Department of Electrical, Computer, and Systems Engineering,

Rensselaer Polytechnic Institute

JEC 6044, Troy, NY 12180-3590

Abstract

This paper describes a real-time prototype computer vision system for monitoring driver vigilance. It uses a remotely located CCD camera equipped with an active IR illuminator to acquire video images of the driver. Various visual cues typically characterizing the level of alertness of a person are extracted in real time and systematically combined to infer the fatigue level of the driver. The visual cues used include eyelid movement, gaze movement, head movement, and facial expression. A probabilistic model is developed to model human fatigue and to predict fatigue based on the visual cues obtained. The simultaneous use of multiple visual cues and their systematic combination yields a much more robust and accurate fatigue characterization than using a single visual cue. The system was validated under real life fatigue conditions with human subjects of different ethnic backgrounds, different genders, ages, with/without glasses, and under different illumination conditions, and it was found reasonably robust, reliable and accurate in fatigue characterization.

Keywords

Driver Vigilance, Human Fatigue, Probabilistic Model, Visual Cues

I. Introduction

The ever-increasing number of traffic accidents in the U.S. due to a diminished driver's vigilance level has become a problem of serious concern to society. Drivers with a diminished vigilance level suffer from a marked decline in their abilities of perception, recognition, and vehicle control and therefore pose serious danger to their own life and the lives of other people. Statistics show that a leading cause for fatal or injury-causing traffic accidents is due to drivers with a diminished vigilance level. In the trucking industry, 57% fatal truck accidents are due to driver fatigue. It is the number 1 cause for heavy truck crashes. 70% of American drivers report driving fatigued. The National Highway Traffic Safety Administration (NHTSA) [1] estimates that there are 100,000 crashes, which are caused by drowsy drivers and result in more than 1,500 fatalities and 71,000 injuries each year in U.S. With the ever-growing traffic conditions, this problem will further deteriorate. For this reason, developing systems actively monitoring a driver's level of vigilance and alerting the driver of any insecure driving conditions is essential to accident prevention.

Many efforts [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21] have been reported in the literature for developing active safety systems for reducing the number of automobile accidents due to reduced vigilance. The techniques can be classified into the following categories [22].

- Readiness-to-perform and fitness-for-duty technologies

These technologies [10], [11], [12] attempt to assess the vigilance capacity of an operator before the work is performed. The tests conducted to assess the vigilance level of the operator consist of two groups: performance-based or measuring ocular physiology.

- Mathematical models of alertness dynamics joined with ambulatory technologies

The technologies use the mathematical models to predict operator alertness and performance at different times based on interactions of sleep, circadian, and related temporal antecedents of fatigue [13], [14], [15].

- Vehicle-based performance technologies

These technologies detect the behavior of the driver by monitoring the transportation hardware systems under the control of the driver, such as driver's steering wheel movements, driver's acceleration, braking and gear changing [16], [17], [18].

- In-vehicle, on-line, operator status monitoring technologies

The technologies in this category seek to real-time record some bio-behavioral dimension(s) of an operator, such as feature of the eyes, face, head, heart, brain activity, reaction time etc., during driving [19], [20], [21]. According to the different methods used for measurements, the technologies can be further divided into three types. The first type employs electroencephalograph measures (EEG), based on which most of successful equipments developed are off-line versions. Also, there is an on-line version called "Mind Switch" that uses a headband device, in which, the electrodes are embedded to make contact with the driver's scalp so as to measure the brain waves. Ocular measures are used in the second type, which is considered as the most suitable way for on-line monitoring. So far, many eye blinking, pupil response, eye closure and eye movement monitors have been developed. Other physiological/bio-behavioral measures used in the third type include tone of facial muscles (facial expression), body postures and head noddings.

Among different techniques, the best detection accuracy is achieved with techniques that measure physiological conditions like brain waves, heart rate, and pulse rate [9], [23]. Requiring physical contact with drivers (e.g., attaching electrodes) to perform, these techniques are intrusive, causing annoyance to drivers. Good results have also been reported with techniques that monitor eyelid movement and eye gaze with a head-mounted eye tracker or special contact lens. Results from monitoring head movement [24] with a head-mount device are also encouraging. These techniques, though less intrusive, are still not practically acceptable. A driver's state of vigilance can also be characterized by the behaviors of the vehicle he/she operates. Vehicle behaviors including speed, lateral position, turning angle, and moving course are good indicators of a driver's alertness level. While these techniques may be implemented non-intrusively, they are, nevertheless, subject to several limitations including the vehicle type, driver experiences, and driving conditions [3].

People in fatigue exhibit certain visual behaviors easily observable from changes in facial features like the eyes, head, and face. Visual behaviors that typically reflect a person's level of fatigue include eyelid movement, gaze, head movement and facial expression. To make use of these visual cues, another increasingly popular and non-invasive approach

for monitoring fatigue is to assess a driver's vigilance level through visual observation of his/her physical conditions using a remote camera and state-of-the-art technologies in computer vision. Techniques using computer vision are aimed at extracting visual characteristics that typically characterize a driver's vigilance level from his/her video images. In a recent workshop [25] sponsored by the Department of Transportation (DOT) on driver's vigilance, it is concluded that computer vision represents the most promising non-invasive technology to monitor driver's vigilance.

Many efforts have been reported in the literature on developing active real-time image-based fatigue monitoring systems [26], [2], [3], [6], [8], [9], [4], [5], [27], [28], [29], [30], [31], [32], [33]. These efforts are primarily focused on detecting driver fatigue. For example, Ishii [8] et al introduced a system for characterizing a driver's mental state from his facial expression. Saito et al [2] proposed a vision system to detect a driver's physical and mental conditions from line of sight (gaze). Boverie et al [4] described a system for monitoring driving vigilance by studying the eyelid movement. Their preliminary evaluation revealed promising results of their system for characterizing a driver's vigilance level using eyelid movement. Ueno et al [3] described a system for drowsiness detection by recognizing whether a driver's eyes are open or closed, and, if open, computing the degree of eye openness. Their study showed that the performance of their system is comparable with those of techniques using physiological signals.

Although the success of the existing approaches/systems for extracting characteristics of a driver using computer vision technologies, current efforts in this area, however, focus on using only a single visual cue such as eyelid movement or line of sight or head orientation to characterize driver's state of alertness. The system relying on a single visual cue may encounter difficulty when the required visual features cannot be acquired accurately or reliably. For example, drivers with glasses could pose serious problem to those techniques based on detecting eye characteristics. Glasses can cause glare and may be totally opaque to light, making it impossible for camera to monitor eye movement. Furthermore, the degree of eye openness may vary from people to people. Another potential problem with the use of a single visual cue is that the obtained visual feature is often ambiguous, therefore can not always be indicative of one's mental conditions. For example, the irregular head

movement or line of sight (like briefly look back or at the minor) may yield false alarms for such a system.

All those visual cues, however imperfect they are individually, if combined systematically, can provide an accurate characterization of a driver's level of vigilance. It is our belief that simultaneous extraction and use of multiple visual cues can reduce the uncertainty and resolve the ambiguity present in the information from a single source. The systematic integration of these visual parameters, however, requires a fatigue model that models the fatigue generation process and is able to systematically predict fatigue based on the available visual as well as the relevant contextual information. The system we propose can simultaneously, non-intrusively, and in real time monitor several visual behaviors that typically characterize a person's level of alertness while driving. These visual cues include eyelid movement, pupil movement, head movement and facial expression. The fatigue parameters computed from these visual cues are subsequently combined probabilistically to form a composite fatigue index that could robustly, accurately, and consistently characterize one's vigilance level. Figure 1 gives an overview of our driver vigilance monitoring system.

The paper consists of three parts. First, the paper focuses on discussion of the computer vision algorithms and the necessary hardware components to extract the needed visual cues. Second, after extracting these visual cues, the issue of sensory data fusion and fatigue modeling and inference is discussed. Finally, experiments under real life conditions are conducted to validate our driver vigilance monitoring system.

II. Eye Detection and Tracking

Fatigue monitoring starts with extracting visual parameters typically characterizing a person's level of vigilance. This is accomplished via a computer vision system. In this section, we discuss the computer vision system we developed to achieve this goal. Figure 2 provides an overview of our visual cues extraction system for driver fatigue monitoring. The system consists of two cameras: one wide angle camera focusing on the face and another narrow angle camera focusing on the eyes. The wide angles camera monitors head movement and facial expression while the narrow angle camera monitors eyelid and gaze movements. The system starts with eye detection and tracking.

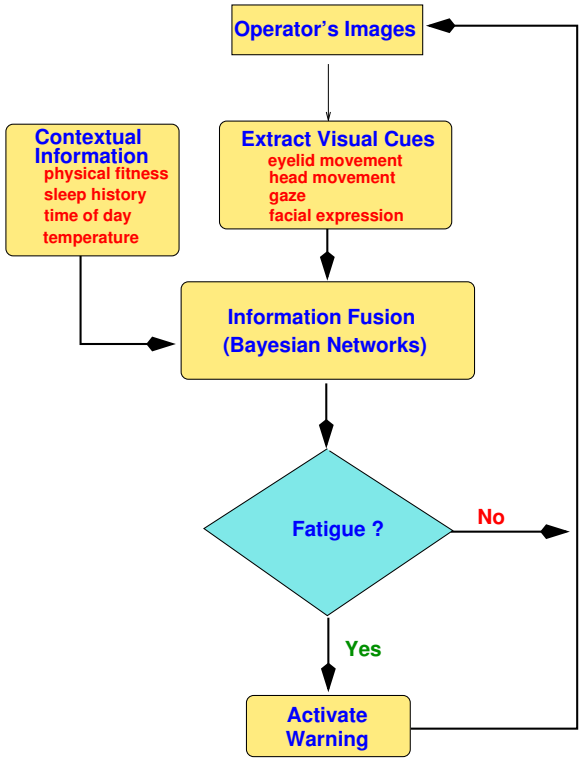


Fig. 1. A flowchart of the proposed driver vigilance monitoring system

The goal of eye detection and tracking is for subsequent eyelid movement monitoring, gaze determination, face orientation estimation and facial expression analysis. A robust, accurate, and real-time eye tracker is therefore crucial. In this research, we propose real-time robust methods for eye tracking under variable lighting conditions and face orientations, based on combining the appearance-based methods and the active IR illumination approach. Combining the respective strengths of different complementary techniques and overcoming their shortcomings, the proposed method uses active infrared illumination to brighten subject's faces to produce the bright pupil effect. The bright pupil effect and appearance of eyes (statistic distribution based on eye patterns) are utilized simultaneously for eyes detection and tracking. The latest technologies in pattern classification recognition (the Support Vector Machine) and in object tracking (the mean-shift) are employed for eye detection and tracking based on eyes appearance.

Our method consists of two parts: eye detection and eye tracking. Figure 3 summarizes our eye detection and tracking algorithm. Some of the ideas presented in this paper have

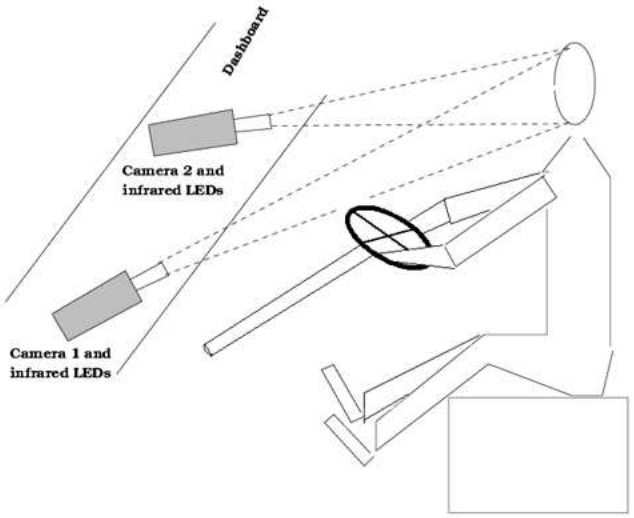


Fig. 2. Overview of the driver vigilance monitoring system

been reported in [34],[35]. In the sections to follow, we summarize our eye detection and tracking algorithms.

A. Image Acquisition System

Image understanding of visual behaviors starts with image acquisition. The purpose of image acquisition is to acquire the video images of the driver face in real time. The acquired images should have relatively consistent photometric property under different climatic/ambient conditions and should produce distinguishable features that can facilitate the subsequent image processing. To this end, the person's face is illuminated using a near-infrared illuminator. The use of infrared illuminator serves three purposes: first it minimizes the impact of different ambient light conditions, therefore ensuring image quality under varying real-world conditions including poor illumination, day, and night; second, it allows to produce the bright/dark pupil effect, which constitutes the foundation for detection and tracking the proposed visual cues. Third, since near infrared is barely visible to the driver, this will minimize any interference with the driver's driving.

Specifically, our IR illuminator consists of two sets of IR LEDs, distributed evenly and symmetrically along the circumference of two coplanar concentric rings as shown in Figure 4. The center of both rings coincides with the camera optical axis. These IR LEDs will

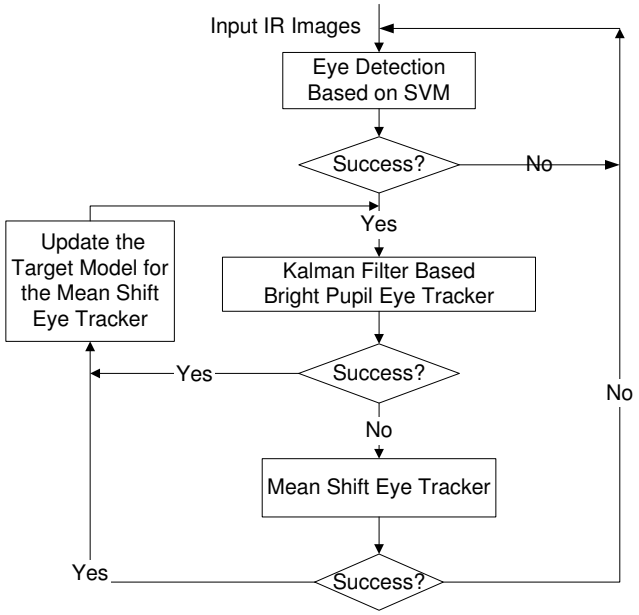


Fig. 3. The combined eye tracking flowchart

emit non-coherent IR energy in the 800 to 900 nanometer region of the spectrum.

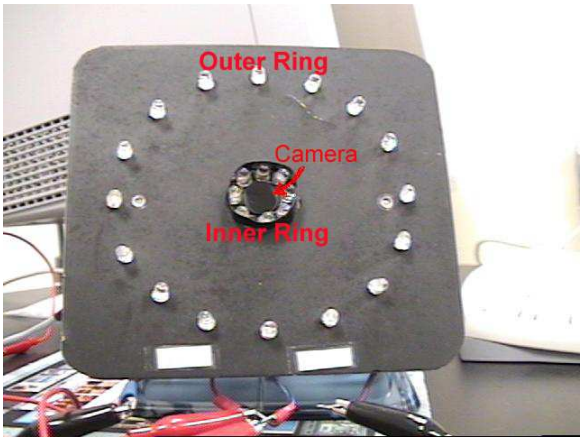
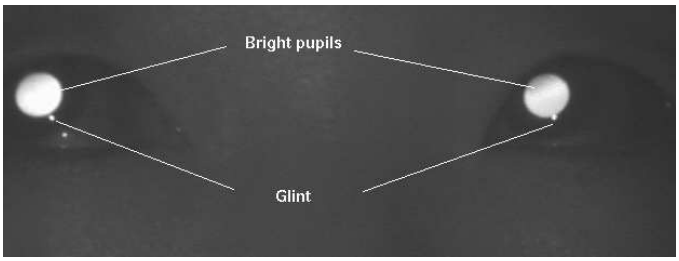


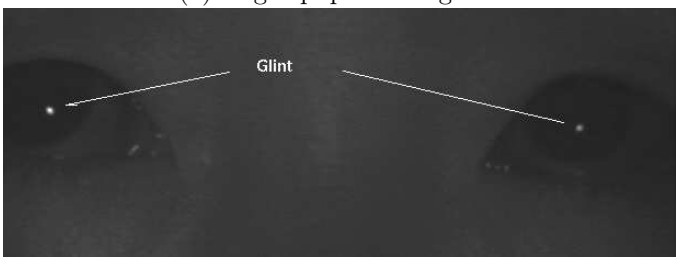
Fig. 4. An actual photo of the two rings IR illuminator configuration

The bright pupil image is produced when the inner ring of IR LEDs is turned on and the dark pupil image is produced when the outer ring is turned on. This is controlled via a video decoder. An example of the bright/dark pupils is given in Figure 5. Note the glint [1] appears on both the dark and bright pupil images.

[1]the small bright spot near the pupil, produced by cornea reflection of the IR light.



(a) bright pupils with glints



(b) dark pupils with glints

Fig. 5. Bright and dark pupil images with glints

B. Eye Detection

Eyes tracking starts with eyes detection. Figure 6 gives a flowchart of the eye detection procedure. Eyes detection is accomplished via pupils detection due to the use of active IR illumination.

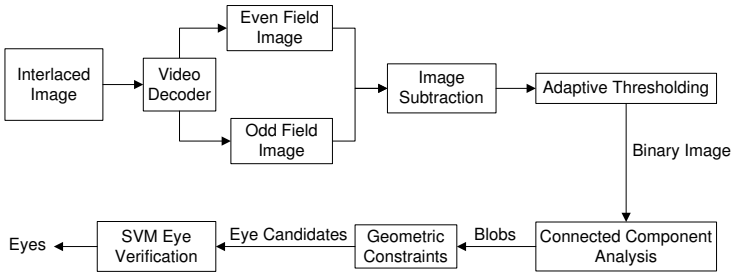


Fig. 6. Eye detection block diagram

Specifically, to facilitate pupils detection, we have developed a circuitry to synchronize the inner ring of LEDs and outer ring of LEDs with the even and odd fields of the interlaced image respectively so that they can be turned on and off alternately. The interlaced input image is de-interlaced via a video decoder, producing the even and odd field images as shown in Figure 7 (a) and (b). While both images share the same background and external

illumination, pupils in the even images look significantly brighter than in the odd images. To eliminate the background and reduce external light illumination, the odd image is subtracted from the even image, producing the difference image as shown in Figure 7 (c), with most of the background and external illumination effects removed. The difference image is subsequently thresholded. A connected component analysis is then applied to the thresholded difference image to identify binary blobs that satisfy certain size and shape constraints as shown in Figure 8 (a). From Figure 8 (a), we can see that there are still several non-pupil blobs left because they are so similar in shape and size that we can not distinguish the real pupil blobs from them. So we have to use other features.

(a) (b) (c)

Fig. 7. (a) Even field image (b) Odd field image (c) The difference image

From the dark pupil image as shown in Figure 8 (b), we observed that each pupil is surrounded by the eye region, which has an unique intensity distribution and appears different from other parts of the face. The appearance of an eye can therefore be utilized to separate it from non-eyes. We map the locations of the remaining binary blobs to the dark pupil images and then apply the Support Vector Machine (SVM) classifier [36], [37] to automatically identify the binary blobs that correspond to eyes. A large number of training images including eyes and non-eyes were used to train the SVM classifier. Figure 8 (c) shows that the SVM eye classifier correctly identify the real eye regions as marked and remove the spurious ones. Details on our eye detection algorithm may be found in [34].

(a) (b) (c)

Fig. 8. (a) The thresholded difference image marked with possible pupil candidates; (b) The image marked with possible eye candidates according to the positions of pupil candidates; (c) The image marked with identified eyes

C. Eye Tracking Algorithm

The detected eyes are then tracked frame to frame. We have developed the following algorithm for the eye tracking by combining the bright pupil based Kalman filter eye tracker with the mean shift eye tracker [35]. While Kalman filtering accounts for the dynamics of the moving eyes, mean shift tracks eyes based on the appearance of the eyes. We call this two-stage eye tracking.

After locating the eyes in the initial frames, the Kalman filtering is activated to track bright pupils. Kalman filter pupil tracker works reasonably well under frontal face orientation with eye open. However, it will fail if the pupils are not bright due to oblique face orientations, eye closures or external illumination interferences. Kalman filter also fails when sudden head movement occurs because the assumption of smooth head motion has been violated. Therefore, we propose to use mean shift tracking to augment Kalman filtering tracking to overcome this limitation. If Kalman filtering tracking fails in a frame, eye tracking based on mean shift will take over. Mean shift tracking is an appearance based object tracking method which tracks the eye regions according to the intensity statistical distributions of the eye regions and doesn't need bright pupils. It employs the mean shift analysis to identify an eye candidate region, which has the most similar appearance to the given eye model in terms of intensity distribution. Therefore, the mean shift eye tracking can track the eyes successfully under eye closure or under oblique face orientations. Also, it is fast and handles noise well. But it does not have the capability of self-correction and

the errors therefore tend to accumulate and propagate to subsequent frames as tracking progresses and eventually the tracker drifts away.

To overcome these limitations with mean shift tracker, we propose to combine the Kalman filter tracking with the mean shift tracking to overcome their respective limitations and to take advantage of their strengths. Specifically, we take the following measures. First, two channels (eye images with dark pupil and bright pupil) are used to characterize the statistical distributions of the eyes. Second, the eye's model is continuously updated by the eyes detected by the last Kalman filtering tracker to avoid the error propagation with Mean Shift tracker. Finally, the experimental determination of the optimal window size and quantization level for mean shift tracking further enhance the performance of our technique.

The two trackers are activated alternately. The Kalman tracker is first initiated, assuming the presence of the bright pupils. When the bright pupils appear weak or disappear, the mean shift tracker is activated to take over the tracking. Mean shift tracking continues until the reappearance of the bright pupils, when the Kalman tracker takes over. Eye detection will be activated if the mean shift tracking fails. These two stage eye trackers work together and they complement each other. The robustness of the eye tracker is improved significantly. The Kalman filtering and mean shift tracking algorithms are discussed in [38], [34].

The eye detection and tracking algorithm is tested with different subjects under different face orientations and different illuminations. These experiments reveal that our algorithm is more robust than the conventional Kalman filter based bright pupil tracker, especially for the closed eyes and partially occluded eyes due to the face orientations. Even under strong external illuminations, we have achieved good results. Video demos are available at *http://www.ecse.rpi.edu/~cvrl/Demo/demo.html*.

III. Eyelid Movement Parameters

Eyelid movement is one of the visual behaviors that reflect a person's level of fatigue. The primary purpose of eye tracking is to monitor eyelid movements and compute the relevant eyelid movement parameters. Here, we focus on two ocular measures to characterize the eyelid movement. The first one is Percentage of Eye Closure Over Time (*PERCLOS*)

and the second is Average Eye Closure Speed (*AECS*). *PERCLOS* has been validated and found to be the most valid ocular parameter for monitoring fatigue [26].

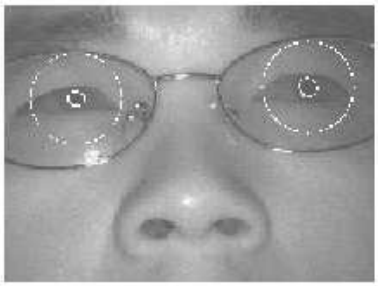
The eye closure/opening speed is a good indicator of fatigue. It's defined as the amount of time needed to fully close the eyes or to fully open the eyes. Our previous study indicates that the eye closure speed of a drowsy person is distinctively different from that of an alert person [38].

The degree of eye opening is characterized by the shape of pupil. It is observed that as eyes close, the pupils start getting occluded by the eyelids and their shapes get more elliptical. So, we can use the ratio of pupil ellipse axes to characterize degree of eye opening. The cumulative eye closure duration over time excluding the time spent on normal eye blinks is used to compute *PERCLOS*. To obtain a more robust measurement for these two parameters, we compute their running average (time tracking). To obtain running average of *PERCLOS* measurement, for example, the program continuously tracks the person's pupil shape and monitors eye closure at each time instance. We compute these two parameters in 30 seconds window and output them onto the computer screen in real time, so we can easily analyze the alert state of the driver. The plots of the two parameters over time are shown in Figure 9. Also, video demos are available at *http://www.ecse.rpi.edu/~cvrl/Demo/demo.html*.

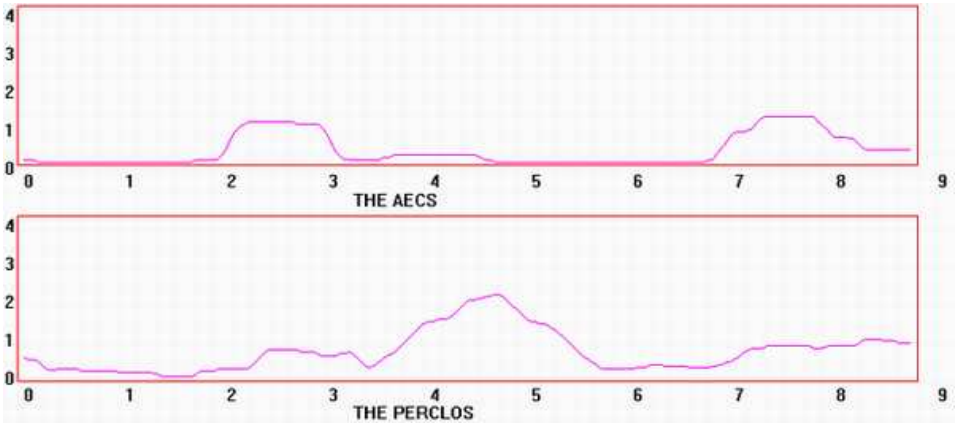
IV. Face (Head) Orientation Estimation

Face (head) pose contains information about one's attention, gaze, and level of fatigue. Face pose determination is concerned with computation of the 3D face orientation and position to detect such head movements as head tilts. Frequent head tilts indicate the onset of fatigue. Furthermore, the nominal face orientation while driving is frontal. If the driver faces in the other directions (e.g., down or sideway) for an extended period of time, this is due to either fatigue or inattention. Face pose estimation, therefore, can indicate both fatigued and inattentive drivers. For this study, we focus on the former, i.e., detection of frequent head tilts.

We present a new technique to perform the 2D face tracking and 3D face pose estimation synchronously. In our method, 3D face pose is tracked by Kalman Filtering. The initial estimated 3D pose is used to guide face tracking in the image, which is subsequently used



(a)



(b)

Fig. 9. (a) Detected eyes and pupils (b) Plots for eyelid movement parameters: the top one displays *AECS* parameter and the bottom one displays *PERCLOS* parameter.

to refine the 3D face pose estimation. Face detection and pose estimation work together and benefit from each other. Weak perspective projection is assumed so that face can be approximated as a planar object with facial features, such as eyes, nose and mouth, located symmetrically on the plane. Figure 10 summarizes our approach. Initially, we automatically detect a fronto-parallel face view based on the detected eyes [35] and some simple anthropometric statistics. The detected face region is used as the initial 3D planar face model. The 3D face pose is then tracked starting from the fronto-parallel face pose. During tracking, the 3D face model is updated dynamically, and the face detection and face pose estimation are synchronized and kept consistent with each other.

We will discuss our face pose tracking algorithm briefly as follows.

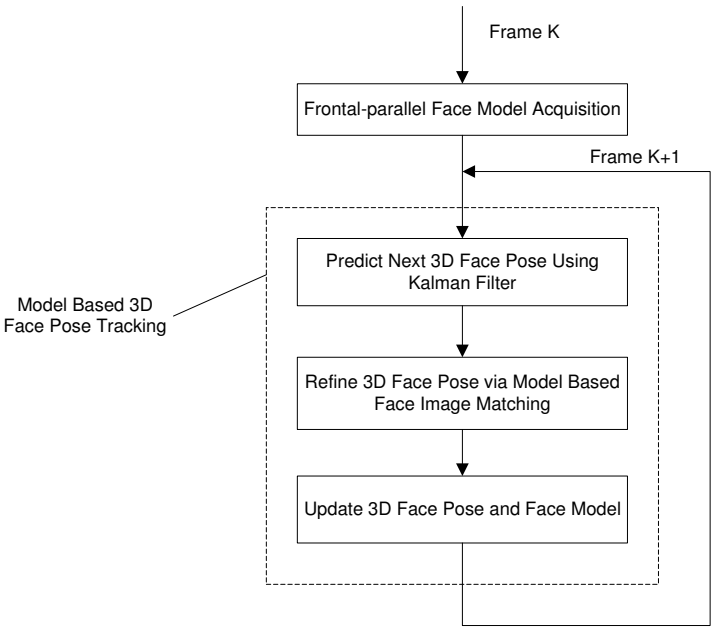


Fig. 10. The flowchart of face pose tracking

A. Automatic 3D Face Model and Pose Initialization

In our algorithm, we should have a fronto-parallel face to represent the initial face model. This initialization is automatically accomplished by using the eye tracking technique we have developed [35]. Specifically, the subject starts in fronto-parallel face pose position with the face facing directly to the camera as shown in figure 11. The eye tracking technique is then activated to detect eyes. After detecting the eyes, the first step is to compute the distance d_{eyes} between two eyes. Then, the distance between the detected eyes, eyes locations and the anthropometric proportions are used to estimate the scope and the location of the face in the image automatically. Experiments show that our face detection method works well for all the faces we tested. Example of the detected frontal face region is shown in Figure 11. Once the face region is decided, we will treat it as our initial face model, whose pose parameters are used as initial face pose parameters.

Compared with the existing frontal face detection methods, ours takes full advantage of the detected eyes to guide the detection of the frontal face, and it is simple, robust and automatic. In [39], we will also demonstrate the tolerance of our face initialization to slight deviations from the fronto-parallel face pose and to perturbations of initial positions

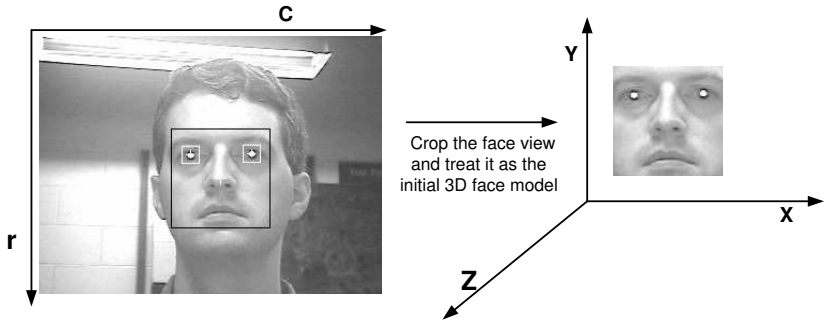


Fig. 11. The initial face model

of the face.

B. Face Pose Tracking Algorithm

Given the initial face image and its pose in the first frame, the task of finding the face location and the face pose in subsequent frames can be implemented as simultaneous 3D face pose tracking and face detection described in [39].

Based on the detected face pose in the previous frame, we can use the Kalman Filtering to predict the face pose in the next frame. But the prediction based on Kalman Filtering assumes smooth face movements. The prediction will be off significantly if head undergoes a sudden rapid movement. To handle this issue, we propose to approximate the face movement with eyes movement since eyes can be reliably detected in each frame. Then the final predicted face pose is based on combining the one from Kalman Filtering with the one from eyes. The simultaneous use of Kalman Filtering and eye's motion allows to perform accurate face pose prediction even under significant or rapid head movements. Details on our face pose estimation and tracking algorithm may be found at [39].

The proposed algorithm is tested with numerous image sequences of different people. The image sequences include a person rotating his head before an un-calibrated camera, which is approximately 1.5 meter from the person. Figure 12 shows some tracking results under different face rotations. It is shown that the estimated pose is very visually convincing over a large range of head orientations and changing distance between the face and camera. Plots of three face pose angles ω, ϕ, κ are shown in Figure 13, from which we can see that three face pose angles vary consistently and smoothly as the head rotates. Video demos of our system may be found at *http://www.ecse.rpi.edu/~cvrl/Demo/demo.html*.

Fig. 12. Face and face pose tracking results for images randomly selected from one video sequence. The white rectangle indicates the tracked face region and the white line represents the normal of the face plane, which is drawn according to the estimated face pose.

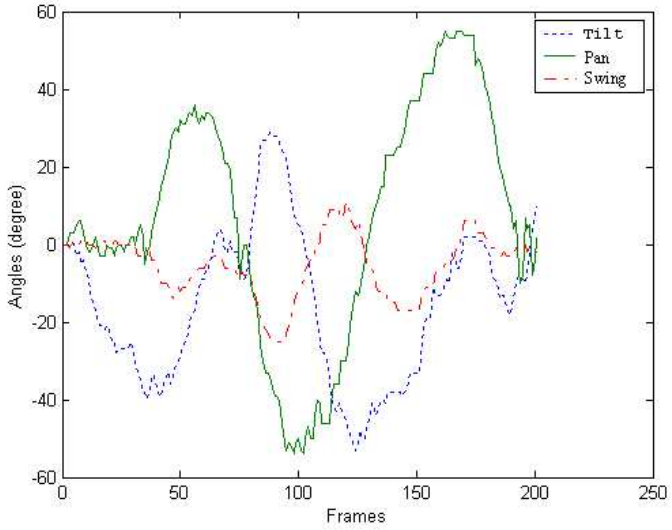


Fig. 13. The results of face pose tracking. The plots show the sequences of three estimated rotation angles through the image sequence.

To quantitatively characterize one's level of fatigue by face pose, we introduce a new fatigue parameter called *NodFreq*, which measures the frequency of head tilts over time. Figure 14 shows the running average of the estimated head tilts for a period of 140 seconds. As can be seen, our system can accurately detect head tilts, which are represented in the curve by the up-and-down bumps.

V. Eye Gaze Determination and Tracking

Gaze has the potential to indicate a person's level of vigilance. A fatigued individual tends to have a narrow gaze. Gaze may also reveal one's needs and attention. The direction

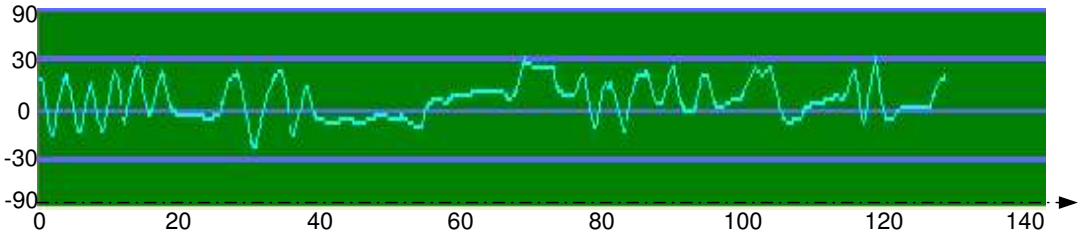


Fig. 14. Head tilts monitoring over time (seconds)

of a person's gaze is determined by two factors: the orientation of the face (face pose), and the orientation of eye (eye gaze). Face pose determines the global direction of the gaze, while eye gaze determines the local direction of the gaze. Global gaze and local gaze together determine the final gaze of the person. So far, the most common approach for ocular-based gaze estimation is based on the determination of the relative position between pupil and the glint (cornea reflection) via a remote IR camera [40], [41], [42], [43], [44], [45], [46]. While contact-free and non-intrusive, these methods work well only for a static head, which is a rather restrictive constraint on the part of the user. Even a chin rest is often used to maintain the head still because minor head movement can fail these techniques. This poses a significant hurdle for practical application of the system. Another serious problem with the existing eye and gaze tracking systems is the need to perform a rather cumbersome calibration process for each individual. Often re-calibration is even needed for the same individual who already underwent the calibration procedure, whenever his/her head moved. This is because only local gaze is accounted for while global gaze due to face pose is ignored.

In view of these limitations, we present a gaze estimation approach [47] that accounts for both the local gaze computed from the ocular parameters and the global gaze computed from the head pose. The global gaze (face pose) and local gaze (eye gaze) are combined together to obtain the precise gaze information of the user. Our approach, therefore, allows natural head movement while still estimating gaze accurately. Another effort is to make the gaze estimation calibration free. New users or the existing users who have moved, do not need undergo a personal gaze calibration before using the gaze tracker. Therefore, the proposed gaze tracker can perform robustly and accurately without calibration and under

natural head movements.

An overview of our algorithm is given in Figure 15.

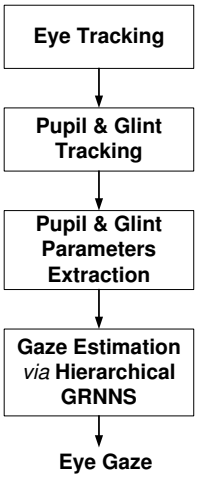


Fig. 15. Major components of the proposed system

A. Gaze Estimation

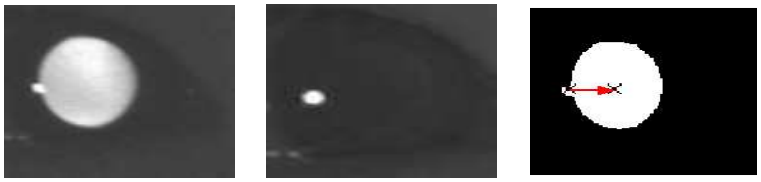
Our gaze estimation algorithm consists of three parts: pupil-glint detection and tracking, gaze calibration, and gaze mapping.

Gaze estimation starts with pupil & glint detection and tracking. For gaze estimation, we continue using the IR illuminator as shown in Figure 4. To produce the desired pupil effects, the two rings are turned on and off alternately via the video decoder we developed to produce the so-called bright and dark pupil effect as shown in Figure 5 (a) and (b). The pupil detection and tracking technique can be used to detect and track glint from the dark images. Figure 16 (c) shows the detected glints and pupils.

Given the detected glint and pupil, we can use their properties to obtain local gaze and global gaze. Figure 16 shows the relationship between the local gaze and the relative position between the glint and the pupil, i.e., the pupil-glint vector.

Our study in [48] shows there exists a direct correlation between 3D face pose and the geometric properties of the pupils. Specifically, pupil size, shape, and orientation vary with face pose. It is therefore to capture 3D face pose using pupil geometric properties. 3D face pose provides the global gaze.

In order to obtain the final gaze, the factors accounting for the head movements and



(a) (b) (c)

Fig. 16. Relative spatial relationship between glint and bright pupil center used to determine local gaze position. (a) bright pupil images, (b) glint images; (c) pupil-glint vector indicating local gaze direction.

those affecting the local gaze should be combined. Hence, six parameters are chosen for the gaze calibration to get the parameters mapping function: Δx, Δy, r, θ, g_x and g_y. Δx and Δy are the pupil-glint displacement. r is the ratio of the major to minor axes of the ellipse that fits to the pupil. θ is the pupil ellipse orientation and g_x and g_y are the glint image coordinates. The choice of these factors is based on the following rational. Δx and Δy account for the relative movement between the glint and the pupil, representing the local gaze. The magnitude of the glint-pupil vector can also relate to the distance of the subject to the camera. r is used to account for out-of-plane face rotation. The ratio should be close to one when the face is frontal. The ratio becomes larger or less than 1 when the face turns either up/down or left/right. Angle θ is used to account for in-plane face rotation around the camera optical axis. Finally, (g_x, g_y) is used to account for the in-plane head translation.

The use of these parameters accounts for both head and pupil movement since their movements will introduce corresponding changes to these parameters. This effectively reduces the head movement influence. Given the six parameters affecting gaze, we now need to determine the mapping function that maps the parameters to the actual gaze. This mapping function can be approximated by the Generalized Regression Neural Networks (GRNN) [49]. GRNN features fast training times, can model non-linear functions, and has been shown to perform well in noisy environments given enough data. Specifically, the input vector to the GRNN is

$$\mathbf{g} = \begin{bmatrix} \Delta x & \Delta y & r & \theta & g_x & g_y \end{bmatrix}$$

A large amount of training data under different head positions is collected to train the GRNN. During the training data acquisition, the user is asked to fixate his/her gaze on

TABLE I

GAZE CLASSIFICATION RESULTS FOR THE GRNN GAZE CLASSIFIER. AN AVERAGE OF GAZE CLASSIFICATION ACCURACY OF (96% ACCURACY) WAS ACHIEVED FOR 480 TESTING DATA NOT INCLUDED IN THE TRAINING DATA FOR THE GAZE CLASSIFIER.

ground truth	estimated result (mapping target #)				correctness
(target #)	1	2	3	4	rate (%)
1	114	4	2	0	95
2	0	117	0	3	97.5
3	2	3	111	4	92.5
4	0	0	0	120	100

each predefined gaze region. After training, given an input vector, the GRNN can then approximate the user's actual gaze.

Experiments were conducted to study the performance of our gaze estimation technique. Table I shows some results. An average of gaze classification accuracy of (96% accuracy) was achieved for 480 testing data not included in the training data as shown in the confusion table I.

Given the gaze, we can compute a new fatigue parameter named *GAZEDIS*, which represents the gaze distribution over time to indicate the driver's fatigue or attention level. GAZEDIS measures the driver's situational awareness. Another fatigue parameter we compute is PERSAC, which is the percentage of saccade eye movement over time. Saccade eye movements represent the deliberate and conscious driver action to move eye from one to another place. It therefore can measure the degree of alertness. The value of PERSAC is very small for a person in fatigue. Figure 17 plots the PERSAC parameter over 30 seconds.

VI. Facial Expression Analysis

Besides eye and head movements, another visual cue that can potentially capture one's level of fatigue is his/her facial expression. In general, people tend to exhibit different

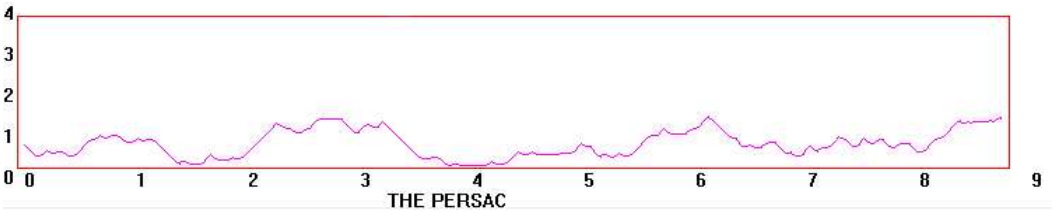


Fig. 17. Plot of PERSAC parameter over 30 seconds.

facial expressions under different levels of vigilance. The facial expression of a person in fatigue or in the onset of fatigue can usually be characterized by lagging facial muscles, expressionless, and frequent yawnings.

Our recent research has led to the development of a feature-based facial expression analysis algorithm. The facial features around eyes and mouth represent the most important spatial patterns composing the facial expression. Generally, these patterns with their changes in spatio-temporal spaces can be used to characterize facial expressions. For the fatigue detection application, in which there are only limited facial expressions, the facial features around eyes and mouth include enough information to capture these limited expressions. So in our research, we focus on the facial features around eyes and mouth. We use 22 fiducial features and three local graphs as the facial model (shown in Fig.18).

In our method, the multi-scale and multi-orientation Gabor wavelet is used to represent and detect each facial feature. For each pixel in the image, a set of Gabor coefficients in the complex form can be obtained by convolution with the designed Gabor kernels. These coefficients can be used to represent this pixel and its vicinity [50]. After training, these coefficients are subsequently used for facial feature detection.

After detecting each feature in the first frame, a Kalman Filter-based method with the eye constraints is proposed to track them. The Kalman filter is used to predict the current feature positions from the previous locations. It puts a smooth constraint on the motion of each feature. The eye positions from our eye tracker provide a strong and reliable information that gives where a rough location of face is and how the head moves between two consecutive frames. By combining the head motion information inferred from the detected eyes with the predicted locations from the Kalman Filtering, we can obtain a very accurate and robust prediction of feature locations in the current frame, even under

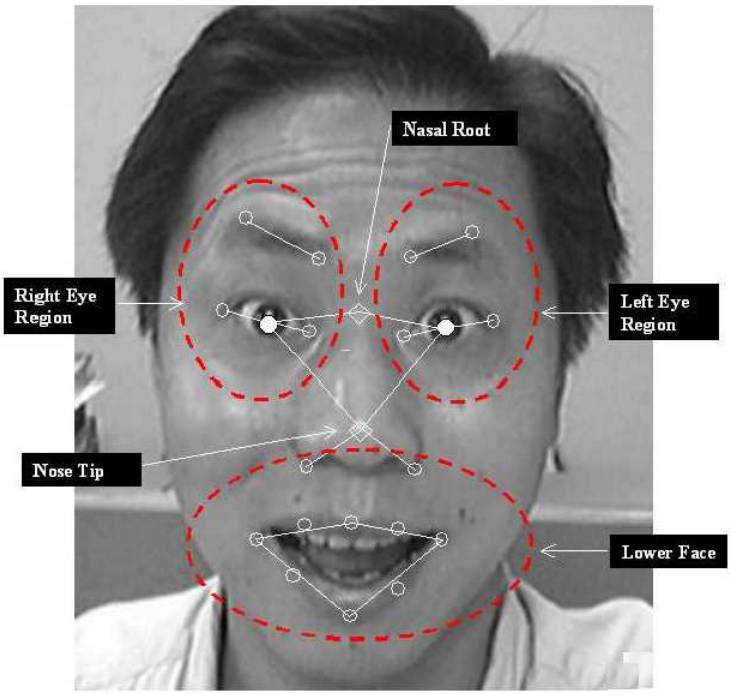


Fig. 18. The facial features and local graphs

rapid head movement. The detected features and their spatial connections are used to characterize facial expressions. Details can be found in [51].

A series of experiments are conducted in [51] , and good results are achieved under large head movements, self-occlusion and different facial expressions. Figure 19 shows the results of a typical sequence of a person in fatigue. It consists of blended facial expressions. The person in the scene yawned from the neutral state, then moved the head rapidly from the frontal view to the large side view and back to the other direction, raised the head up and finally returned to the neutral state. During the head movements, the facial expression changes dramatically.

For now, we focus on monitoring mouth movement to detect yawning. A yawning is detected if the features around mouth significantly deviate from its closed configuration, especially in vertical direction. There are 8 tracked facial features around the mouth as shown in Figure 20. Also, as shown in Figure 20, the height of the mouth is represented by the distance between the upper-lip and the lower-lip, and the width of the mouth is represented by the distance between the left and right mouth corners. The degree of mouth opening is characterized by the shape of mouth. Therefore, the openness of the mouth

Fig. 19. Tracked facial features and local graphes.

can be represented by the ratio of mouth height and width.

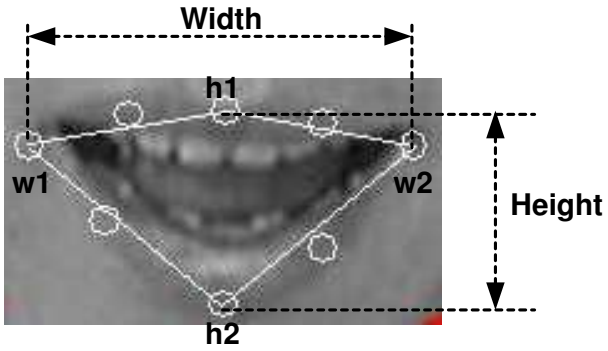


Fig. 20. The facial features to be tracked around the mouth and the mouth width and height used to represent the openness of the mouth.

We develop a new measure of facial expression, *YawnFreq*, which computes the occurrence frequency of yawning over time. Figure 21 shows the plot of *YawnFreq* over time, and a yawning is represented by an up-and-down bump.

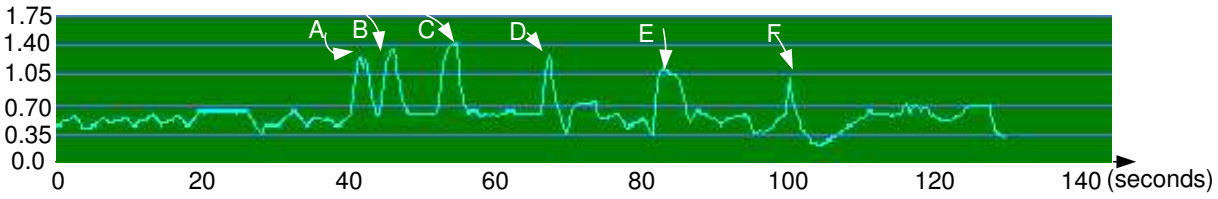


Fig. 21. The plot of the openness of the mouth over time. The bumps A, B, C, D, E, F are the detected yawns.

VII. Fatigue Modeling Using Bayesian Networks

As we discussed above, human fatigue generation is a very complicated process. Several uncertainties may be present in this process. First, fatigue is not observable and it can only be inferred from the available information. In fact, fatigue can be regarded as the result of many contextual variables such as working environments, health and sleep history. Also, it is the cause of many symptoms, e.g. the visual cues, such as irregular eyelid movements, yawning and frequent head tilts. Second, human's visual characteristics vary significantly with age, height, health and shape of face. To effectively monitor fatigue, a system that integrates evidences from multiple sources into one representative format is needed. Naturally, a Bayesian Networks (BN) model is the best option to deal with such an issue.

A BN provides a mechanism for graphical representation of uncertain knowledge and for inferring high level activities from the observed data. Specifically, a BN consists of nodes and arcs connected together forming a directed acyclic graph (DAG) [52]. Each node can be viewed as a domain variable that can take a set of discrete values or a continuous value. An arc represents a probabilistic dependency between the parent node and the child node.

The main purpose of a BN model is to infer the unobserved events from the observed or contextual data. So, the first step in BN modeling is to identify those hypothesis events and group them into a set of mutually exclusive events to form the target hypothesis variable. The second step is to identify the observable data that may reveal something about the hypothesis variable and then group them into information variables. There are also other hidden states which are needed to link the high level hypothesis node with the low level information nodes. For fatigue modeling, fatigue is obviously the target

hypothesis variable that we intend to infer. While other contextual factors, which could cause fatigue, and visual cues, which are symptoms of fatigue, are information variables. Among many factors that can cause fatigue, the most significant ones are sleep history, circadian, work condition, work environment, and physical condition. The most profound factors that characterize work environment are temperature, weather and noise; the most significant factors that characterize physical condition are age and sleep disorders; the significant factors characterizing circadian are time of day and time zone change; the factors affecting work conditions include workload and type of work. Furthermore, factors affecting sleep quality include sleep environment and sleep time. The sleep environment includes random noise, background light, heat and humidity.

The vision system discussed in previous sections can compute several visual fatigue parameters. They include *PERCLOS* & *ACSE* for eyelid movement, *NodFreq* for head movement, *GAZEDIS* and *PERSAC* for gaze movement, and *YawnFreq* for facial expression. Putting all these factors together, the BN model for fatigue is constructed as shown in Fig.22. The target node is fatigue. The nodes above the target node represent various major factors that could lead to one's fatigue. They are collectively referred to as contextual information. The nodes below the target node represent visual observations from the output of our computer vision system. These nodes are collectively referred to as observation nodes.

A. Construction of Conditional Probability Table (CPT)

Before using the BN for fatigue inference, the network needs to be parameterized. This requires specifying the prior probability for the root nodes and the conditional probabilities for the links. Usually, probability is obtained from statistical analysis of a large amount of training data. For this research, training data come from three different sources. First, we obtain some training data from the human subjects study we conducted. These data are used to train the lower part of the BN fatigue model. Second, several large-scale subjective surveys [1], [53], [54], [55] provide additional such data. Despite the subjectivity with these data, we use them to help parameterize our fatigue model. They were primarily used to train the upper part of the fatigue model. Since these surveys were not designed for the parameterizations of our BN model, not all needed probabilities are available and some

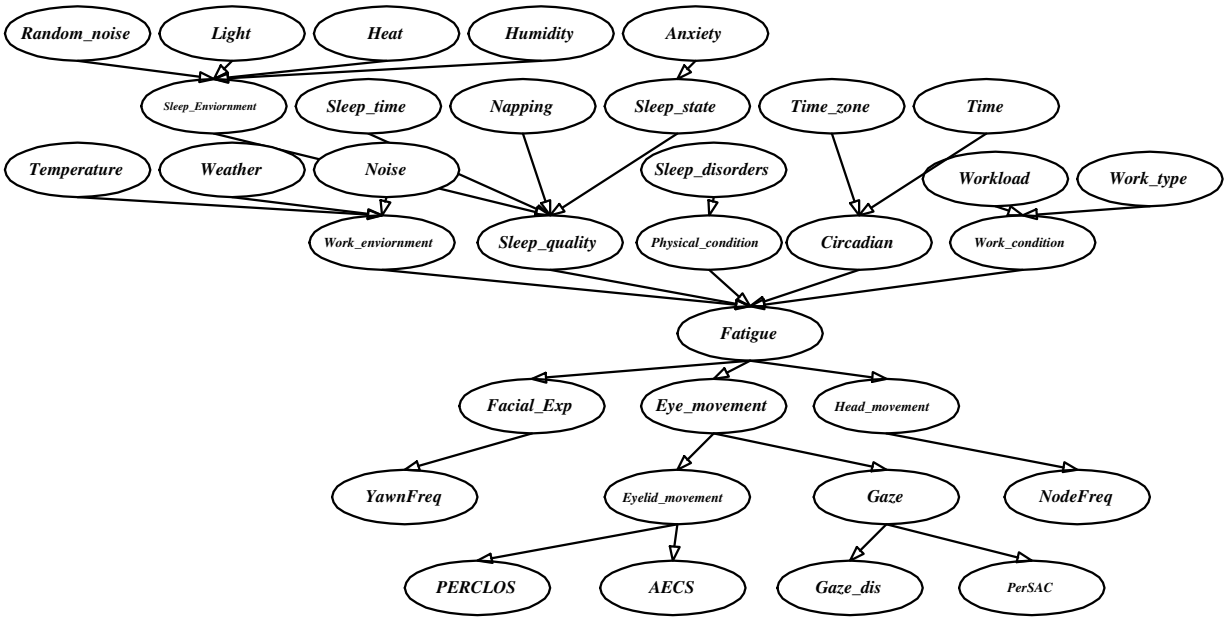


Fig. 22. Bayesian Network model for monitoring human fatigue

conditional probabilities are therefore inferred from the available data using the so-called noisy or principle [56]. Third, still some prior or conditional probabilities are lacking in our model, they are obtained by subjective estimation methods [56]. With the methods discussed above, all the prior and conditional probabilities in our BN model are obtained, part of which are summarized in Table II.

B. Fatigue Inference

Given the parameterized model, fatigue inference can then commence upon the arrival of visual evidences via belief propagation. MSBNX software [57] is used to perform the inference and both top-down and bottom-up belief propagations are performed. Here we use some typical combination of evidences and their results are summarized in Table III.

From Table III, we can see that the prior probability of fatigue (e.g. when there is not any evidence) is about 0.5755 (ref#1). The observation of single visual evidence does not usually provide conclusive finding since the estimated fatigue probability is less than the critical value 0.95 (ref#2 and ref#3). Even when PERCLOS is instantiated, the fatigue probability reaches 0.8639, which is still below the threshold 0.95. This indicates that one visual cue is not sufficient to conclude if the person is fatigued. On the other hand,

TABLE II

Prior Probability Table

Nodes	State	Probability	Notes
Random_noise	yes	0.15	average of [1] [53]
	no	0.85	
Light	on	0.13	average of [1] [53]
	off	0.87	
Heat	high	0.24	average of [1] [53]
	normal	0.76	
Humidity	high	0.19	average of [1] [53]
	normal	0.81	
Sleep_time	sufficient($> 6h$)	0.90	[53]
	loss($< 6h$)	0.1	
Napping	$> 30min.$	0.22	[53]
	No	0.78	
Anxiety	yes	0.28	average of [1] [53]
	no	0.72	
Sleep_disorder	yes	0.08	average of [1] [53]
	no	0.92	
Workload	heavy	0.15	[53]
	normal	0.85	
Time	drowsy_time	0.26	[1]
	Active_time	0.74	
Time_zone	changed	0.17	[1]
	no	0.83	
Temperature	high	0.15	average of [1] [53]
	normal	0.85	
Weather	abnormal	0.10	average of [1] [53]
	normal	0.90	
Noise	high	0.15	average of [1] [53]
	normal	0.85	
Work_type	tedious	0.2	average of [1] [53]
	normal	0.8	

TABLE III

THE INFERENCE RESULTS OF FATIGUE BAYESIAN NETWORK MODEL

Ref. No.	Evidences Instantiated	Fatigue Probability
1	No any evidence	0.5755
2	YawnFreq (high)	0.8204
3	PERCLOS (high)	0.8639
4	AECS (slow), Sleep time (insufficient), Time (drowsy time)	0.9545
5	YawnFreq (high), AECS(slow)	0.9552
6	Sleep time (insufficient), Time (drowsy time), Temperature (high)	0.8363

when combined with some contextual evidences, any visual parameter can lead to a high fatigue probability (ref#4). This demonstrates the importance of contextual information. The simultaneous observation of abnormal values for two visual parameters (ref#5) such as NodeFreq and PerSAC can lead to a fatigue probability higher than 0.95. This makes sense since they quantify fatigue from two different perspectives: one is gaze and the other is head movement. Any simultaneous observation of abnormal values of three or more visual parameters guarantees that the estimated fatigue probability exceeds the critical value. The simultaneous presence of several contextual evidences only leads to a high probability of fatigue, even in the absence of any visual evidence. These inference results, though preliminary and synthetic, demonstrate the utility of the proposed framework for predicting and modelling fatigue.

C. Interfacing with the Vision System

To perform real-time driver's fatigue monitoring, the visual module and the fusion module must be combined via an interface program such that the output of the vision system can be used by the fusion module to update its belief in fatigue in real time. Such an interface has been built. Basically, the interface program periodically (every 0.03 second)

examines the output of the vision module to detect any output change. If a change is detected, the interface program instantiates the corresponding observation nodes in the fusion module, which then activates its inference engine. The interface program then displays the inference result, plus current time as shown in Fig. 23. Besides displaying current fatigue level, the interface program also issues a warning beep when the fatigue level reaches a critical level.

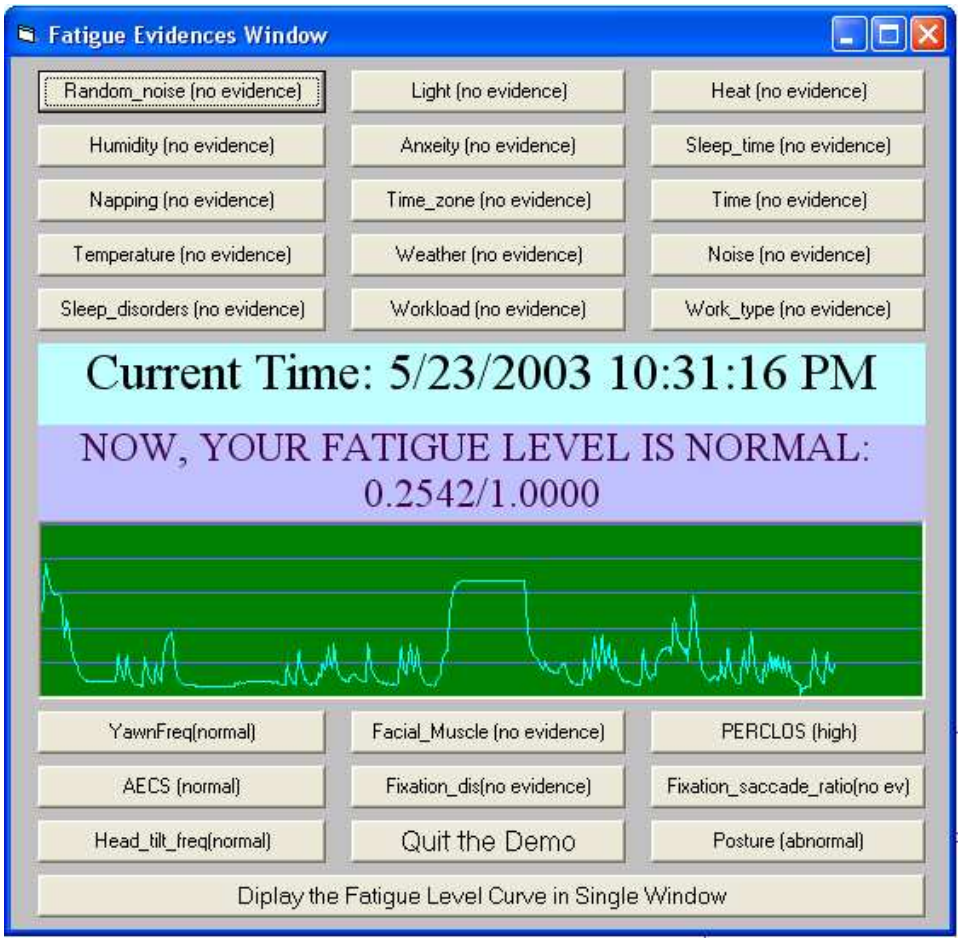


Fig. 23. The visual interface program panel. It displays the composite fatigue score over time.

VIII. System Validation

The last part of this research is to experimentally and scientifically demonstrate the validity of the computed fatigue parameters as well as the composite fatigue index. The validation consists of two parts. The first part involves the validation of the measurement accuracies of our computer vision techniques, and the second part studies the validity

of the fatigue parameters and the composite fatigue index that our system computes in characterizing fatigue.

A. Validation of the Measurement Accuracy

We present results to quantitatively characterize the measurement accuracies of our computer vision techniques in measuring eyelid movement, gaze, face pose and facial expressions. The measurements from our system are compared with those obtained either manually or using conventional instruments.

This section summarizes the eye detection and tracking accuracy of our eye tracker. For this study, we randomly selected an image sequence that contains $13,620$ frames, and manually identified the eyes in each frame. The manually labelled data serves as the ground-truth data, and they are compared with the eye detection results from our eye tracker. The study shows that our eye tracker is quite accurate, with a false alarm rate of 0.05% and a misdetection rate of 4.2%.

Further, we studied the positional accuracy of the detected eyes as well as the accuracy of the estimated pupil size (pupil axes ratio). The ground-truth data are produced by manually determining the locations of the eyes in each frame as well as the size of the pupil. The study shows that the detected eye positions match very well with manually detected eye positions, with a RMS position errors of 1.09 and 0.68 pixels for x and y coordinates respectively. The estimated size of pupil has an average RMS error of 0.0812.

Finally, we study the accuracy of the estimated face pose. To do so, we use a head-mount head tracker that tracks head movements. The output of the head-mount head tracker is used as the ground-truth. Quantitatively, the RMS errors for the pan and tilt angles are 1.92 degrees and 1.97 degrees respectively. This experiment demonstrates that our face pose estimation technique is sufficiently accurate.

B. Validation of Fatigue Parameters and the Composite Fatigue Score

To study the validity of the proposed fatigue parameters and that of the composite fatigue index, we performed a human subject study. The study included a total of 8 subjects. Two test bouts were performed for each subject. The first test was done when they first arrived in the lab at 9 pm and when they were fully alert. The second test was

performed about 12 hours later early in morning about 7 am the following day, after the subjects have been deprived of sleep for a total of 25 hours.

During the study, the subjects are asked to perform a TOVA (Test of Variables of Attention) test. The TOVA test consists of a 20-minute psychomotor test, which requires the subject to sustain attention and respond to a randomly appearing light on a computer screen by pressing a button. TOVA test was selected as the validation criterion because driving is primarily a vigilance task requiring psychomotor reactions, and psychomotor vigilance. The response time is used as a metric to quantify the subject's performance.

Figure 24 plots the average response times versus average PERCLOS measurements. The figure clearly shows the approximate linear correlation between PERCLOS and the TOVA response time. This experiment demonstrates the validity of PERCLOS in quantifying vigilance, as characterized by TOVA response time.

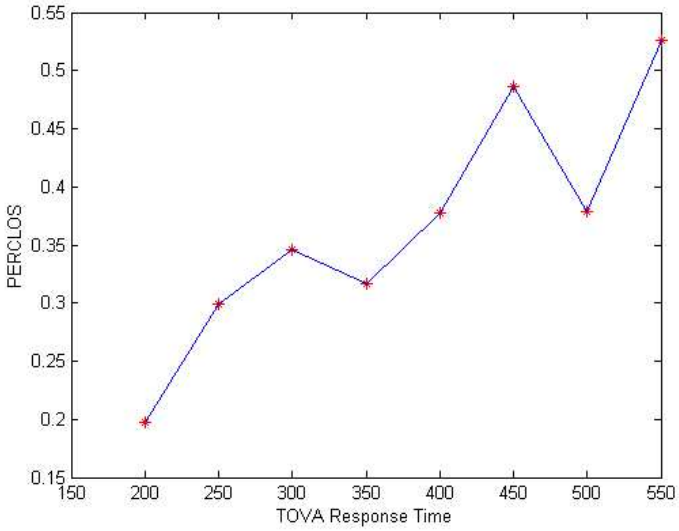


Fig. 24. PERCLOS versus TOVA response time. The two parameters are clearly correlated almost linearly. A larger PERCLOS measurement corresponds to a longer reaction time.

In addition, we want to demonstrate the correlation between PERCLOS and fatigue. For this, we compared the PERCLOS measurements for two bouts for the same individual. The comparison is shown in Figure 25, where it is clear that the PERCLOS measurements for the night bout (when the subject is alert) is significantly lower than the morning bout (subject is fatigued). This not only proves the validity of PERCLOS to characterize fatigue

but also proves the accuracy of our system in measuring PERCLOS. Similar results were obtained for other visual fatigue parameters we proposed.

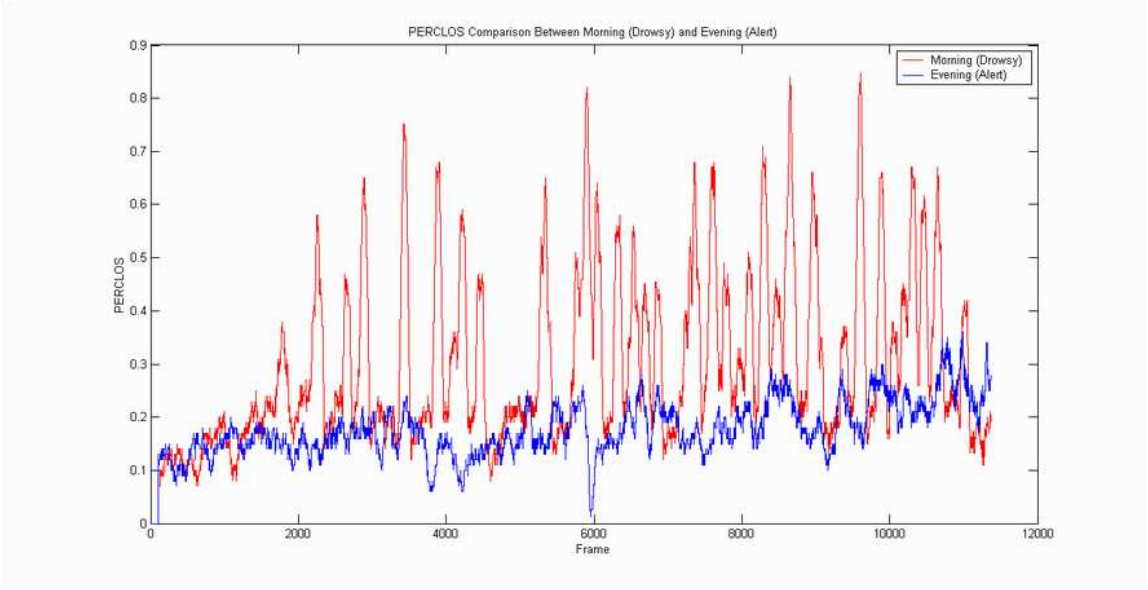


Fig. 25. PERCLOS measurements for evening (blue) and morning (red) bouts.

We also study the validity of the composite fatigue index our fatigue monitor computes. Figure 26 plots the TOVA performance versus the composite fatigue score. It clearly shows that the composite fatigue score (based on combining different fatigue parameters) highly correlates with the subject's response time.

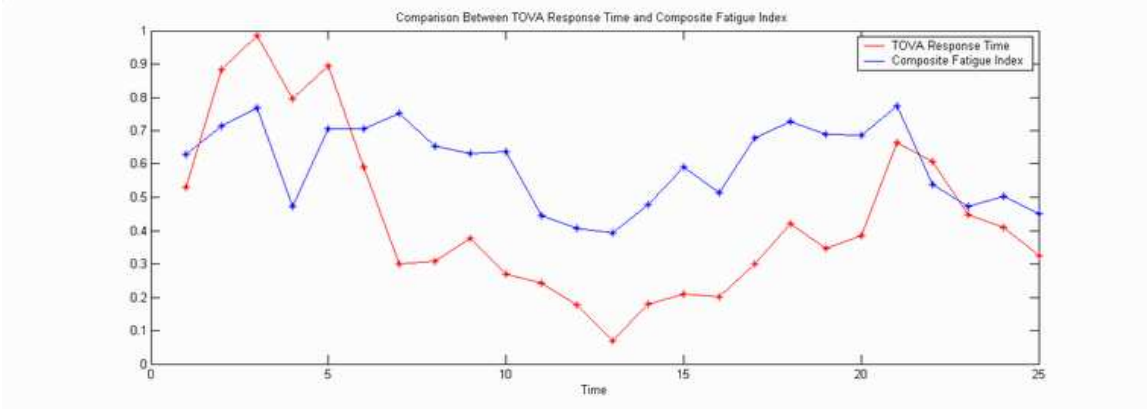


Fig. 26. The estimated composite fatigue index (blue) versus the normalized TOVA response time. The two curves track each other well.

It is clear that the two curves' fluctuations match well, proving their correlation and

co-variation, therefore proving the validity of the composite fatigue score in quantifying performance.

IX. Conclusion

Through research presented in this paper, we developed an non-intrusive prototype computer vision system for real-time monitoring a driver's vigilance. First, the necessary hardware and imaging algorithms are developed to simultaneously extract multiple visual cues that typically characterize a person's level of fatigue. Then, a probabilistic framework is built to model fatigue, which will systematically combine different visual cues and the relevant contextual information to produce a robust and consistent fatigue index.

These visual cues include eyelid movement, gaze, head movement, and facial expression. The main components of the system consist of a hardware system for real time acquisition of video images of the driver and various computer vision algorithms and their software implementations for real time eye tracking, eyelid movement parameters computation, eye gaze estimation, face pose discrimination and facial expression analysis. To effectively monitor fatigue, a BN model for fatigue is constructed to integrate these visual cues and relevant contextual information into one representative format.

Experiment studies in a real life environment with subjects of different ethnic backgrounds, different genders and ages were scientifically conducted to validate the fatigue monitoring system. The validation consists of two parts. The first part involves the validation of the measurement accuracy of our computer vision techniques, and the second part studies the validity of the fatigue parameters we compute in characterizing fatigue. Experiment results show that our fatigue monitor system is reasonably robust, reliable and accurate in characterizing human fatigue. It represents state of the art in real time on-line and non-intrusive fatigue monitoring.

References

[1] M. R. Rosekind, E. L. Co, K. B. Gregory, and D. L. Miller, "Crew factors in flight operations xiii: A survey of fatigue factors in corporate/executive aviation operations," *National Aeronautics and Space Administration, Ames Research Center Moffett Field, California 94035, NASA/TM-2000-209610*, 2000.

[2] H. Saito, T. Ishiwaka, M. Sakata, and S. Okabayashi, "Applications of driver's line of sight to automobiles-what can driver's eye tell," *Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, Aug. 1994*, pp. 21–26, 1994.

[3] H. Ueno, M. Kaneda, and M. Tsukino, "Development of drowsiness detection system," *Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, Aug. 1994*, pp. 15–20, 1994.

[4] S. Boverie, J. M. Leqellec, and A. Hirl, "Intelligent systems for video monitoring of vehicle cockpit," *1998 International congress and exposition ITS: Advanced controls and vehicle navigation systems*, pp. 1–5, 1998.

[5] M. Kaneda et al, "Development of a drowsiness warning system," *11th international conference on enhanced safety of vehicle, Munuch, 1994*, 1994.

[6] Reiner Onken, "Daisy, an adaptive knowledge-based driver monitoring and warning system," *Proceedings of 1994 Vehicle navigation and information systems conference, Yokohama, Japan, Aug. 1994*, pp. 3–10, 1994.

[7] J. Feraric, M. Kopf, and R. Onken, "Statistical versus neural bet approach for driver behavior description and adaptive warning," *11th European annual manual*, pp. 429–436, 1992.

[8] T. Ishii, M. Hirose, and H. Iwata, "Automatic recognition of driver's facial expression by image analysis," *Journal of JSAE*, vol. 41, no. 12, pp. 1398–1403, 1987.

[9] K. Yammamoto and S. Higuchi, "Development of a drowsiness warning system," *Journal of SAE of Japan*, vol. 46, no. 9, 1992.

[10] D.F. Dinges and M.M. Mallis, "Managing fatigue by drowsiness detection: Can technological promises be realised? in hartley, l.r. (ed.) managing fatigue in transportation.," *Proceedings of the Third International Conference on Fatigue and Transportation, Fremantle, Western Australia. Elsevier Science Ltd., Oxford UK*, 1998.

[11] S.G. Charlton and P.H. Baas, "Fatigue and fitness for duty of new zealand truck drivers.," *Road Safety Conference, Wellington, New Zealand.*, 1998.

[12] P. Simpson, "Review of computer based measurements of fitness for work.," *Prime Employee Assistance Services. Unpublished.*, 1998.

[13] T. Akerstedt and S. Folkard, "The three-process model of alertness and its extension to performance, sleep latency and sleep length.," *Chronobiology International. 14(2), pp115-123.*, 1997.

[14] G. Belenky, T.J. Balkin, D.P. Redmond, H.C. Sing, M.L. Thomas, D.R. Thorne, and N.J. Wesensten, "Sustained performance during continuous operations: The us armys sleep management system. in hartley, l.r. (ed.) managing fatigue in transportation.," *Proceedings of the Third International Conference on Fatigue and Transportation, Fremantle, Western Australia. Elsevier Science Ltd., Oxford UK.*, 1998.

[15] D. Dawson, N. Lamond, K. Donkin, and K. Reid, "Quantitative similarity between the cognitive psychomotor performance decrement associated with sustained wakefulness and alcohol intoxication. in hartley, l.r. (ed.) managing fatigue in transportation.," *Proceedings of the Third International Conference on Fatigue and Transportation, Fremantle, Western Australia. Elsevier Science Ltd., Oxford UK.*, 1998.

[16] P. Artaud, S. Planque, C. Lavergne, H. Cara, P. de Lepine, C. Tarriere, and B. Gueguen, "An on-board system for detecting lapses of alertness in car driving," *Proceedings of the Fourteenth International Conference on Enhanced Safety of Vehicles, Vol 1, Munich, Germany.*, 1994.

[17] N.A. Mabbott, M. Lydon, L. Hartley, and P. Arnold, "Procedures and devices to monitor operator alertness whilst operating machinery in open-cut coal mines.," *Stage 1: State-ofthe-art review. ARRB Transport Research Report RC 7433.*, 1999.

[18] C. Lavergne, P. De Lepine, P. Artaud, S. Planque, A. Domont, C. Tarriere, C. Arsonneau, X. Yu, A. Nauwink, C. Laurgeau, J.M. Alloua, R.Y. Bourdet, J.M. Noyer, S. Ribouchon, and C. Confer, "Results of the feasibility study of a system for warning of drowsiness at the steering wheel based on analysis of driver eyelid move-

ments.," *Proceedings of the Fifteenth International Technical Conference on the Enhanced Safety of Vehicles, Vol 1, Melbourne, Australia.*, 1996.

[19] E. Grandjean, "Fitting the task to the man (fourth edition)," *Taylor and Francis: London.*, 1988.

[20] D. Cleveland, "Unobtrusive eyelid closure and visual point of regard measurement system.," *Proceedings of the Ocular Measures of Driver Alertness Technical Conference, Herndon, Virginia, USA.*, 1999.

[21] R. J. (Ed.). Carroll, "Ocular measures of driver alertness: Technical conference proceedings.," *FHWA Technical Report No. FHWA-MC-99-136. Washington, DC: Federal Highway Administration, Office of Motor Carrier and Highway Safety.*, 1999.

[22] L. Hartley, T.J. Horberry, N. Mabbott, and G. Krueger, "Review of fatigue detection and prediction technologies," *National Road Transport Commission report. ISBN 0 642 54469 7*, 2000.

[23] S. Saito, "Does fatigue exist in a quantitative of eye movement ?," *Ergonomics*, vol. 35, 1992.

[24] Anon, "Perclos and eyetracking: Challenge and opportunity," *Technical report, Applied Science Laboratories, Bedford, MA 01730*, 1999.

[25] Anon, "Conference on ocular measures of driver alertness, washington dc," vol. April, 1999.

[26] David F. Dinges, M. Mallis, G. Maislin, and J. W. Powell, "Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management," *Department of Transportation Highway Safety publication 808 762*, vol. April, 1998.

[27] Richard Grace, "A dowsy driver detection system for heavy vehicles," *Conference on ocular measures of driver alertness*, vol. April, 1999.

[28] Dixon Cleveland, "Unobtrusive eyelid closure and visual of regard measurement system," *Conference on ocular measures of driver alertness*, vol. April, 1999.

[29] J. Fukuda, K. Adachi, M. Nishida, and Akutsu E, "Development of driver's drowsiness detection technology," *Toyota technical review*, vol. 45, pp. 34–40, 1995.

[30] J. H. Richardson, "The development of a driver alertness monitoring system," *Fatigue and Driving: Driver impairment, driver fatigue and driver simulation, Taylor and Francis; London, L. Harrtley (Ed)*, 1995.

[31] J. Dowdall, I. Pavlidis, and G. Bebis, "A face detection method based on multi-band feature extraction in the near-ir spectrum," *IEEE Workshop on Computer Vision Beyond the Visible Spectrum, December 14, 2001, Hawaii*, 2001.

[32] A. Haro, M. Flickner, and I. Essa, "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance," in *Proceedings IEEE CVPR 2000*, Hilton Head Island, South Carolina, 2000.

[33] W. Mullally and M. Betke, "Preliminary investigation of real-time monitoring of a driver in city traffic," *IEEE International Conference on Intelligent Vehicles,Dearborn, MI*, 2000.

[34] Zhiwei Zhu, Qiang Ji, Kikuo Fujimura, and Kuang chih Lee, "Combining kalman filtering and mean shift for real time eye tracking under active ir illumination," in *International Conference on Pattern Recognition*, Quebec, Canada, 2002.

[35] Zhiwei Zhu, Kikuo Fujimura, and Qiang Ji, "Real-time eye detection and tracking under various light conditions," in *Symposium on Eye Tracking Research and Applications*, New Orleans, LA, USA, 2002.

[36] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.

[37] J. Huang, D Ii, X. Shao, and H. Wechsler, "Pose discrimination and eye detection using support vector machines (svms)," in *Proceeding of NATO-ASI on Face Recognition: From Theory to Applications*, 1998, pp. 528–536.

[38] Q. Ji and X. Yang, "Real time visual cues extraction for monitoring driver vigilance," in *Proc. of International Workshop on Computer Vision Systems*, Vancouver, Canada, 2001.

[39] Zhiwei Zhu and Qiang Ji, "Active 3d face pose tracking from an uncalibrated monocular camera," *IEEE Transactions of Pattern Analysis and Machine Intelligence (submitted)*, 2003.

[40] Y. Ebisawa, "Unconstrained pupil detection technique using two light sources and the image difference method," *Visualization and Intelligent Design in Engineering*, pp. 79–89, 1989.

[41] Thomas E. Hutchinson, "Eye movement detection with improved calibration and speed," *United States Patent [19]*, , no. 4,950,069, 1988.

[42] Thomas E. Hutchinson, K.Preston White, J. R. Worthy, N. Martin, C. Kelly, Reichert Lisa, , and A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Transaction on systems,man,and cybernetics*, vol. 19, no. 6, pp. 1527–1533, 1989.

[43] Takehiko Ohno, Naoki Mukawa, and Atsushi Yoshikawa, "Freegaze: A gaze tracking system for everyday gaze interaction," *Eye Tracking Research and Applications Symposium, 25-27 March, New Orleans, LA, USA*, 2002.

[44] David Koons and Myron Flickner, "Ibm blue eyes project," *http://www.almaden.ibm.com/cs/blueeyes*.

[45] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner, "Frame-rate pupil detector and gaze tracker," *IEEE ICCV'99 FRAME-RATE WORKSHOP*, 1999.

[46] Y. Ebisawa, "Improved video-based eye-gaze detection method," *IEEE Transcations on Instrumentation and Measruement*, vol. 47, no. 2, pp. 948–955, 1998.

[47] Qiang Ji and Zhiwei Zhu, "Eye and gaze tracking for interactive graphic display," in *2nd International Symposium on Smart Graphics*, Hawthorne, NY, USA, 2002.

[48] Qiang Ji and Xiaojie Yang, "Real time 3d face pose discrimination based on active ir illumination (oral)," *International Conference on Pattern Recognition*, 2002.

[49] D. F. Specht, "A general regression neural network," *IEEE Transcations on Neural Networks*, vol. 2, pp. 568–576, 1991.

[50] T.S. Lee, "Image representation using 2d gabor wavelets," *IEEE Transactions of PAMI*, pp. 959–971, 1996.

[51] Haisong Gu, Qiang Ji, and Zhiwei Zhu, "Active facial tracking for fatigue detection," *IEEE Workshop on Applications of Computer Vision,Orlando, Florida*, 2002.

[52] M. I. Jordan, "Learning in graphical models," *MIT press*, 1999.

[53] E. L. Co, K. B. Gregory, J. M. Johnson, and M. R. Rosekind, "Crew factors in flight operations xi: A survey of fatigue factors in regional airline operations," *National Aeronautics and Space Administration, Ames Research Center Moffett Field, California 94035, NASA/TM-1999-208799*, 1999.

[54] P. Sherry, "Fatigue countermeasures in the railroad industry-past and current developments," *Counseling Psychology Program, Inter-modal Transportation Institute, University of Denver*, 2000.

[55] M. R. Rosekind, K. B. Gregory, E. L. Co, D. L. Miller, and D. F. Dinges, "Crew factors in flight operations xii: A survey of sleep quantity and quality in on-board crew rest facilities," *National Aeronautics and Space Administration, Ames Research Center Moffett Field, California 94035, NASA/TM-2000-209611*, 2000.

[56] F. V. Jensen, "Bayesian networks and decision graphs," *Statistics for Engineering and Information Science,Springer*, 2001.

[57] Microsoft Research Center, "online msbnx editor manual and software download," *http://research.microsoft.com/adapt/MSBNx/*.