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) 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.

Mathematical models of alertness dynamics joined with ambulatory technologies:
1) 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.
2) 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.
3) In-vehicle, online, operator-status-monitoring technologies:

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. 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 movement 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.

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.

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.

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. 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 detect the eyes. After detecting the eyes, the distance between the detected eyes. 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.

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. In our algorithm, we should have a front-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 front-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 e_{yes} between two eyes. Then, the distance between the detected eyes. 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. 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. 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.

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. 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. 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.

Gaze Estimation

Our gaze-estimation algorithm consists of three parts:
pupil-glint detection and tracking, gaze calibration, and
gaze mapping. 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
dark pupil effect, The pupil-detection and -tracking
technique can be used to detect and track glint from the
dark images

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: \(A_x, A_y, r, \theta, g_y, \) and \(g_x.\) \(A_x\) and \(A_y\) are the
pupil-glint displacement. the ratio of the major-to-minor
axes of the ellipse that fits to the pupil. \(\theta\) 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. \(A_x\) and \(A_y\) account for the relative movement
between the glint and the pupil, representing the local gaze.

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), 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

\[
g = [\Delta x \quad \Delta y \quad r \quad \theta \quad y_x \quad g_x]
\]

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.

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.
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. 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.

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.

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). 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.

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, provide additional data of this type.

Fatigue Inference
Given the parameterized model, fatigue inference can then commence upon the arrival of visual evidences via belief propagation. MSBNX software is used to perform the inference and both top-down and bottom-up belief propagations are performed.

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 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, 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.

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 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 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.

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.

Plots the average response times versus average PERCLOS 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, 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. 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 co variation and, therefore, proving the validity of the composite fatigue score in quantifying performance.

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