Real-Time Nonintrusive Monitoring and Prediction of Driver Fatigue

Qiang Ji, Zhiwei Zhu, and Peilin Lan

Abstract—**This paper describes a real-time online prototype driver-fatigue monitor. It uses remotely located charge-coupled-device cameras equipped with active infrared illuminators to acquire video images of the driver. Various visual cues that typically characterize 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 employed characterize 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. This system was validated under real-life fatigue conditions with human subjects of different ethnic backgrounds, genders, and ages; with/without glasses; and under different illumination conditions. It was found to be reasonably robust, reliable, and accurate in fatigue characterization.**

Index Terms—**Driver vigilance, human fatigue, probabilistic model, visual cues.**

I. INTRODUCTION

THE EVER-INCREASING number of traffic accidents in the United States that are 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 perception, recognition, and vehicle-control abilities and, therefore, pose a serious danger to their own life and the lives of other people. Statistics show that a leading cause of fatal or injury-causing traffic accidents is due to drivers with a diminished vigilance level. In the trucking industry, 57% of fatal truck accidents are due to driver fatigue. It is the number one cause of heavy truck crashes. Seventy percent of American drivers report driving fatigued. The National Highway Traffic Safety Administration (NHTSA) [1] estimates that there are 100 000 crashes that are caused by drowsy drivers and result in more than 1500 fatalities and 71 000 injuries each year in U.S. With the ever-growing traffic conditions, this problem will further increase. For this reason, developing systems that actively monitoring a driver's level of vigilance and alerting the driver of any insecure driving conditions is essential for accident prevention.

Q. Ji and Z. Zhu are with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: qji@ecse.rpi.edu; zhuz@rpi.edu).

P. Lan is with the Department of Computer Science, University of Nevada at Reno, Reno, NV 89507 USA (e-mail: plan@cs.unr.edu).

Many efforts [2]–[20] have been reported in the literature for developing an active safety systems for reducing the number of automobile accidents due to reduced vigilance. These techniques can be classified into the following categories [21].

- Readiness-to-perform and fitness-for-duty technologies:
  These technologies [10] 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:
  These technologies use mathematical models to predict operator alertness and performance at different times based on interactions of sleep, Circadian, and related temporal antecedents of fatigue [12]–[14].
- 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, acceleration, braking, and gear changing [15]–[17].
- In-vehicle, online, operator-status-monitoring technologies:

The technologies in this category seek to real-time record some biobehavioral dimension(s) of an operator, such as features of the eyes, face, head, heart, brain activity, reaction time, etc., during driving [18]–[20]. According to the methods used for measurements, the technologies can be further divided into three types. The first employs electroencephalograph (EEG) measures, on which most successful equipments developed for offline fatigue monitoring are based. Also, there is an online version called "mind switch" that uses a headband device in which the electrodes are embedded to make contact with the driver's scalp to measure brain waves. Ocular measures are used in the second type, which is considered to be the most suitable way for online monitoring. So far, many eye-blinking, pupil-response, eye-closure, and eye-movement monitors have been developed. Other physiological/biobehavioral measures that are used in the third type include the 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 such as brain waves, heart rate, and pulse rate [9], [22]. Requiring physical contact with drivers (e.g., attaching electrodes), 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 [23] with a head-mount device are also encouraging. These techniques, although less intrusive, still are 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 changing course are good indicators of a driver's alertness level. While these techniques may be implemented nonintrusively, they are, nevertheless, subject to several limitations, including the vehicle type, driver experiences, and driving conditions [3].

People in fatigue exhibit certain visual behaviors that are easily observable from changes in facial features such as 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 noninvasive approach for monitoring fatigue is to assess a driver's vigilance level through the visual observation of his/her physical conditions using a remote camera and state-of-the-art technologies in computer vision. Techniques that use 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 on driver's vigilance [24] sponsored by the Department of Transportation (DOT), it is concluded that computer vision represents the most promising noninvasive technology to monitor driver's vigilance.

Many efforts on developing active real-time image-based fatigue-monitoring systems have been reported in the literature [2]–[6] [8], [9], [25],[26]–[32]. These efforts are primarily focused on detecting driver fatigue. For example, Ishii *et al.* [8] 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 eyelid movement. Their preliminary evaluations 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.

Despite the success of the existing approaches/systems for extracting characteristics of a driver using computer vision technologies, current efforts in this area, however, focus only on using a single visual cue, such as eyelid movement, line of sight, or head orientation, to characterize driver's state of alertness. The system that relies 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 a 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 a camera to monitor eye movement. Furthermore, the degree of eye openness may vary from person to person. Another potential problem with the use of a single visual cue is that the obtained visual feature is
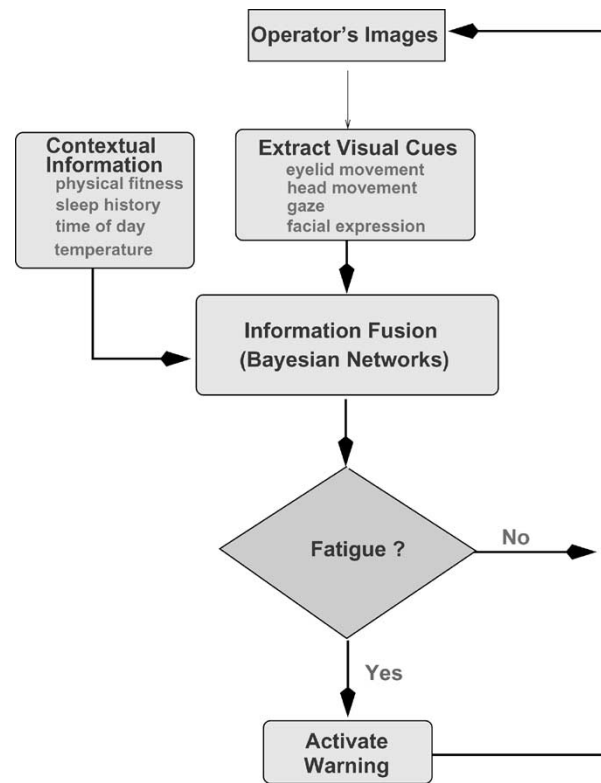

Fig. 1. Flowchart of the proposed driver-vigilance-monitoring system.

often ambiguous and, therefore, cannot always be indicative of one's mental conditions. For example, the irregular head movement or line of sight (such as briefly look back or at the minor) may yield false alarms for such a system.

All those visual cues, however imperfect they are individually, can provide an accurate characterization of a driver's level of vigilance if combined systematically. It is our belief that simultaneous extraction and the 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 information, as well as the relevant contextual information. The system we propose can simultaneously, nonintrusively, 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. Fig. 1 gives an overview of our driver-vigilance-monitoring system.

This paper consists of three parts. First, it focuses on a discussion of the computer vision algorithms and the hardware components that are necessary 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.
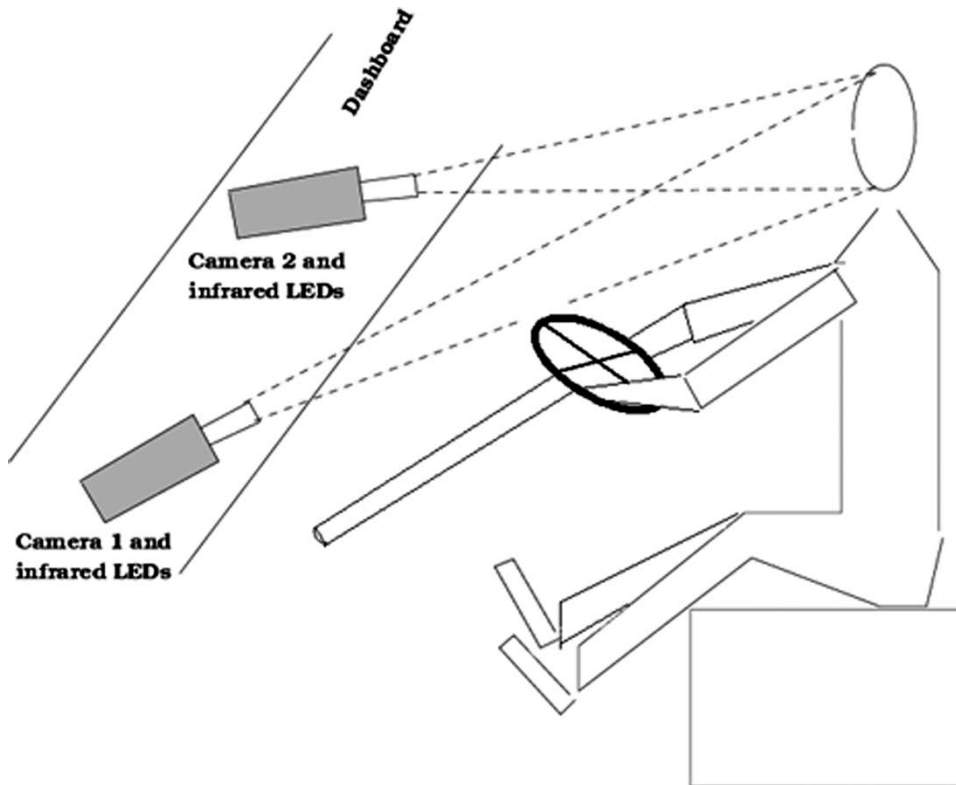
Fig. 2.   Overview of the driver-vigilance-monitoring system.

II. EYE DETECTION AND TRACKING

Fatigue monitoring starts with extracting visual parameters that typically characterize 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. Fig. 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-angle 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.

The goal of eye detection and tracking is for subsequent eyelid-movement monitoring, gaze determination, facial-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 facial orientations, based on combining the appearance-based methods and the active infrared (IR) illumination approach. Combining the respective strengths of different complementary techniques and overcoming their shortcomings, the proposed method uses active IR 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 eye appearance.
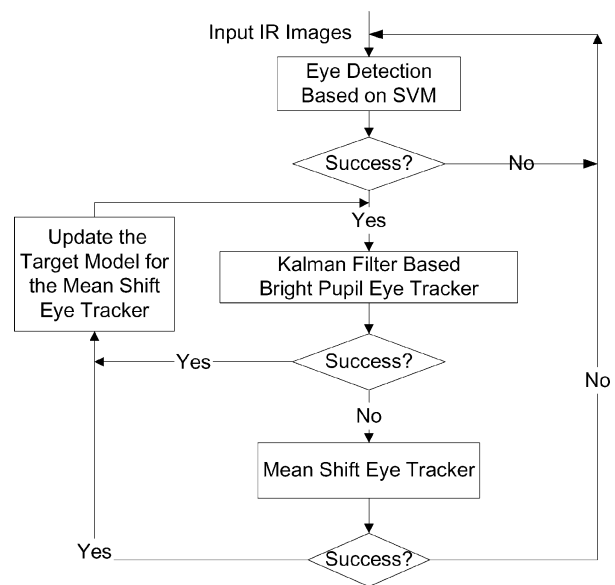

Fig. 3.   Combined eye-tracking flowchart.

Our method consists of two parts: eye detection and eye tracking. Fig. 3 summarizes our eye-detection and -tracking algorithm. Some of the ideas presented in this paper have been reported in [33] and [34]. In the sections that 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's face in real time. The acquired
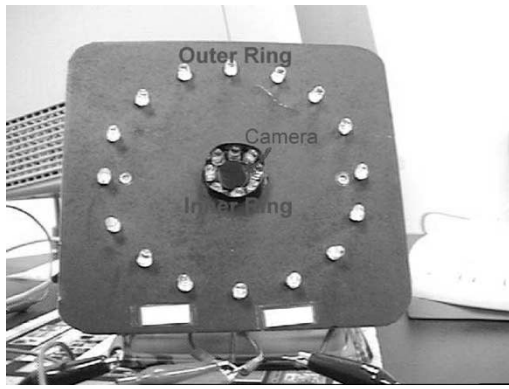
Fig. 4.   Actual photograph of the two-ring IR illuminator configuration.

images should have a 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-IR illuminator. The use of an IR 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 us to produce the bright/dark pupil effect, which constitutes the foundation for detection and tracking of the proposed visual cues. Third, since near IR 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 light-emitting diodes (LEDs), distributed evenly and symmetrically along the circumference of two coplanar concentric rings, as shown in Fig. 4. The center of both rings coincides with the camera optical axis. These IR LEDs will emit noncoherent IR energy in the 800–900-nm region of the spectrum.

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, which is controlled via a video decoder. An example of the bright/dark pupils is given in Fig. 5. Note that the glint, the small bright spot near the pupil, produced by cornea reflection of the IR light, appears on both the dark and bright pupil images.

B. Eye Detection

Eye-tracking starts with eyes detection. Fig. 6 gives a flowchart of the eye-detection procedure. Eye-detection is accomplished via pupil detection due to the use of active IR illumination.

Specifically, to facilitate pupil detection, we have developed a circuitry to synchronize the inner and outer rings 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 deinterlaced via a video decoder, producing the even and odd field images as shown in Fig. 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
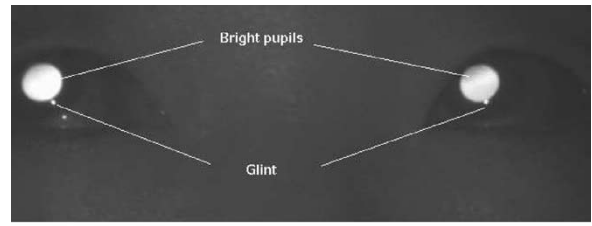
(a) bright pupils with glints
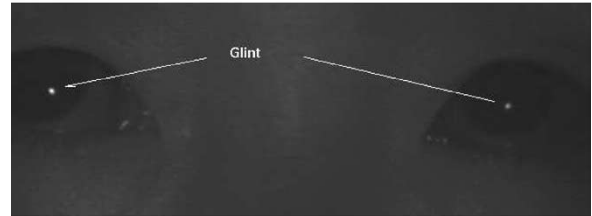
(b) dark pupils with glints

Fig. 5.   Bright and dark pupil images with glints.

Fig. 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 Fig. 8(a). From Fig. 8(a), we can see that there still are several nonpupil blobs left, because they are so similar in shape and size that we cannot distinguish them from the real pupil blobs, so we have to use other features.

From the dark pupil image, as shown in Fig. 8(b), we observed that each pupil is surrounded by the eye region, which has a 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 noneyes. We map the locations of the remaining binary blobs to the dark pupil images and then apply the support vector machine (SVM) classifier [35], [36] to automatically identify the binary blobs that correspond to eyes. A large number of training images including eyes and noneyes were used to train the SVM classifier. Fig. 8(c) shows that the SVM eye classifier correctly identifies the real eye regions as marked and removes the spurious ones. Details on our eye-detection algorithm may be found in [33].

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 [34]. 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. The Kalman filter pupil tracker works reasonably well under frontal face orientation with open eyes. 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
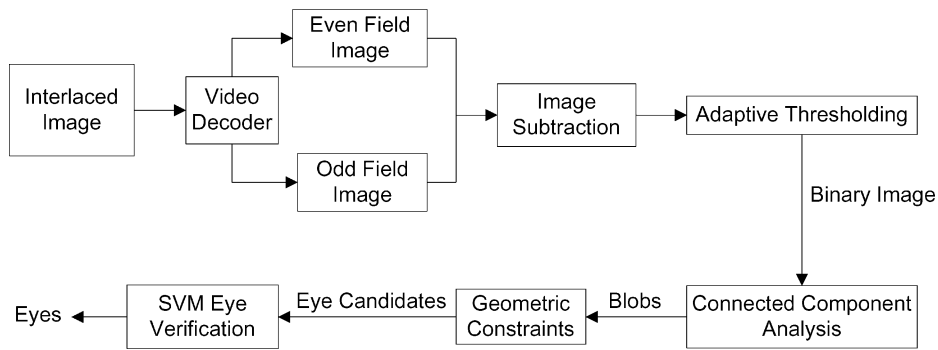
Fig. 6.   Eye-detection block diagram.

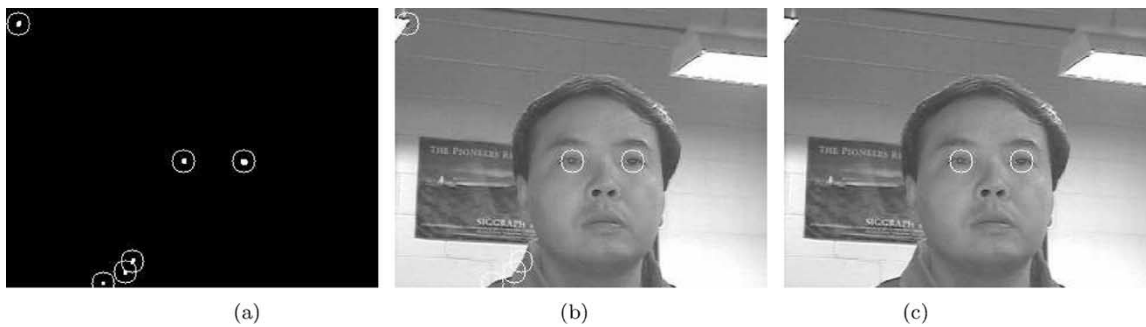Fig. 7.   (a) Even field image, (b) odd field image, and (c) the difference image.

Fig. 8.   (a) Thresholded difference image marked with possible pupil candidates, (b) image marked with possible eye candidates according to the positions of pupil candidates, and (c) image marked with identified eyes.

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 that tracks the eye regions according to the intensity statistical distributions of the eye regions and does not 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 oblique face orientations. Also, it is fast and handles noise well, but it does not have the capability to self-correction and, therefore, the errors tend to accumulate and propagate to subsequent frames as tracking progresses. Eventually, the tracker drifts away.

To overcome these limitations with the 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 and bright pupils) 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 error propagation with the 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 initiated first, 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 complement each other. The robustness of the eye tracker is improved significantly. The Kalman

filtering and mean shift-tracking algorithms are discussed in [33] and [37].

The eye-detection and -tracking algorithm is tested with different subjects under different facial orientations and illuminations. These experiments reveal that our algorithm is more robust than the conventional Kalman-filter-based bright pupil tracker, especially for the closed and partially occluded eyes due to the face orientations. Even under strong external illuminations, we have achieved good results. Video demonstrations 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 to compute the relevant eyelid-movement parameters. Here, we focus on two ocular measures to characterize the eyelid movement. The first 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 [25].

The eye-closure/opening speed is a good indicator of fatigue. It is defined as the amount of time needed to fully close or 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 [37].

The degree of eye opening is characterized by the shape of the pupil. It is observed that, as eyes close, the pupils start to get occluded by the eyelids and their shapes get more elliptical. So, we can use the ratio of pupil ellipse axes to characterize the 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 a 30-s 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 Fig. 9. Also, video demos are available at http://www.ecse.rpi.edu/~cvrl/Demo/demo.html.

IV. FACE (HEAD) ORIENTATION ESTIMATION

The facial (head) pose contains information about one's attention, gaze, and level of fatigue. Facial-pose determination is concerned with computation of the three-dimensional (3-D) facial orientation and position to detect head movements such 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 another directions (e.g., down or sideway) for an extended period of time, this is due to either fatigue or inattention. Facial-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.
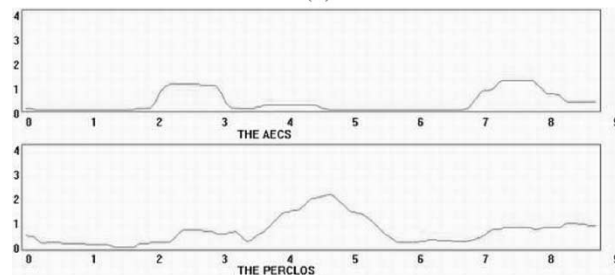
Fig. 9. (a) Detected eyes and pupils. (b) Plots for eyelid-movement parameters: The top displays the AECS parameter and the bottom displays the PERCLOS parameter.

We present a new technique to perform the two-dimensional (2-D) facial tracking and 3-D facial-pose estimation synchronously. In our method, a 3-D facial pose is tracked by Kalman filtering. The initial estimated 3-D pose is used to guide facial tracking in the image, which is subsequently used to refine the 3-D face pose estimation. Facial detection and pose estimation work together and benefit from each other. Weak perspective projection is assumed so that the face can be approximated as a planar object with facial features, such as eyes, nose, and mouth, located symmetrically on the plane. Fig. 10 summarizes our approach.

Initially, we automatically detect a fronto-parallel facial view based on the detected eyes [34] and some simple anthropometric statistics. The detected face region is used as the initial 3-D planar facial model. The 3-D facial pose is then tracked, starting from the fronto-parallel facial pose. During tracking, the 3-D face model is updated dynamically and the facial-detection and facial-pose estimation are synchronized and kept consistent with each other.

We will discuss our facial-pose-tracking algorithm briefly, as follows.

A. Automatic 3-D Facial Model and Pose Initialization

In our algorithm, we should have a fronto-parallel face to represent the initial facial model. This initialization is automatically accomplished by using the eye-tracking technique we have developed [34]. Specifically, the subject starts in the fronto-parallel facial pose position with the face facing directly at the camera, as shown in Fig. 11. The eye-tracking technique is then activated to detect the 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 automatically estimate the scope and location of the face in the image. Experiments show that our facial-detection method works well for all the
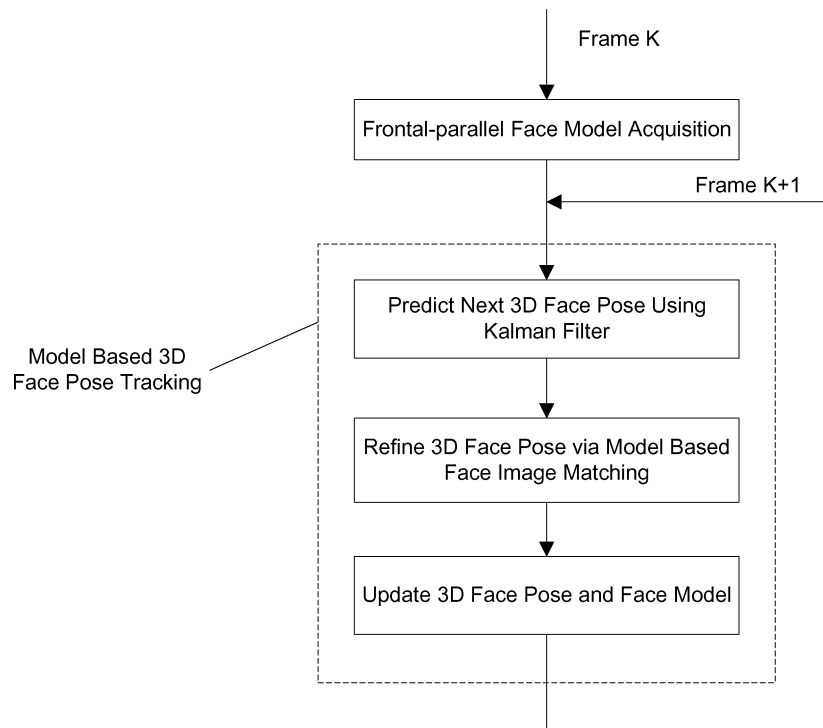
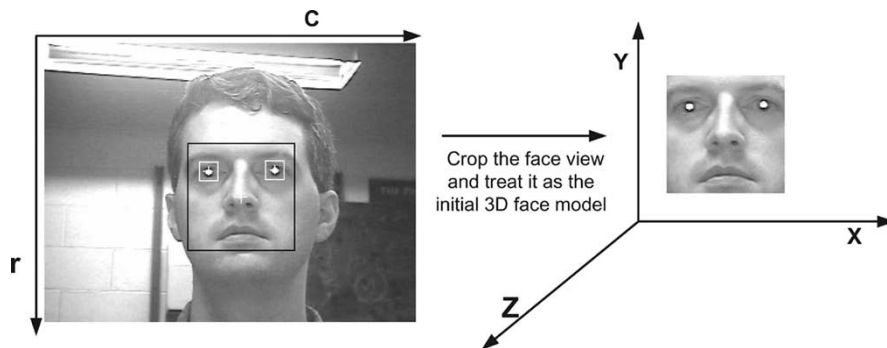Fig. 10.   Flowchart of facal-pose tracking.

Fig. 11.   Initial facial model.

faces we tested. An example of the detected frontal facial region is shown in Fig. 11. Once the facial region is decided, we will treat it as our initial facial model, whose pose parameters are used as initial facial-pose parameters.

Compared with the existing frontal facial-detection methods, ours takes full advantage of the detected eyes to guide the detection of the frontal face and is simple, robust, and automatic. In [38], we also demonstrate the tolerance of our facial initialization to slight deviations from the fronto-parallel facial pose and to perturbations of initial positions of the face.

B. Facial-Pose-Tracking Algorithm

Given the initial facial image and its pose in the first frame, the task of finding the facial location and facial pose in subsequent frames can be implemented as simultaneous 3-D facial-pose tracking and faceial detection described in [38].

Based on the detected facial pose in the previous frame, we can use Kalman filtering to predict the facial pose in the next frame, but the prediction based on Kalman filtering

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

assumes smooth facial movements. The prediction will be off significantly if the head undergoes a sudden rapid movement. To handle this issue, we propose to approximate the facial movement with eyes movement, since eyes can be reliably detected in each frame. Then, the final predicted facial pose is based on combining the one from Kalman filtering with the one from the eyes. The simultaneous use of Kalman filtering and eye's motion allows us to perform accurate facial-pose prediction even under significant or rapid head movements.
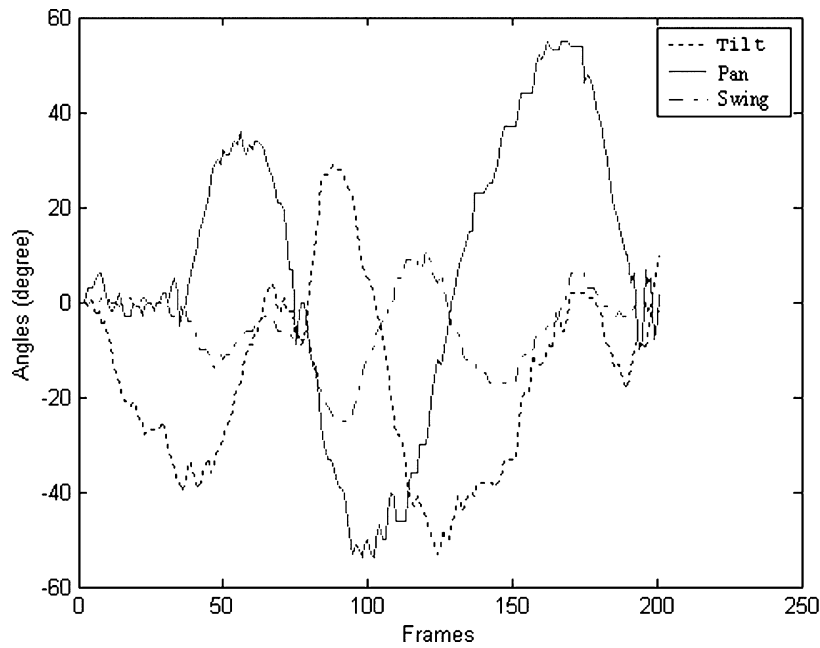
Fig. 13. Results of facial-pose tracking. The plots show the sequences of three estimated rotation angles through the image sequence.
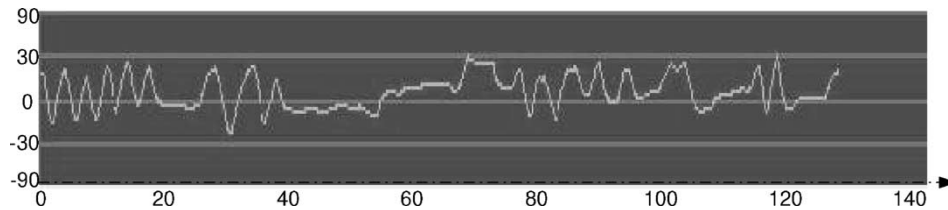

Fig. 14. Head-tilt monitoring over time (seconds).

Details on our facial-pose-estimation and -tracking algorithm may be found at [38].

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

To quantitatively characterize one's level of fatigue by facial pose, we introduce a new fatigue parameter called *NodFreq*, which measures the frequency of head tilts over time. Fig. 14 shows the running average of the estimated head tilts for a period of 140 s. 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 of a person's gaze is determined by two factors: the orientation of the face (facial pose) and the orientation of eye (eye gaze). Facial pose determines the global direction of the gaze, while eye gaze determines the local direction of the gaze. Global and local gazes 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 [39]–[45]. While contact free and nonintrusive, 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 keep the head still, because minor head movement can make these techniques fail. 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, recalibration is needed even for the same individual who already underwent the calibration procedure, whenever his/her head moved. This is because only the local gaze is accounted for, while global gaze due to facial pose is ignored.

In view of these limitations, we present a gaze-estimation approach [46] that accounts for both the local gaze computed from the ocular parameters and the global gaze computed from the head pose. The global gaze (facial pose) and local gaze (eye gaze) are combined together to obtain the precise gaze information of the user. Our approach, therefore, allows natural head
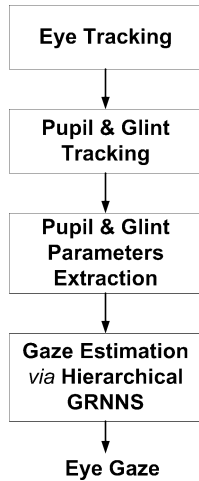
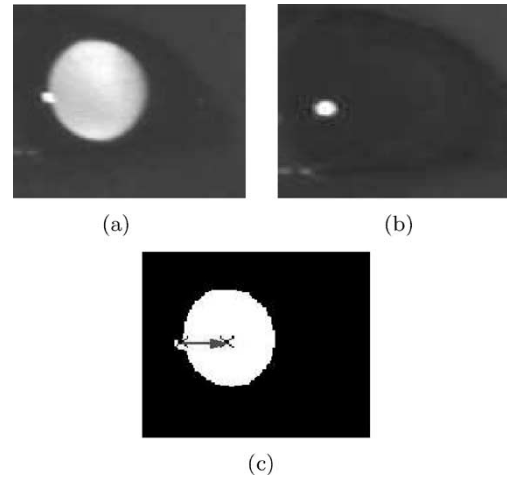Fig. 15. Major components of the proposed system.

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

movement while still estimating gaze accurately. Another effort is to make the gaze estimation calibration free. New or existing users who have moved do not need to 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 Fig. 15.

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

Given the detected glint and pupil, we can use their properties to obtain local and global gazes. Fig. 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 [47] shows that there exists a direct correlation between a 3-D facial pose and the geometric properties of the pupils. Specifically, pupil size, shape, and orientation vary with facial pose. Therefore, the pupil geometric properties are used to capture the 3-D face pose, which will serve as the global gaze.

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

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

the distance of the subject to the camera. r is used to account for out-of-plane facial rotation. The ratio should be close to 1 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 inplane facial rotation around the camera optical axis. Finally, (g_x, g_y) is used to account for the inplane head translation.

The use of these parameters accounts for both head and pupil movement, since their movements will introduce corresponding changes to these parameters, which 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) [48], which features fast training times, can model nonlinear 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 each predefined gaze region. After training, given an input vector, the GRNN can then approximate the user's actual gaze.
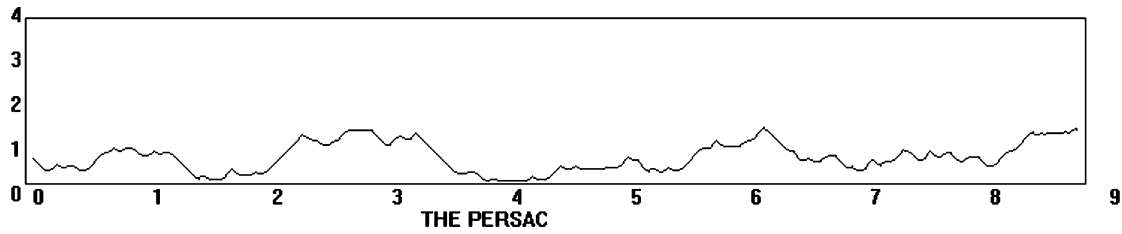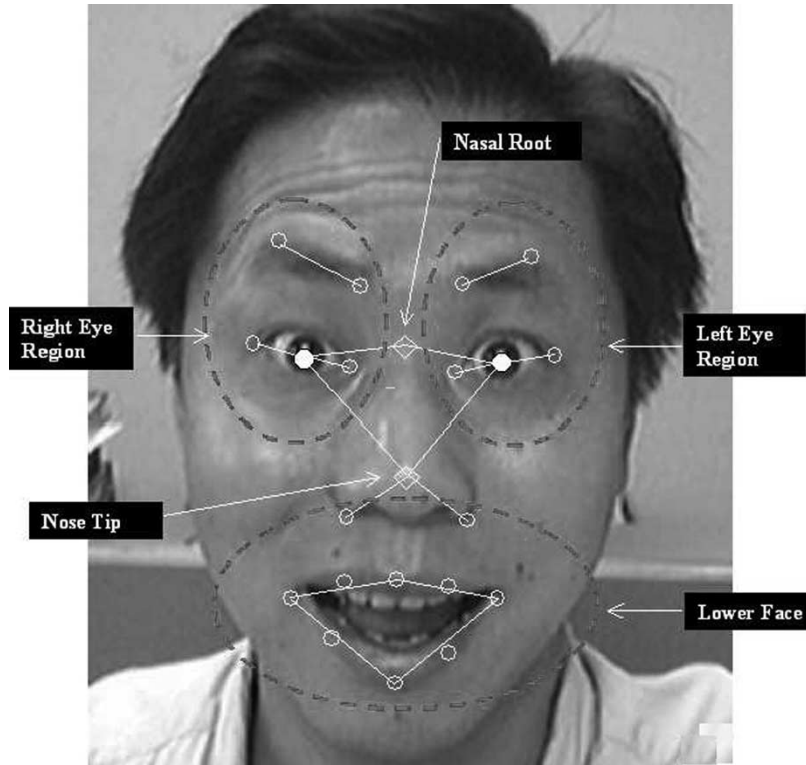
Fig. 17. Plot of PERSAC parameter over 30 .

Fig. 18. Facial features and local graphs.

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. Details on our gaze-estimation algorithm can be found in [46].

Given the gaze, we can compute a new fatigue parameter called *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 deliberate and conscious driver action to move an eye from one place to another. Therefore, it can measure the degree of alertness. The value of PERSAC is very small for a person in fatigue. Fig. 17 plots the PERSAC parameter over 30 s.

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 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 having lagging facial muscles, being expressionless, and yawning frequently.

Our recent research has led to the development of a feature-based facial-expression-analysis algorithm. The facial features around the 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 the eyes and mouth include enough information to capture these limited expressions. So, in our research, we focus on the facial features around the eyes and mouth. We use 22 fiducial features and three local graphs as the facial model (shown in Fig. 18).

In our method, the multiscale and multiorientation 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 [49]. 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 strong and reliable information that gives a rough location of where the 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 rapid head movement. The detected features and their spatial connections are used to characterize facial expressions. Details can be found in [50].

A series of experiments are conducted in [50] and good results are achieved under large head movements, self-occlusion, and different facial expressions. Fig. 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 movements to detect yawning. Yawning is detected if the features around te mouth significantly deviate from its closed configuration, especially in the vertical direction. There are eight tracked facial features around the mouth, as shown in Fig. 20. Also, as shown in Fig. 20, the height of the mouth is represented by the distance between the upper and lower lips 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 the mouth. Therefore, the openness of the mouth can be represented by the ratio of mouth height and width.

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

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 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, a 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.

Fig. 19.   Tracked facial features and local graphes.
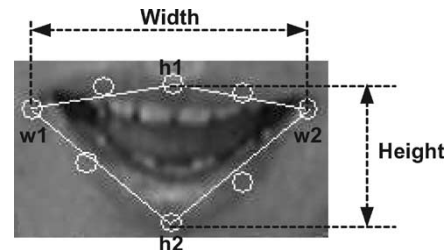

Fig. 20.   Facial features to be tracked around the mouth and the mouth width and height used to represent the openness of the mouth.

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) [51]. 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 and the child nodes.

A. Fatigue Modeling With BN

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 also are other hidden states that 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. 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 are sleep history, Circadian, work conditions, 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 and ACSE for eyelid movement, NodFreq for head movement, GAZEDIS and PERSAC for gaze movement, and YawnFreq
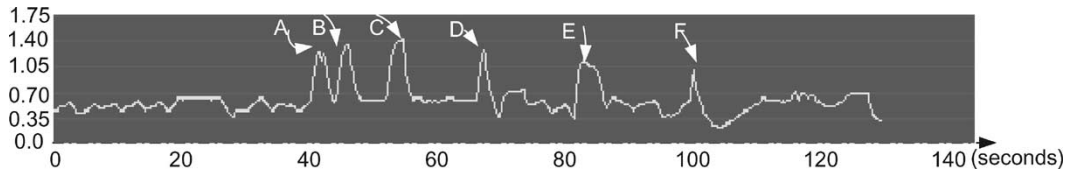
Fig. 21. Plot of the openness of the mouth over time. The bumps A, B, C, D, E, and F are the detected yawns.
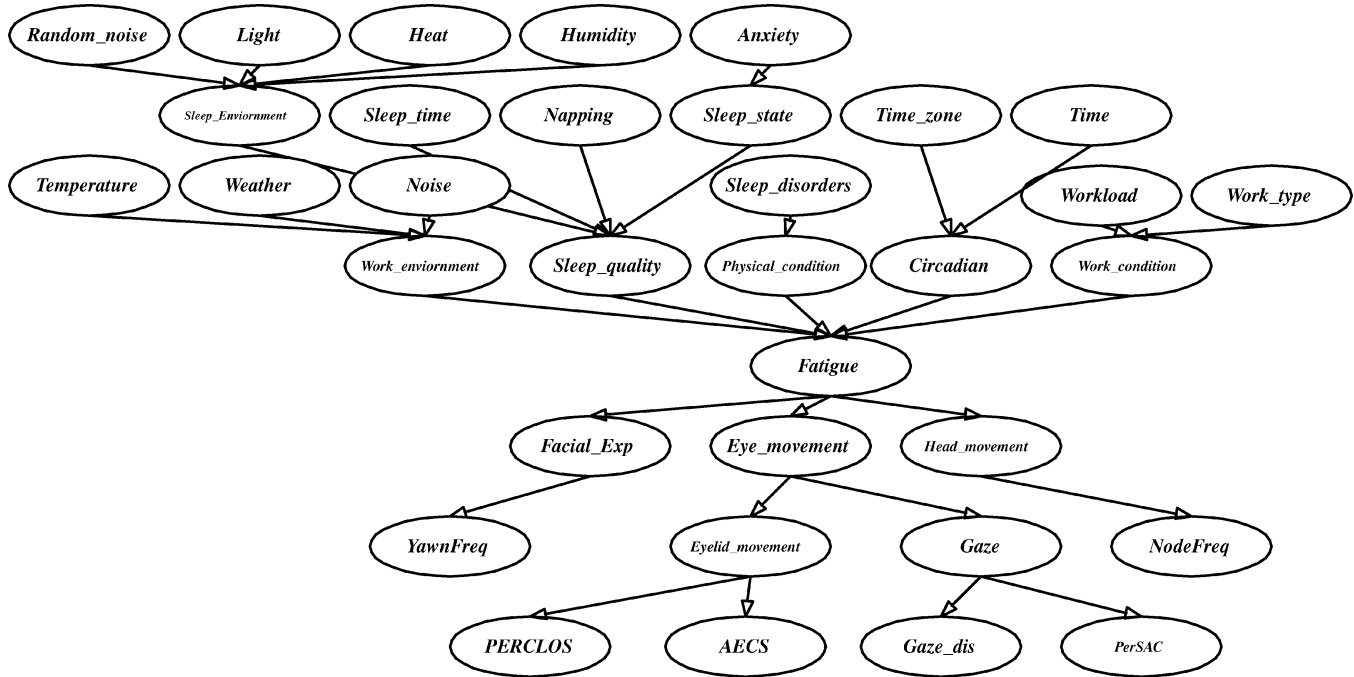

Fig. 22. BN model for monitoring human fatigue.

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.

B. 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], [52]–[54] provide additional data of this type. 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 conditional probabilities are, therefore, inferred from the available data

using the so-called noisy or principle [55]. Third, still some prior or conditional probabilities are lacking in our model, which are obtained by subjective estimation methods [55]. 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.

C. Fatigue Inference

Given the parameterized model, fatigue inference can then commence upon the arrival of visual evidences via belief propagation. MSBNX software [56] 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; 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 [1]. The observation of a single visual evidence does not usually provide a conclusive finding, since the estimated fatigue probability is less than the critical value 0.95, which is a hypothesized critical fatigue level[1] [2]–[3]. Even when PERCLOS is instantiated, the fatigue probability reaches 0.8639, which still is below the threshold of 0.95. This indicates that one visual cue is not sufficient to conclude if the person is fatigued. On the other hand, when combined with some contextual evidences, any visual parameter can lead to a high fatigue probability [4]. This

[1]This level may vary from application to application.

TABLE II
PRIOR PROBABILITY TABLE

Nodes	State	Prob.	Notes
Random_noise	yes	0.15	average
	no	0.85	of [1] [53]
Light	on	0.13	average
	off	0.87	of [1] [53]
Heat	high	0.24	average
	normal	0.76	of [1] [53]
Humidity	high	0.19	average
	normal	0.81	of [1] [53]
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
	no	0.72	[1] [53]
Sleep_disorder	yes	0.08	average
	no	0.92	of [1] [53]
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
	normal	0.85	of [1] [53]
Weather	abnormal	0.10	average
	normal	0.90	of [1] [53]
Noise	high	0.15	average
	normal	0.85	of [1] [53]
Work_type	tedious	0.2	average
	normal	0.8	of [1] [53]

TABLE III
INFERENCE RESULTS OF FATIGUE BN MODEL

Ref. No.	Evidences Instantiated	Fatigue Prob.
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

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 s) 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 a current fatigue level, the interface program also issues a warning beep when the fatigue level reaches a critical level.

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 involves the validation of the measurement accuracies of our computer vision techniques and the second 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, facial 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. This manually labeled data serves as the ground-truth data and are compared with the eye-detection results from our eye tracker. This 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. This study shows that the detected eye

demonstrates the importance of contextual information. The simultaneous observation of abnormal values for two visual parameters [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 modeling fatigue.

D. Interfacing with the Vision System

To perform real-time driver-fatigue monitoring, the visual and fusion modules must be combined via an interface program such that the output of the vision system can be used by the

Fig. 23. Visual interface program panel, which displays the composite fatigue score over time.
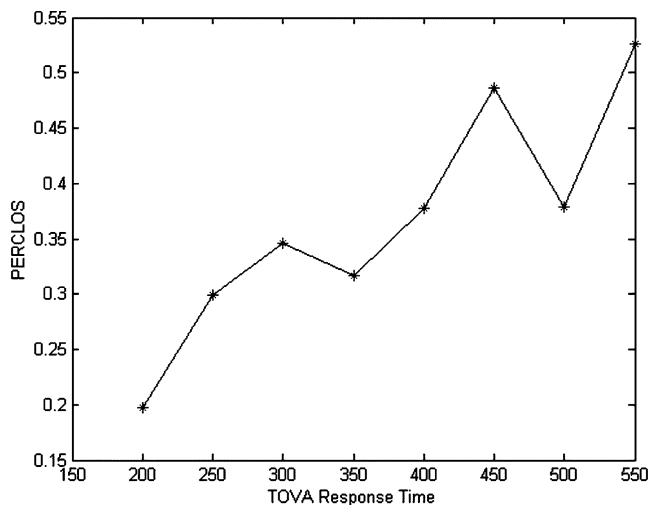

Fig. 24. PERCLOS versus the TOVA response time. The two parameters are clearly correlated almost linearly. A larger PERCLOS measurement corresponds to a longer reaction time.

positions match very well with manually detected eye positions, with a root-mean-square (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 facial 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 $1.92°$ and $1.97°$, respectively. This experiment demonstrates that our facial-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 eight subjects and two test bouts were performed for each subject. The first test was done when they first arrived in the laboratory at 9:00 PM and when they were fully alert. The second test was performed about 12 hours later, early in morning at about 7:00 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 test of variables of attention (TOVA) test. The TOVA test consists of a 20-min psychomotor test, which requires the subject to sustain attention and respond to a randomly appearing light on a computer screen by pressing a button. The TOVA test was selected
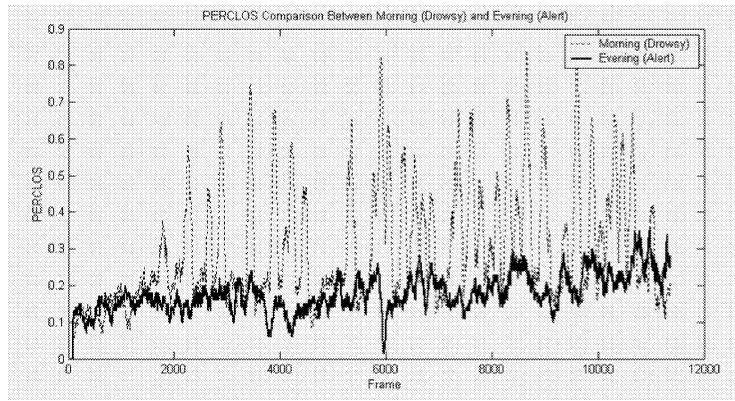
Fig. 25.   PERCLOS measurements for evening (solid line) and morning (dotted line) bouts.
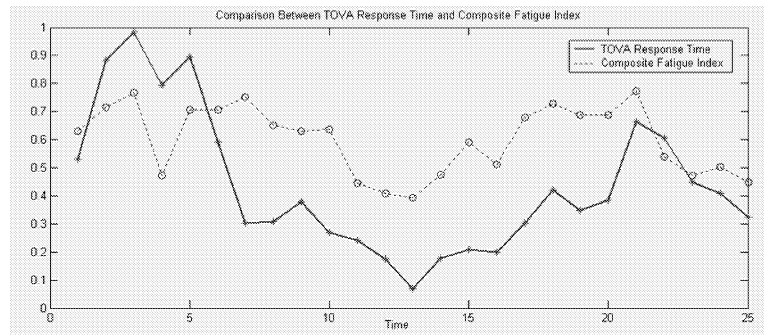

Fig. 26.   Estimated composite fatigue index (dotted line) versus the normalized TOVA response time (solid line). The two curves track each other well.

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.

Fig. 24 plots the average response times versus average PERCLOS measurements. This 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 the TOVA response 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. This comparison is shown in Fig. 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.

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

It is clear that the two curves' fluctuations match well, proving their correlation and covariation and, therefore, proving the validity of the composite fatigue score in quantifying performance.

IX. CONCLUSION

Through research presented in this paper, we developed an nonintrusive prototype computer vision system for real-time monitoring of 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 systematically combines different visual cues and the relevant contextual information to produce a robust and consistent fatigue index.

These visual cues characterize eyelid movement, gaze, head movement, and facial expression. The main components of the system consist of a hardware system for the 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, facial-pose determination, 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, genders, and ages were scientifically conducted to validate the fatigue-monitoring system. The validation consists of two parts. The first involves the validation of the measurement accuracy of our computer vision techniques and the second studies the validity of the fatigue parameters that 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

the state of the art in real-time, online, and nonintrusive 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, CA, NASA/TM-2000-209 610, 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," in *Proc. Vehicle Navigation Information Systems Conf.*, Yokohama, Japan, Aug. 1994, pp. 21–26.

[3] H. Ueno, M. Kaneda, and M. Tsukino, "Development of drowsiness detection system," in *Proc. Vehicle Navigation Information Systems Conf.*, Yokohama, Japan, Aug. 1994, pp. 15–20.

[4] S. Boverie, J. M. Leqellec, and A. Hirl, "Intelligent systems for video monitoring of vehicle cockpit," in *Proc. Int. Congr. Expo. ITS: Advanced Controls Vehicle Navigation Systems*, 1998, pp. 1–5.

[5] M. K. *et al.*, "Development of a drowsiness warning system," presented at the Proc. 11th Int. Conf. Enhanced Safety Vehicle, Munich, Germany, 1994.

[6] R. Onken, "Daisy, an adaptive knowledge-based driver monitoring and warning system," in *Proc. Vehicle Navigation Information Systems Conf.*, Yokohama, Japan, Aug. 1994, pp. 3–10.

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

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

[9] K. Yammamoto and S. Higuchi, "Development of a drowsiness warning system," *J. Soc. Automotive Eng. Japan*, vol. 46, no. 9, pp. 127–133, 1992.

[10] D. Dinges and M. Mallis, "Managing fatigue by drowsiness detection: Can technological promises be realized? in managing fatigue in transportation," in *Managing Fatigue in Transportation: Selected Papers from the 3rd Fatigue in Transportation Conference, Fremantle, Western Australia*, L. R. Hartley, Ed. Oxford, U.K.: Elsevier, 1998, pp. 209–229.

[11] S. Charlton and P. Baas, "Fatigue and fitness for duty of New Zealand truck drivers," presented at the Road Safety Conf., Wellington, New Zealand, 1998.

[12] T. Akerstedt and S. Folkard, "The three-process model of alertness and its extension to performance, sleep latency and sleep length," *Chronobio. Int.*, vol. 14, no. 2, pp. 115–123, 1997.

[13] G. Belenky, T. Balkin, D. Redmond, H. Sing, M. Thomas, D. Thorne, and N. Wesensten, "Sustained performance during continuous operations: The us armys sleep management system. In managing fatigue in transportation," in *Managing Fatigue in Transportation: Selected Papers from the 3rd Fatigue in Transportation Conference, Fremantle, Western Australia*, L. R. Hartley, Ed. Oxford, U.K.: Elsevier, 1998.

[14] 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 managing fatigue in transportation," in *Managing Fatigue in Transportation: Selected Papers from the 3rd Fatigue in Transportation Conference, Fremantle, Western Australia*, L. R. Hartley, Ed. Oxford, U.K.: Elsevier, 1998, pp. 231–256.

[15] 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," presented at the 14th Int. Conf. Enhanced Safety of Vehicles, vol. 1, Munich, Germany, 1994.

[16] N. 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-of-the-art review," ARRB Transport Res. Rep. RC 7433, 1999.

[17] C. Lavergne, P. De Lepine, P. Artaud, S. Planque, A. Domont, C. Tarriere, C. Arsonneau, X. Yu, A. Nauwink, C. Laurgeau, J. Alloua, R. Bourdet, J. 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 movements," presented at the Proc. 15th Int. Tech. Conf. Enhanced Safety Vehicles, vol. 1, Melbourne, Australia, 1996.

[18] E. Grandjean, *Fitting the Task to the Man*, 4th ed. London, U.K.: Taylor & Francis, 1988.

[19] D. Cleveland, "Unobtrusive eyelid closure and visual point of regard measurement system," presented at the Proc. Ocular Measures Driver Alertness Tech. Confe., Herndon, VA, Apr. 26–27, 1999.

[20] R. J. E Carroll, "Ocular measures of driver alertness technical conference proceedings," Federal Highway Administration, Office of Motor Carrier and Highway Safety, Washington, DC, FHWA Tech. Rep. FHWA-MC-99-136, 1999.

[21] L. Hartley, T. Horberry, N. Mabbott, and G. Krueger, *Review of Fatigue Detection and Prediction Technologies*. Melbourne, Australia: National Road Transport Commission, 2000.

[22] S. Saito, "Does fatigue exist in a quantitative of eye movement ?," *Ergonomics*, vol. 35, pp. 607–615, 1992.

[23] Appl. Sci. Lab., "PERCLOS and eyetracking: Challenge and opportunity," Tech. Rep., Appl. Sci. Lab., Bedford, MA, 1999.

[24] *Conf. Ocular Measures of Driver Alertness*, Washington, DC, Apr. 26–27, 1999.

[25] D. 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," Dept. Transp. Highway Safety, pub. 808 762, 1998.

[26] R. Grace, "A drowsy driver detection system for heavy vehicles," presented at the Conf. Ocular Measures Driver Alertness, Apr. 1999.

[27] D. Cleveland, "Unobtrusive eyelid closure and visual of regard measurement system," presented at the Conf. Ocular Measures Driver Alertness, Apr. 1999.

[28] J. Fukuda, K. Adachi, M. Nishida, and A. E. , "Development of driver's drowsiness detection technology," *Toyota Techn. Rev.*, vol. 45, pp. 34–40, 1995.

[29] J. H. Richardson, "The development of a driver alertness monitoring system," in *Fatigue and Driving: Driver Impairment, Driver Fatigue and Driver Simulation*, L. Harrtley, Ed. London, U.K.: Taylor & Francis, 1995.

[30] J. Dowdall, I. Pavlidis, and G. Bebis, "A face detection method based on multi-band feature extraction in the near-IR spectrum," presented at the IEEE Workshop Computer Vision Beyond Visible Spectrum, Honolulu, HI, Dec. 14, 2001.

[31] A. Haro, M. Flickner, and I. Essa, "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance," presented at the Proc. IEEE Conf. Computer Vision and Pattern Recognition, Hilton Head Island, SC, 2000.

[32] W. Mullally and M. Betke, "Preliminary investigation of real-time monitoring of a driver in city traffic," presented at the IEEE Int. Conf. Intelligent Vehicles, Dearborn, MI, 2000.

[33] Z. Zhu, Q. Ji, K. Fujimura, and K. c. Lee, "Combining Kalman filtering and mean shift for real time eye tracking under active IR illumination," presented at the Int. Conf. Pattern Recognition, Quebec, PQ, Canada, 2002.

[34] Z. Zhu, K. Fujimura, and Q. Ji, "Real-time eye detection and tracking under various light conditions," presented at the Symp. Eye Tracking Research Applications, New Orleans, LA, 2002.

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

[36] J. Huang, D. Ii, X. Shao, and H. Wechsler, "Pose discrimination and eye detection using support vector machines (SVM's)," in *Proc. NATO Advanced Study Institute (ASI) Face Recognition: From Theory to Applications*, 1998, pp. 528–536.

[37] Q. Ji and X. Yang, "Real time visual cues extraction for monitoring driver vigilance," presented at the Proc. Int. Workshop Computer Vision Systems, Vancouver, BC, Canada, 2001.

[38] Z. Zhu and Q. Ji, "3D face pose tracking from an uncalibrated monocular camera," presented at the Int. Conf. Pattern Recognition, Cambridge, U.K., 2004.

[39] Y. Ebisawa, "Unconstrained pupil detection technique using two light sources and the image difference method," in *Visualization and Intelligent Design in Engineering and Architecture II*. Boston, MA: Computational Mechanics, 1989, pp. 79–89.

[40] T. E. Hutchinson, "Eye Movement Detection with Improved Calibration and Speed," U.S. Patent 4 950 069, 1988.

[41] T. E. Hutchinson, K. White, J. R. Worthy, N. Martin, C. Kelly, R. Lisa, and A. Frey, "Human-computer interaction using eye-gaze input," *IEEE Trans. Syst., Man, Cybern.*, vol. 19, pp. 1527–1533, Nov./Dec. 1989.

[42] T. Ohno, N. Mukawa, and A. Yoshikawa, "Freegaze: A gaze tracking system for everyday gaze interaction," presented at the Eye Tracking Research Applications Symp., New Orleans, LA, Mar. 25–27, 2002.

[43] D. Koons and M. Flickner. IBM Blue Eyes Project. [Online]. Available: http://www.almaden.ibm.com/cs/blueeyes

[44] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner, "Frame-rate pupil detector and gaze tracker," presented at the IEEE Int. Conf. Computer Vision (ICCV) Frame-Rate Workshop, 1999.

[45] Y. Ebisawa, "Improved video-based eye-gaze detection method," *IEEE Trans. Instrum. Meas.*, vol. 47, pp. 948–955, Apr. 1998.

[46] Q. Ji and Z. Zhu, "Eye and gaze tracking for interactive graphic display," presented at the 2nd Int. Symp. Smart Graphics, Hawthorne, NY, 2002.

[47] Q. Ji and X. Yang, "Real time 3D face pose discrimination based on active IR illumination (oral)," presented at the Int. Conf. Pattern Recognition, 2002.

[48] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Networks*, vol. 3, pp. 568–576, Nov. 1991.

[49] T. Lee, "Image representation using 2D Gabor wavelets," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 959–971, Oct. 1996.

[50] H. Gu, Q. Ji, and Z. Zhu, "Active facial tracking for fatigue detection," presented at the IEEE Workshop Applications of Computer Vision, Orlando, FL, 2002.

[51] M. I. Jordan, *Learning in Graphical Models*. Cambridge, MA: MIT Press, 1999.

[52] 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," Nat. Aeron. Space Admin., Ames Res. Center, Moffett Field, CA, NASA/TM-1999-208 799, 1999.

[53] P. Sherry, *Fatigue Countermeasures in the Railroad Industry-Past and Current Developments*. Denver, CO: Univ. Denver, Intermodal Transportation Institute, Counseling Psychology Program, 2000.

[54] 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," Nat. Aeron. Space Admin., Ames Res. Center, Moffett Field, CA, NASA/TM-2000-209 611, 2000.

[55] F. V. Jensen, "Bayesian networks and decision graphs," in *Statistics for Engineering and Information Science*. New York: Springer-Verlag, 2001.

[56] Online MSBNx Editor Manual and Software Download. M. R. Center. [Online]. Available: http://research.microsoft.com/adapt/MSBNx/

Qiang Ji received the Ph.D. degree in electrical engineering from the University of Washington, Seattle, in 1998.

He currently is an Assistant Professor with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY. He has published more than 60 papers in refereed journals and conferences. His research has been funded by the National Science Foundation, National Istitute of Health, Air Force Office for Scientific Research (AFOSR), Office of Naval Research (ONR), Defence Advanced Research Projects Agency (DARPA), Army Research Office (ARO), and Honda. His research interests include computer vision and probabilistic reasoning for decision-making and information fusion, pattern recognition, and robotics. His latest research focuses on applying computer vision and probabilistic reasoning theories to human–computer interaction, including human fatigue monitoring, user affect modeling and recognition, and active user assistance.

Zhiwei Zhu received the M.S. degree in computer science from the University of Nevada, Reno, in August 2002. He currently is working toward the Ph.D. degree in the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY.

His research interests include computer vision, pattern recognition and human computer interaction.

Peilin Lan received the M.S. degree in mineral engineering from Kunming University of Science and Technology, Kunming, China, in 1987 and the M.S. degree in computer science and the Ph.D. degree in metallurgical engineering from University of Nevada, Reno, in 2001 and 2002, respectively. He is currently working toward the Ph.D. degree in the Department of Computer Science and Engineering, University of Nevada, Reno.

His research interests in computer science and engineering, including uncertainty reasoning and Bayesian networks and their applications to information fusion, data mining, and machine learning. He is also interested in the application of artificial intelligence to mining and metallurgical engineering.