Detection of Driver Fatigue Caused by Sleep Deprivation

Ji Hyun Yang, Zhi-Hong Mao, *Member, IEEE*, Louis Tijerina, Tom Pilutti, Joseph F. Coughlin, and Eric Feron

Abstract—This paper aims to provide reliable indications of driver drowsiness based on the characteristics of driver–vehicle interaction. A test bed was built under a simulated driving environment, and a total of 12 subjects participated in two experiment sessions requiring different levels of sleep (partial sleep-deprivation versus no sleep-deprivation) before the experiment. The performance of the subjects was analyzed in a series of stimulus-response and routine driving tasks, which revealed the performance differences of drivers under different sleep-deprivation levels. The experiments further demonstrated that sleep deprivation had greater effect on rule-based than on skill-based cognitive functions: when drivers were sleep-deprived, their performance of responding to unexpected disturbances degraded, while they were robust enough to continue the routine driving tasks such as lane tracking, vehicle following, and lane changing. In addition, we presented both qualitative and quantitative guidelines for designing drowsy-driver detection systems in a probabilistic framework based on the paradigm of Bayesian networks. Temporal aspects of drowsiness and individual differences of subjects were addressed in the framework.

Index Terms—Bayesian networks (BNs), camouflage, drowsy driving, sleep deprivation, stimulus-response tasks, tracking tasks.

I. INTRODUCTION

UNTIL recently, most safety-related research has focused on methods to reduce damage caused by transportation accidents while they are occurring or after they happen. Passive safety systems such as seat belts, airbags, and crashworthy body structures help reduce the effects of an accident. In contrast, active safety systems help drivers avoid accidents by monitoring the state of the vehicle, the driver, or the surrounding traffic environment and providing driver alerts or control interventions. Examples of active safety technologies include traction control systems, electronic stability control systems, forward-collision warning and lane-departure warning systems, panic brake assist, lane-keeping aids, and automatic braking systems [1]. Systems that monitor driver states such as where the driver is looking or driver drowsiness also fall under the category of active safety systems.

Our research interest centers on the detection of drowsiness among fatigue-related impairments in driving. This paper makes two major contributions. First, the *camouflage nature of drowsiness* is revealed. Drivers that are deprived of sleep can still maintain performance in some routine driving tasks. However, their ability to cope with unusual or unexpected driving situations deteriorates. Second, a probabilistic framework based on Bayesian networks (BNs) for inferring drivers' state of drowsiness is introduced.

Online driver monitoring devices in motor vehicles have received renewed attention for helping detect fatigue in the U.S. and Europe since the late 1990s. These devices rely on a wide range of parameters, as there is no single commonly accepted metric to detect driver fatigue in an operational context [2]. Our study does not employ physiological signals such as EEG or physical changes of drivers such as eye-closure rate, but deals with data from driver–vehicle interaction. There are moments when a driver still looks awake (eyes wide open) but does not process any information [3]. Although explicit sleep-onset episodes can cause serious performance failures, some effects resulting from sleepiness can occur without microsleeps [4]. This implies that falling asleep may not be the only cause of fatigue-related accidents; performance deterioration due to drowsiness may not be induced only by sensor degradation such as eye closure but may be affected by controller degradation such as brain functions associated with sleep deprivation. To address this issue, our approach mainly uses the performance of drivers extracted from the driver–vehicle interaction.

The rest of this paper is organized as follows. After a brief introduction on the background and motivation (Sections I and II), Section III presents the experiment design, which includes a detailed explanation of different sleep-deprivation levels of drivers, simulated driving and nondriving tasks, performance metrics, and laboratory setup. Next, Section IV presents the analysis on experimental results. Section V discusses the application of BNs to detecting drowsy drivers based on the experimental data. Section VI summarizes the study and provides future research ideas.

J. Yang is with the Human Systems Integration Laboratory, Operations Research Department, Naval Postgraduate School, Monterey, CA 93943 USA (e-mail: jyan1@nps.edu).

Z.-H. Mao is with the Department of Electrical and Computer Engineering and the Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA 15261 USA (e-mail: maozh@engr.pitt.edu).

L. Tijerina and T. Pilutti are with the Ford Motor Company, Dearborn, MI 48126 USA (e-mail: ltijeri1@ford.com; tpilutti@ford.com).

J. F. Coughlin is with the Age Lab, Massachusetts Institute of Technology, Cambridge, MA 02139 USA, and also with the New England University Transportation Center, Center for Transportation and Logistics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: coughlin@mit.edu).

E. Feron is with the Daniel Guggenheim School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: feron@gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCA.2009.2018634

II. BACKGROUND

Specification of drowsy driving requires us to understand the sleep–wake mechanism, which, in turn, allows us to clarify the concept of drowsiness. It has been discovered that sleep is a dynamic behavior instead of a passive state caused by a decrease in stimulus level to humans [5]. People have an embedded sleep–wake cycle regulator controlled by a combination of two internal influences: circadian pacemakers and homeostasis [4], [6]. Environmental factors such as stress, noise, light, excitement, anger, pain, and sleep fragment are known to affect the sleep–wake cycle as well. However, in contrast with common belief, they do not *cause* us to sleep, but simply *unmask* any tendency to fall asleep that is already present. It is frequently misunderstood that boredom can cause sleepiness. It may unmask sleep in a human who is either originally sleep-deprived or in circadian sleep peaks, but itself does not cause sleepiness [6]. Masking is a critical concept for understanding sleep in terms of the sleep–wake cycle. It is common that people with chronic sleep deprivation can mask their level of sleepiness at their workplace. However, when they sit still and are deprived of external stimuli, sleep is unmasked and quickly arises [7].

The history of drowsy-driver research dates back to the 1950s, beginning with studies on aircraft pilots [8]. In the 1990s, driver fatigue began to be recognized as a major concern to both automotive industry and public-safety agencies. Recently, a 100-Car Naturalistic Study conducted by the National Highway Traffic Safety and Administration (NHTSA) and the Virginia Tech Transportation Institute has sought to provide precrash data that are necessary for understanding the cause of crashes, to support the development and the refinement of crash-avoidance countermeasures, and to estimate the potential utility of these countermeasures to reduce crashes and their consequences [9]. Europe also has several large-scale programs in progress, e.g., Dedicated Road Infrastructure for Vehicle Safety in Europe, Program for European Traffic with Highest Efficiency and Unprecedented Safety, Impaired Motorists, Methods of Roadside Testing and Assessment for Licensing, and Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation [10]–[13].

Fatigue has been estimated to be involved in 2%–23% of all crashes [14], [15]. The NHTSA conservatively estimates that 100 000 police-reported crashes are caused by drowsy drivers each year. (That is, about 1.5% of all crashes.) These crashes result in more than 1500 fatalities, 71 000 injuries, and an estimated $12.5 billion in diminished productivity and property loss [16]. The 1990 National Transportation Safety Board's (NTSB's) study of 182 heavy-truck accidents fatal to drivers showed that 31% of the accidents in this sample involved fatigue. The NTSB's numbers regarding fatigue-involved accidents are more revealing, as the NTSB's in-depth investigations included surrogate measures such as the 72-h history of rest and duty times, the amount of sleep in the last 24 h, and the regularity of the work schedule. An extensive summary on historical perspectives, facts, and the statistics on drowsy driving can be found in [17].

A. Existing Measures

1) Transportation Policies: The National Center on Sleep Disorders Research and NHTSA expert panels on driver fatigue [18] recommend three priorities for an educational campaign: 1) educate young males (ages 16–24) about drowsy driving and how to reduce lifestyle-related risks; 2) promote shoulder rumble strips as an effective countermeasure for drowsy driving and, in the same context, raise public awareness about drowsy-driving risks and how to reduce them; and 3) educate shift workers about the risks of drowsy-driving and how to reduce them.

2) Law Enforcement: The first federal bill focusing on drowsy driving was introduced in the House of Representatives in October 2002 by Republican Robert Andrews [19]. The bill, i.e., HR 5543, is called *Maggie's Law: National Drowsy Driving Act of 2002*. The law narrowly defines *fatigue* as being without sleep for a period in excess of 24 consecutive hours. Under Maggie's law, anyone causing a fatality after being awake for 24 h or more can be prosecuted for vehicular homicide. Currently, a number of states, including New York, Massachusetts, Tennessee, Oregon, Kentucky, and Illinois, are considering similar drowsy-driving legislation [19].

3) Fatigue Detection Techniques: Along with transportation policies, reliable and applicable drowsy-driving detection techniques may help detect fatigue. Researchers have developed a variety of drowsiness-detection methods, which can be classified in terms of their specific techniques [2], [20], [21]. References [17] and [20] have summarized the detection techniques based on: 1) physiological signals, including pulse rate and EEG; 2) physical changes, including changes of head position, eye-closure rate, and eyelid movement; 3) driver–vehicle data, including steering angle, throttle/brake input, and speed; and 4) secondary tasks that periodically request responses from drivers.

III. DESIGN OF EXPERIMENTS

A. Experimental Objectives

The simulator-based human-in-the-loop experiments are intended to explore the characteristics of drowsy driving. The purpose of the experiments is not to propose another drowsiness-monitoring technique but to understand the in-depth characteristics of driver–vehicle systems under different sleep-deprivation levels of drivers. For instance, we will not only compare driver performances between drowsy and alert drivers but also investigate tasks that show little performance deterioration under sleep deprivation.

Our main working hypothesis is that the deterioration degree of a drowsy-driver performance is greater in "medium-level" switching than in "low-level" regulation tasks. The rationale of this hypothesis came from the physiological basis of the sleep–wake cycle and the motor control of the human nervous system [17].

B. Independent Variable

The independent variable of the experiment is the *sleep-deprivation level* of the human subjects, i.e., the level of homeostatic need for sleep. We consider two sleep-deprivation levels,

i.e., "partial sleep-deprivation" and "no sleep-deprivation." The level of sleep deprivation is determined by the amount of sleep that each subject had before the test day. The non sleep-deprived subjects slept for at least 7–8 h per 24 h for more than a week before the test day. The partially sleep-deprived subjects had less than 7 h in bed two days before the test and less than 4 h in bed on the eve of the test.

Actiwatches, which are small actigraphy-based data loggers that record digitally integrated measure of gross motor activities, allow us to objectively measure the amount and the quality of sleep for several days prior to the experiments. Subjects were asked to wear the MiniMitter AW-16, which can be worn just like a watch, for one week prior to the non sleep-deprived session and for two to three days prior to the partially sleep-deprived session.

C. Simulated Tasks

A series of simulated driving and nondriving tasks were given to the subjects. A full description of the driving scenarios, including road geometry, intervals between tasks, and random order of tasks, can be found in [17].

1) Tracking Tasks: Deterioration in lane-tracking performance can lead to overall driving malfunction. Lane-tracking performance has been considered a main indicator for detecting drowsy drivers [21]–[25]. However, the validity of this indicator is still controversial, as shown in [3] and [22]. Moreover, between alert and drowsy drivers, there have been few studies on their performance in lane tracking under various road or weather conditions. This type of study is nontrivial, as real driving happens under a variety of situations. Thus, for non sleep-deprived and partially sleep-deprived subjects, we examined lateral lane-tracking performance under several conditions to evaluate the drivers' ability to maneuver a vehicle inside the roadway.

Five different tracking tasks were given to each subject in a random order while driving. The drivers drove in the simulator as if they were driving in the real world and were not supposed to change lanes or pass a lead vehicle in front of them. The five tracking tasks involved driving on the following: 1) a curved road; 2) a straight road with changes in steering dynamics; 3) a straight road with a lead vehicle; 4) a straight road without any disturbance; and 5) a straight road with disturbances (e.g., wind gusts), respectively.

LT: Without any disturbances or stimuli, the drivers kept their vehicle centered on a straight road 700 m long under a specified speed range between 80 and 100 km/h. Each test included three straight-lane tracking (LT) events.

LV: When a lead vehicle appeared, the subjects were instructed not to pass that vehicle but to follow the vehicle with a safe margin that they chose. Each test included three straight-lane tracking given a lead vehicle (LV) events, and each lead vehicle was presented for 100 s.

WG: The drivers kept their vehicle centered in a straight road under a specified speed range between 80 and 100 km/h, whereas some external pseudorandom disturbances (specified by multiple sine waves) were introduced. Each test included three straight-lane tracking given wind gusts (WG) events, each lasting for 24 s.

SC: The drivers kept the vehicle centered in the straight road and maintained a specified speed range between 80 and 100 km/h, while the original steering dynamics were *altered*. Each test included three straight-lane tracking given changes of steering characteristics (SC) events, each having a length of 1900 m.

Although we can easily apply external or environmental disturbances such as wind gust, bumpy roads, and fog in simulated driving, we cannot control the presence of the environmental disturbance in real driving. However, this internal disturbance (SC) can easily be generated via drive-by-wire technology in real driving. The SC was introduced to observe how the drivers adapt themselves to disturbances and then to study how to apply this event to real-driving situations. In the SC tasks, we applied some nonlinearity (backlash) in the steering-alignment system, originally modeled and implemented as a linear system.

Backlash is a common nonlinearity that limits the performance of speed and position control, and typically happens in worn-out cars. It causes a phase delay and, thus, a loss of information by clipping the peaks of input signals [26]. The discrete-time version of the backlash model is

$$u(t) = B\left(\nu(t)\right) = \begin{cases} m\left(\nu(t) - c_l\right), & \text{if } \nu(t) \leq \nu_l \\ m\left(\nu(t) - c_r\right), & \text{if } \nu(t) \geq \nu_r \\ u(t-1), & \text{if } \nu_l < \nu(t) < \nu_r \end{cases} \quad (1)$$

where $\nu_l = u(t-1)/m + c_l$ and $\nu_r = u(t-1)/m + c_r$. Coefficients used in the simulation were $m = 1$, $c_l = -10°$, and $c_r = 10°$.

CL: The drivers kept the car centered in a serpentine road under a specified speed range between 80 and 100 km/h. The length of each serpentine was approximately 1500 m. Each test included three curved-lane-tracking (CL) events. Subjects practiced on a curved road before they start the main experiment. The serpentine geometry used in the practice was different from that used in the main experiment.

2) Stimulus-Response Tasks: Four stimulus-response tasks were given to each subject in a random order during the simulated driving. Stimulus can be an auditory ringing signal, a visual red triangular symbol, or an overhead lane-change sign on the driving lane. Definitions of the four stimulus-response tasks are given as follows.

SLCT: Once an overhead lane-change sign appeared, the drivers were supposed to immediately change lanes. Each lane-change sign had an array consisting of an arrow and two X's. The arrow indicates the target lane that the drivers need to move into. The rest of the lanes were marked with X's to indicate that the drivers should not move in those lanes. The single-lane-change task (SLCT) sign refers to the moment when the lane indicated by the arrow is adjacent to the current driving lane. Each experimental session included five SLCT events.

DLCT: The lane-change sign had the same format as described in SLCT. However, the double-lane-change task (DLCT) sign refers to the moment when the lane indicated by the arrow is separated from another lane, and the drivers need to shift two lanes at once. Fig. 1 is a screenshot of the driving simulator, showing one of the lane-change signs. Each experimental session included five DLCT events.

Fig. 1.   Screenshot of the DLCT.

APVT: The drivers were supposed to press a green button on the steering wheel immediately after hearing a ringing tone. The ringing tone lasted for about 1 s. Each experimental session included ten auditory psychomotor vigilance task (APVT) events.

VPVT: The drivers were supposed to press a green button on the steering wheel immediately after recognizing a red stimulus on the screen. The stimulus was shown for 5 s if there was no response from the subjects. Each experimental session included ten visual psychomotor vigilance task (VPVT) events.

D. Dependent Variables

The dependent variables of the experiments are the performance measures of the simulated tasks. In this paper, we consider the following performance measures.

1) RMT: The root-mean-square (RMS) error with threshold (RMT) is a parameterized variation of the conventional RMS error, which is usually used to measure the general tracking performance. We have devised the RMT instead of the RMS to capture common driving characteristics. The drivers tend to ignore a certain level of errors, as they generally try to stay *within* the driving lane instead of trying to follow a single line on the road. This driving characteristic is usually called "good-enough" or "satisfying" characteristics of drivers [20]. Thus, we introduce a threshold γ to the RMS so that the RMT vanishes as long as driving trajectories stay within the threshold γ.

The RMT is defined as

$$\text{RMT} = \sqrt{\frac{\sum_{k=t_0}^{t_f} \max\{|x(k)| - \gamma, 0\}^2}{n}} \qquad (2)$$

where $x(k)$ is the lateral lane position of a driver with respect to the centerline of the driving lane at time k, γ is a threshold value varying from 0% to 50% of the road width, n is the number of data within the sampling window, t_0 is the initial time in the sampling window, and t_f is the terminal time in the sampling window. It is apparent that (2) is reduced to a typical RMS error when $\gamma = 0$.

2) RT: The reaction time (RT) is a measure of how fast a driver reacts to stimuli presented abruptly, i.e.,

$$\text{RT} = t_{\text{action}} - t_{\text{stimulus}} \qquad (3)$$

where t_{stimulus} is the time a stimulus is presented to the driver, and t_{action} is the time the man–vehicle system reacts to the given stimulus. (For SLCT and DLCT, t_{stimulus} is the time when the lane-change sign appears on the screen, and t_{action} is the time when the driver starts to steer toward the lane indicated.)

3) ETL: The effective time delay (ETL) of continuous tasks can be considered equivalent to the RT of discrete tasks [27]. The ETL is estimated by applying McRuer's crossover model to some continuous tasks. McRuer's crossover model [28] claims that, for continuous tracking tasks, humans adapt themselves to the system in such a way as to make the total open-loop transfer function behave as a first-order system with *gain* and *effective time delay*. Thus, the total open-loop transfer function can be expressed as

$$Y_H Y_P = \frac{Ke^{-\tau s}}{s} \qquad (4)$$

where Y_H models the human operation, Y_P is a plant, K is a gain, and τ is the ETL.

4) CRR: The accuracy of the drivers' response is measured by the correct response rate (CRR), i.e.,

$$\text{CRR} = \frac{\sum_{i=1}^{n} 1(R_i)}{n} \qquad (5)$$

where n is the total number of responses under consideration, and $1(R_i)$ is equal to 1 when R_i (denoting the ith response) is correct, i.e., the drivers do what they are supposed or instructed to do so. On the other hand, $1(R_i)$ is equal to 0 when the ith response is incorrect. The CRR measures the accuracy of the drivers' response, whereas the RT measures the speed.

E. Laboratory Setup and Experiment Procedure

We developed a simulation test bed allowing human subjects to drive with predesigned driving scenarios. Simulated driving was programmed and run by STISIM Drive v2.06.07. A projector and a wide screen (45 × 60 in) were implemented in the lab, showing a driving scene to the driver. A six-way power adjustable driver's seat was used for the experiment.

Each test session lasted for an hour, including a 10-min demo drive during which a driver learned the tasks described in Section III-C except SC and WG. SC and WG were not included in the demo, as they were not rule-based tasks but emergency-type ones, for which drivers should react with their instinct. An experimenter explained how the subjects should drive or react for relevant tasks. The drivers were asked to practice each task until an adequate performance level was achieved. The instructional and demonstrational session was followed by the main experiment scenario. The initial 20 000 m (approximately 10 min) of the main experiment was the unmasking period [17]. Then, randomly placed and ordered tasks, as described in Section III-C, were successively given to the driver. No two tasks were simultaneously given to a driver. The driver was not informed of the time and duration of the
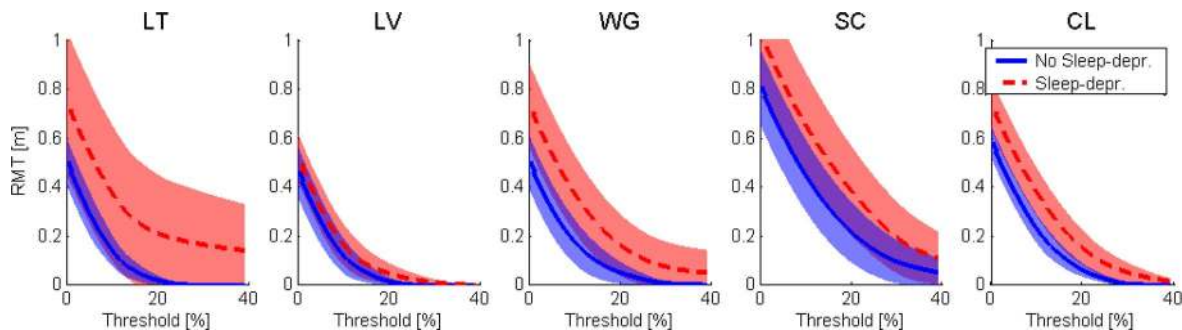
Fig. 2.   RMT sampled over 350 m.

tasks, including the masking period before or during the main simulation. This main simulation ran for up to 40 min.

Non sleep-deprived subjects participated in the experiment between 9:00 A.M. and 12:00 P.M., whereas partially sleep-deprived subjects participated between 2:00 P.M. and 4:00 P.M. considering their circadian rhythm. An experimenter was present in the laboratory throughout the whole test session. Conversation between the subject and the experimenter was prohibited during the main scenario unless an emergency occurred. Sleep-deprived subjects were provided with transportation on the test day to avoid safety-related accidents.

IV. EXPERIMENTAL RESULTS

A total of 12 male subjects participated in the experiment twice: under conditions of partial sleep-deprivation and non sleep-deprivation, respectively. The ages of these subjects ranged from 29 to 49 years, with an average of 41.7 years and a standard deviation σ of 6.4 years. The average amount of sleep these subjects had (measured by MiniMitter AW-16) per day for two weeks prior to the sleep monitoring was 6 h 42 min ($\sigma = 47$ min). None of the subjects had any serious health problems or were under any medication.

For the non sleep-deprived condition, the subjects had an average of 7 h 22 min ($\sigma = 46$ min) of sleep per day for seven days prior to the experiment. For the partially sleep-deprived condition, the subjects had 6 h 11 min of sleep on the night two days before the experiment ($\sigma = 35$ min) and had 3 h 27 min of sleep on the eve of the experiment ($\sigma = 26$ min) by average. The order of the experiment was counterbalanced, and no subject yielded a crash during the simulation.

A. Within-Subject Performance Analysis

We are mainly interested in the following: which driver–vehicle data show more or earlier differences between two groups? Which data show the least amount of differences? Are there any commonalities in the tasks showing performance degradation when drivers are sleep-deprived? Each subject participated in the experiment twice with different sleep-deprivation levels, and this section compares *within*-subject performance differences. We used a nonparametric statistical test called the Wilcoxon signed-rank test [29] to determine whether there exist significant performance differences between the two groups. The p-values listed in the following sections were obtained from the Wilcoxon signed-rank test,

unless otherwise mentioned. The significance level α was set at 0.1, as was conventionally used in general human-factor studies.

1) Lane-Tracking Performance: Fig. 2 shows the average performance (lines) and one standard deviation (shaded area) of the 12 subjects sampled over a distance of 350 m (approximately 13–15 s) of the LT, LV, WG, SC, and CL tasks. The x-axis represents the threshold γ from the centerline of the road, which is given in percentage unit with 100% as the road width. (Fifty percent threshold covers the whole road width.) The y-axis represents the RMT. The solid lines indicate data from the non sleep-deprived group, and the dotted line indicates data from the partially sleep-deprived group.

Fig. 2 shows that the dotted lines are above the solid lines for each lane-tracking task. This implies that the performance of the partially sleep-deprived group was worse than that of the non sleep-deprived group for all tracking tasks. (Analysis will follow to show if there exist statistically significant differences between the two groups.) Clearly, the RMT is a function of the threshold; as the threshold increases, the value of the RMT decreases. This is reasonable since the wider the threshold is, the easier the driver stays within the threshold, and, thus, the smaller the lane-tracking error becomes. Moreover, the RMT versus the threshold is "L"-shaped: The RMT rapidly varies around 0%–10% of the threshold, and the rate of change decreases as the threshold increases. This implies that the lane-tracking performance is more sensitive to smaller threshold values. Such trend is compatible with the drivers' "good-enough" strategy for lane tracking.

Fig. 2 shows possibilities of performance degradation in drivers that are deprived of sleep. Nonparametric statistical tests further confirmed the significance of the results in Fig. 2. Fig. 3 shows the p-values calculated from the data in Fig. 2. The x-axis represents the same threshold as in the previous figure, and the y-axis represents the p-values. The p-values of LT, LV, WG, SC, and CL were plotted in lines of different styles and colors. We set $\alpha = 0.1$, and the p-values less than α were considered to be significant.

Whereas the p-values of WG, SC, and CL are smaller than α for a good range of thresholds, those of LT and LV are *not*. This result implies that sleep-deprived drivers performed worse than non sleep-deprived drivers in WG, SC, and CL, although their performances were not significantly different in LT and LV.
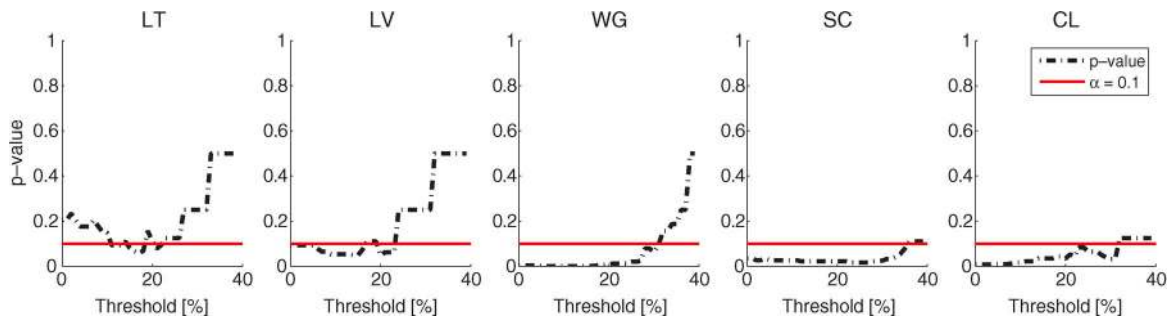
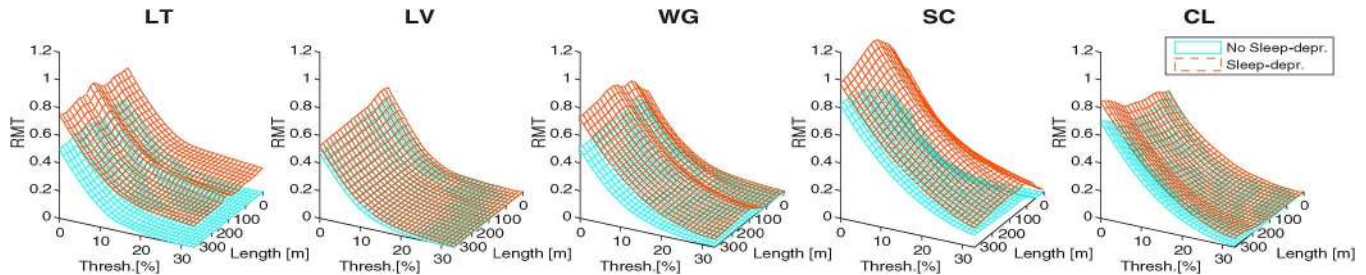Fig. 3. *p*-values calculated from the RMT values in Fig. 2.
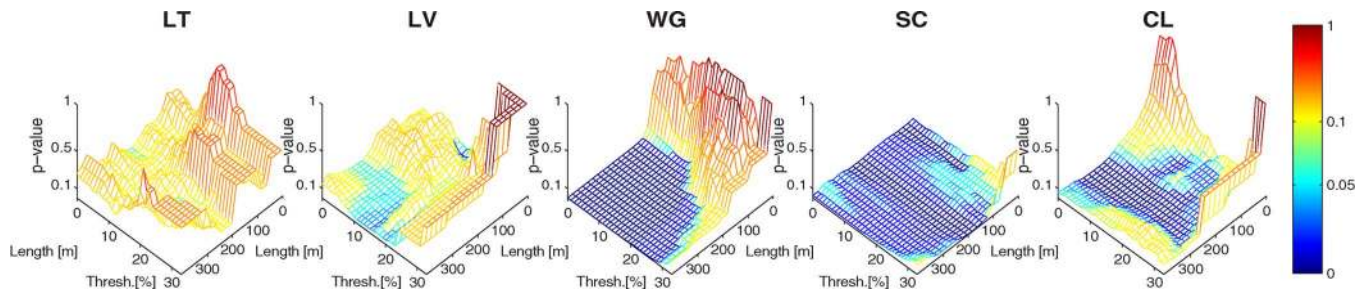

Fig. 4. Threshold versus sampling length versus RMT.

Fig. 5. *p*-values calculated from the RMT values in Fig. 4.

We also want to know how quickly the performance differences can be detected between the two groups (non sleep-deprived and partially sleep-deprived) and how the differences propagate as time evolves. We introduced a parameter called sampling length, which indicates the length of data that need to be collected to detect the differences between the two groups. The results are shown in Fig. 4. The x-axis represents the threshold same as in Fig. 2, the y-axis is the sampling length (ranging from 5 to 400 m), and the z-axis represents the RMT values. Surfaces with solid (or dotted) lines indicate the data from the non sleep-deprived group (or the partially sleep-deprived group). Fig. 4 shows similar trends of the RMT as in Fig. 2, but plots RMT surfaces instead of lines. Clearly, the RMT depended on the threshold: The RMT value decreased as the threshold increased, and the RMT surfaces were "L"-shaped with some folds between edges. Since it is difficult to see the effect of the sampling length in Fig. 4, we turned to statistical tests, which may provide more insights.

Fig. 5 presents the p-values calculated from the RMT values in Fig. 4. (Fig. 3 is just one of the vertical cross sections of Fig. 5.) The p-values were plotted in color, and the color map (the rightmost bar) had high resolution for the p-value ranging from 0 to $0.5\alpha = 0.05$, medium resolution

for $p = 0.05 - 0.1$, and low resolution from $p = 0.1 - 1$. In doing so, we emphasized on the p-values that are smaller than α.

In the subplots for LT and LV, most of the p-values were greater than α. The plots had a noise-like shape, which implies that it is difficult to differentiate the performance between the non sleep-deprived and partially sleep-deprived groups.

In the subplots for WG, SC, and CL, most of the p-values were smaller than α. This result suggests that the performance differences between non sleep-deprived and partially sleep-deprived groups were significant in WG, SC, and CL. Furthermore, WG, SC, and CL were robust to the threshold and sampling length since the statistical significance was observed for a wide range of the sampling length and threshold. For WG, the p-value decreased as the sampling length increased. This may indicate that the subjects, when not deprived of sleep, adapted themselves to disturbances faster than when they were partially deprived of sleep. For SC, the p-value was approximately 0 for most values of the sampling length and threshold. This suggests that the SC task demonstrated distinct and significant performance differences between non sleep-deprived and partially sleep-deprived groups. For CL, the differences were present, but not as significant as in WG or SC.

TABLE I
MEAN, σ, AND SPEARMAN'S COEFFICIENTS OF RT

Mean value [sec]

	APVT	VPVT	SLCT	DLCT
No sleep-depr.	0.741	0.907	0.763	0.722
Sleep-depr.	0.927	1.078	0.814	0.857

σ [sec]

	APVT	VPVT	SLCT	DLCT
No sleep-depr.	0.248	0.251	0.099	0.099
Sleep-depr.	0.359	0.351	0.170	0.188

Spearman's correlation coefficients (ρ) between tasks.

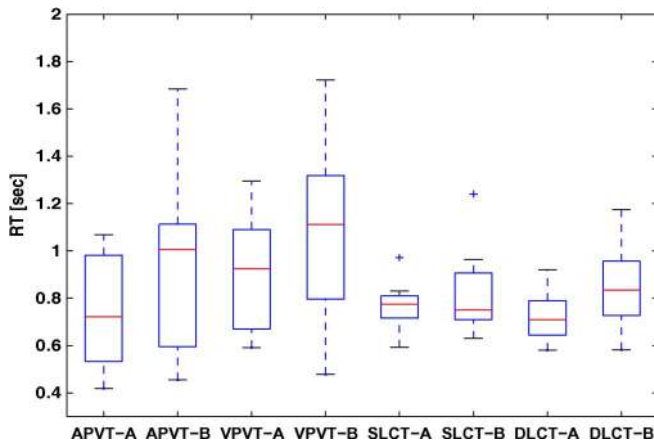|	APVT	VPVT	SLCT	DLCT
APVT	.		.	.
VPVT	**0.9313**	.	.	.
SLCT	0.3200	0.1800	.	.
DLCT	**0.6270**	**0.6400**	**0.4574**	.

Fig. 6. Boxplot of RT data from stimulus-response tasks.

We can see from the above results that, when encountering abrupt changes caused by environment (WG), vehicle (SC), or challenging road geometry (CL), the drivers showed significantly different adaptability to changes under different levels of sleep deprivation. However, if there was no excitation or disturbances introduced (LT or LV), the drivers' performance showed little difference between different sleep-deprivation levels. This implies that drowsy drivers may drive as well as alert drivers for common driving tasks without any disturbances. More in-depth explanation in conjunction with other performances is given in Section IV-B.

2) Stimulus-Response Task Performance: Table I shows the mean and the standard deviation σ of the RT for each task under different sleep-deprivation levels. The results were also visualized in a boxplot (Fig. 6). Each box contained horizontal line segments at the lower quartile, the median, and the upper quartile of the RT data. The whisker extended to the most extreme value within 150% of the interquartile range. Outliers with values beyond the ends of the whiskers were marked with +. Tasks under non-sleep-deprivation and partial sleep-deprivation were labeled A and B, respectively.

Table I and Fig. 6 show that the RT of the partially sleep-deprived group was slower than that of the non sleep-deprived group in general. This observation was confirmed by statistical tests as RTs of APVT, VPVT, and DLCT all showed significant differences between the two groups, with p-values of 0.0034, 0.0269, and 0.0049, respectively. (Previous papers [30], [31]

reported consistent results in psychomotor vigilance tasks.) However, SLCT did *not* show a significant difference between the two sleep-deprivation levels.

We are also interested in the correlations between tasks. These correlations can be used to find alternatives to some simulated tasks that might not be available in real driving. A matrix containing the pairwise Spearman's rank correlation coefficient ρ is given in Table I. Tasks with high ρ, i.e., tasks associated with p-values less than α, were marked with bold font. APVT, VPVT, and DLCT showed high correlations with each other, whereas SLCT was only marginally correlated with DLCT and showed low correlations with APVT and VPVT. Thus, the tasks revealing significant performance differences between the sleep-deprived and non sleep-deprived groups (APVT, VPVT, and DLCT) were highly correlated with each other as well.

On the other hand, the RT data can directly be compared among these tasks. The RT of VPVT was significantly different from those of APVT, SLCT, and DLCT, with p-values of 0.0001, 0.0043, and 0.0015, respectively. A reaction to VPVT was slower than those to AVPT, SLCT, or DLCT. The difference between the RT data in APVT and VPVT can be explained by the stimulus modality. It is known that drivers react faster to auditory stimuli than to visual cues [32]. This is consistent with our data. The difference between the RT data in VPVT and SLCT/DLCT can be explained by the difference between the drivers' actions in the tasks. In the response to VPVT, the drivers need to make an additional hand/arm movement by pressing a button on the steering wheel. In contrast, the drivers only need to rotate their arms to turn the steering wheel to perform lane changes in SLCT/DLCT.

An interesting result is that the performance of SLCT did not degrade even when the subjects were partially deprived of sleep. However, the performance of DLCT did degrade. Note that the stimulus and task protocols of DLCT were identical to those of SLCT, except that the amplitude of lane changes of DLCT was twice that of SLCT. Since SLCT is much more common than DLCT in real driving, our result suggests that drowsy drivers are able to maintain their normal level of performance in most lane-change tasks.

3) Application of the Crossover Model to Lane-Change Tasks: We considered SLCT and DLCT stimulus-response tasks, and the performance metric was the RT. Meanwhile, lane-change tasks can be considered tracking tasks—tracking unit step input. We now employ the McRuer's crossover model to analyze SLCT and DLCT as continuous tracking tasks.

In the crossover model [28] for SLCT and DLCT, the human operator is the driver, the plant is the vehicle, the reference signal is the demanded change of lane, the output signal is the lateral position of the vehicle, and the perceived error is the difference between the reference and output signals. The open-loop transfer function can be simplified as (4), and we can estimate its gain K and ETL τ for each subject for SLCT and DLCT. With these estimated parameters, we can run statistical tests to determine whether any significant differences exist between the non sleep-deprived and partially sleep-deprived groups. MATLAB 6.5.1 System Identification Toolbox was used to estimate the gain and the ETL, and the result showed no significant differences in gain K for SLCT and DLCT.
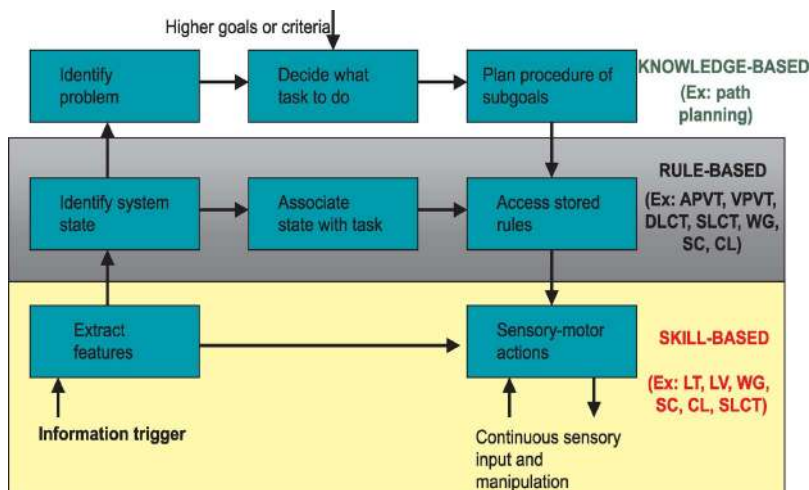
Fig. 7. Rasmussen skill-, rule-, and knowledge-based classifications with simulated task examples [13].

However, in DLCT, we observed an apparent difference in ETL τ (the p-value was 0.0408). This is consistent with the result using the RT as the performance metric in DLCT. Here, the ETL of continuous tasks can be viewed as the RT of stimulus-response tasks.

4) CRR: In addition to the RT, which measures how fast the drivers respond to given stimuli, we are also interested in the validity of the drivers' responses. For APVT and VPVT, the drivers were supposed to press a green button on the steering wheel immediately after a given stimulus. For SLCT and DLCT, the drivers were supposed to change lanes immediately after seeing an overhead lane-change sign. The data showed that all stimulus-response tasks examined during the experiment had no significant differences in the CRR between the non sleep-deprived and partially sleep-deprived groups. No learning or order effect was found for the RT and the CRR.

5) Other Parameters: The lateral position of the vehicle can approximately be modeled as the double integration of the steering movement of drivers [32]. Thus, we might expect that the performance analysis of steering may provide a result that is similar but more sensitive to the sleep-deprivation levels of the drivers than the lane-tracking performance. However, our data showed that the performance differences in steering control between the non sleep-deprived and partially sleep-deprived groups were not as apparent as in lane tracking. This suggests that the drivers' control strategies might not be affected by sleep deprivation.

We were not able to find significant differences in either longitudinal velocities or throttle-and-brake control under the two sleep-deprivation conditions in LT, LV, WG, SC, and CL. The metric used for the steering behavior and the longitudinal control was the RMT as in the tracking tasks.

B. Camouflage Nature of Drowsy Driving

We can analyze the experimental results within the framework of information processing introduced by Rasmussen [33]. The Rasmussen classification claims that human behavior is based on a *skill*, *rule*, and *knowledge* hierarchy. Skill-based tasks need the least cognitive resources; little or no conscious control is used to perform an action. These tasks include highly automated tasks such as walking. Rule-based tasks need more information-processing stages to properly perform relevant tasks. They require us to identify the system state to execute the appropriate rules. Knowledge-based tasks are the highest level tasks involving advanced problem-solving or decision-making.

Fig. 7 depicts the Rasmussen hierarchy. Obvious examples of skill-based tasks from our simulated driving include LT and LV. Both are common driving tasks for which the drivers need little "conscious" control. However, other driving tasks such as WG, SC, and CL hold the characteristics of both skill- and rule-based tasks. The drivers need to pay more attention to control their vehicles in these tasks since unexpected external disturbances may occur. The stimulus-response tasks have the characteristics of the rule-based tasks. In APVT, VPVT, SLCT, or DLCT, a set of rules are assigned, and drivers are supposed to follow the rules. However, SLCT also possesses skill-based characteristics because the drivers perform SLCT very often and are familiar with this maneuver. We have not included any knowledge-based task in our experiments, but an example of it could be path planning.

Recall that the sleep-deprived drivers performed worse only in tasks 1) WG, SC, and CL (tracking tasks with disturbances), and 2) APVT, VPVT, and DLCT (stimulus-response tasks). Performance differences cannot be differentiated in LT, LV, and SLCT between the two levels of sleep deprivation. When connecting these observations with the classification introduced in Fig. 7, we claim that the performance of the sleep-deprived drivers *mainly degrades in rule-based tasks* rather than in skill-based tasks. This result implies that drowsiness has greater effect on the tasks related to the rule-based (medium-level) cognitive functions than skill-based (low-level) cognitive functions. This interpretation uncovers important characteristics of drowsy driving since most driving tasks are skill-based tasks. Drowsy drivers are robust enough to perform the routine tasks such as lane tracking or single-lane changing, and drowsy driving is unobservable in those skill-based tasks. This
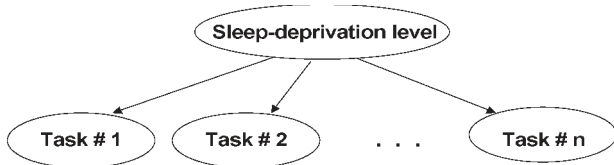
Fig. 8.   Bipartite static Bayesian network for drowsiness detection.

camouflage nature of drowsy driving suggests that we should avoid using skill-based tasks in the detection of drowsy driving.

V. DETECTION OF SLEEP-DEPRIVED DRIVERS

In Section IV, we focused on finding the performance differences between drivers with different levels of sleep deprivation. In this section, our focus is on inferring the drivers' states based on their performances. This will be made possible by using a probabilistic graphical model, i.e., the BN [34]. In the following, we provide quantitative guidelines for the application of a static BN (SBN) and a dynamic BN (DBN).

We utilize the BN paradigm because it is capable of incorporating prior information, explicitly modeling uncertainties and temporal aspects of the problem, and modeling data at different levels of abstraction. The BN has recently been introduced in drowsiness detection based on monitoring the physical behavior of the drivers [34], and it has been shown that the BN is able to capture dynamics associated with fatigue.

A. Formulation of the SBN for Drowsiness Detection

We first utilized the SBN to infer the driver's drowsiness based on our experimental data. A detailed introduction of the SBN can be found in [35].

Fig. 8 shows a bipartite structure with only one parent node. This structure directly reflects our experimental setup described in Section III, where the only independent variable was the sleep-deprivation level of drivers, and the dependent variables were task performances. The sleep-deprivation level was modeled as a parent node in Fig. 8. Each task performance such as the performance of APVT, CL, DLCT, EM, LT, LV, SC, SLCT, VPVT, or WG was a child node of the sleep-deprivation level. They are shown as Task #1, #2, ..., #n in Fig. 8. We did not include any direct links among the task performances because there were no causal dependencies among them.

The parent node has two states, defined as T = partial sleep-deprivation and F = non sleep-deprivation, respectively. A child node, e.g., the node corresponding to task j (one of the driving tasks), also has two states, denoted T_{task_j} and F_{task_j}, respectively. For example, when task j is LT, the two states are denoted T_{LT} and F_{LT}. A child node corresponding to a tracking task is in state T_{task_j} if $\text{RMT}_{\text{task}_j} \geq \theta_{\text{task}_j}$ (θ_j is a preset threshold for task j), and in state F_{task_j} otherwise. Table II presents a conditional probability table (CPT) calculated from the experimental data for a sampling length of 350 m and an RMT threshold of 20%. The CPT for stimulus-response tasks can be found in [17].

TABLE II
CPT

Parent node	LT		LV	
	T_{LT}	F_{LT}	T_{LV}	F_{LV}
T	0.500	0.500	0.286	0.714
F	0.357	0.643	0.286	0.714
WG		SC		CL
T_{WG}	F_{WG}	T_{SC}	F_{SC}	T_{CL}	F_{CL}
0.571	0.429	0.429	0.571	0.500	0.500
0.286	0.714	0.286	0.714	0.357	0.643

TABLE III
ESTIMATED PROBABILITY OF SLEEP-DEPRIVATION
BASED ON EACH TASK PERFORMANCE

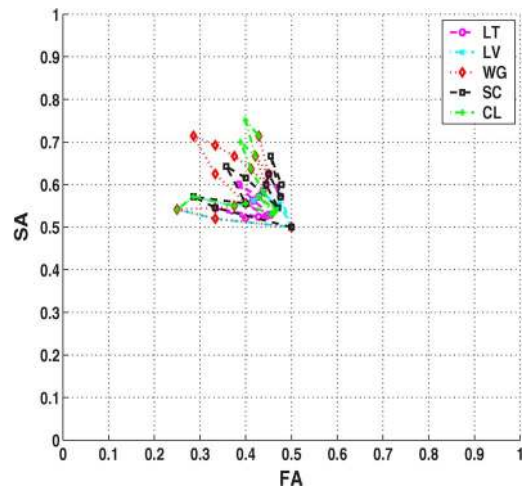| Successful Alarm		False Alarm	
$p(T\|T_{LT})$	0.5833	$p(T\|F_{LT})$	0.4375
$p(T\|T_{LV})$	0.5000	$p(T\|F_{LV})$	0.5000
$p(T\|T_{WG})$	0.6667	$p(T\|F_{WG})$	0.3750
$p(T\|T_{SC})$	0.6000	$p(T\|F_{SC})$	0.4444
$p(T\|T_{CL})$	0.5833	$p(T\|F_{CL})$	0.4375

Fig. 9.   SOC curve of tracking tasks.

B. SBN Simulation Results

Successful alarm (SA) and *false alarm* (FA). We first examined the positive and negative outcomes of the inference based on the driving performance. SA is a positive consequence and FA is a negative consequence of the inference [17]. Table III shows both the SA and the FA when the performance of only one driving task was considered. We can also calculate the inference when we have the performance of multiple tasks. A system operating characteristics (SOC) curve [36] is shown in Fig. 9, where the SA and the FA from the different lapse thresholds based on a single task are presented. The thresholds varied from the mean value of non sleep-deprived drivers minus 100% of σ to the mean plus 150% of σ. We also formulated the SBN for each individual subject. Figs. 10 and 11 show two examples of the SOC curves from subjects 9 and 10, respectively.

C. Formulation of DBN

We have investigated the SBN mainly considering *causality* between variables, but the SBN did not consider the
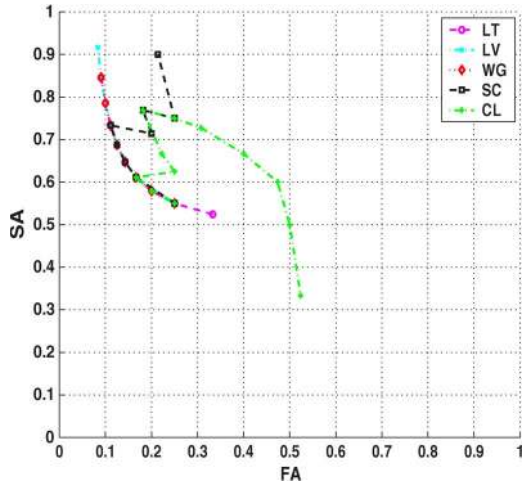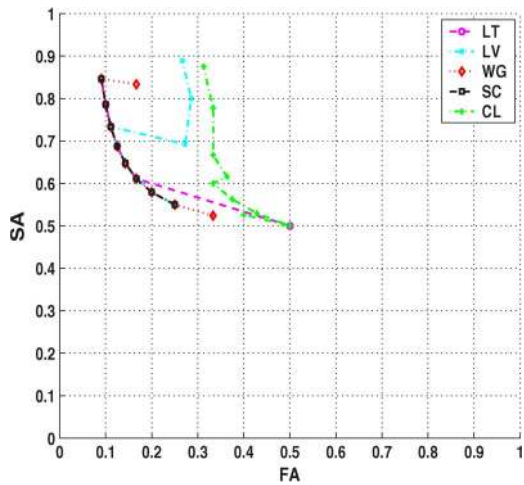
Fig. 10.   SOC curve: subject 9.
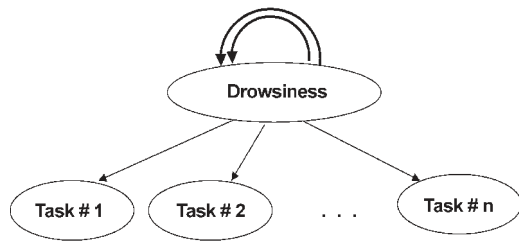

Fig. 11.   SOC curve: subject 10.

Fig. 12.   Temporal model of drowsiness detection. (Double arrow) Temporal link from time slice t to slice $t + 1$.

temporal aspects of the drowsy-driving detection. Drowsiness may be caused by either sleep deprivation or circadian rhythm; many other factors such as boredom or motivation can also mask/unmask sleepiness (Section II). Therefore, drowsiness is time-dependent, and the temporal aspect of it should be considered in the detection. We assume that a DBN for drowsy-driving detection is repetitive in the sense that the structures of time slices, temporal links, and conditional probabilities are time invariant. Formulation of the DBN is explained in [17], and Fig. 12 shows a sketch of the DBN.
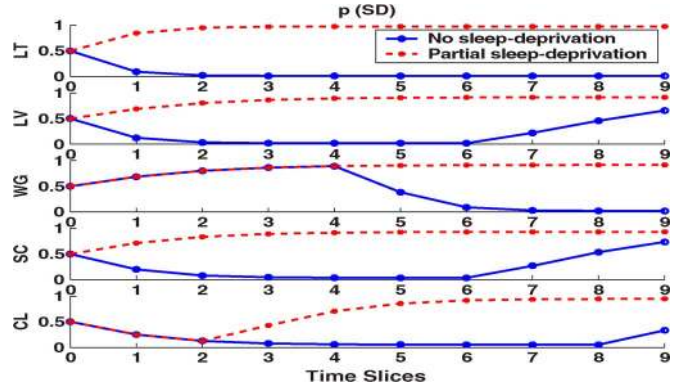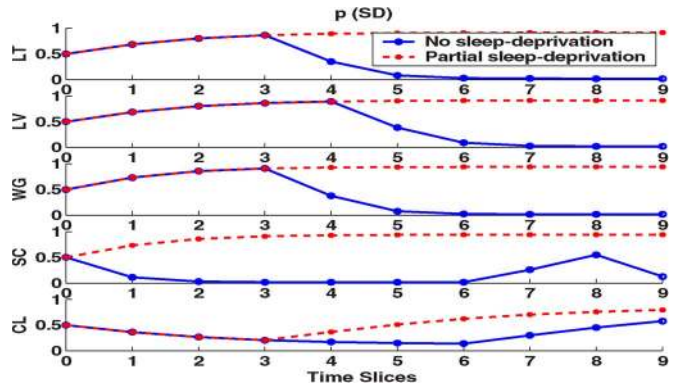

Fig. 13.   DBN curve: subject 9.

Fig. 14.   DBN curve: subject 10.

D. DBN Simulation Results

Figs. 13 and 14 show two examples of the probabilities of drowsiness estimated from the DBN for subjects 9 and 10, respectively. We created time slices for every 40 m starting from the first 150 m in each tracking task. In each figure, the data corresponding to the non sleep-deprived and partially sleep-deprived conditions were plotted in solid and dashed lines, respectively. For subject 9, we can distinguish the driver's states of drowsiness based on his performances in LT, LV, and SC. For subject 10, only SC can provide us with sufficient information to distinguish the driver's drowsiness. For each one of the other subjects, at least one driving task can significantly reveal the driver's drowsiness in driving.

VI. CONCLUSION AND FUTURE WORK

This paper has revealed the characteristics of drowsy driving through simulator-based human-in-the-loop experiments. We have observed that drowsiness has greater effect on rule-based driving tasks than on skill-based tasks. We have confirmed this finding by inferring driver alertness using the BN paradigm. Based on this paper, we suggest that the driving performance of the rule-based tasks should be investigated further for the effective design of drowsy-driver detection systems. The rule-based tasks examined in our experiments were RT tasks and tracking tasks with unexpected disturbances. Other rule-based tasks such as stopping at traffic signals should be examined. Skill-based tasks, which cover most driving tasks, should also be considered in the detection system. Although skill-based
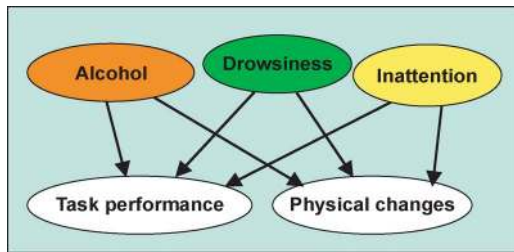
Fig. 15. Detection of impaired drivers.

tasks cannot be used to provide early indicators of drowsy driving, deterioration of such tasks may indicate the existence of other driving impairments such as inebriation.

We can generalize the methods developed in this paper to detect other driving impairments, once we understand the characteristics of impaired driving under the influence of alcohol, motion sickness, stress, or inattention. For example, we can include the associated nodes for each impaired driving condition in the BN structure to assess the driver's state. Fig. 15 presents an extended version of Fig. 8, including alcohol and inattention nodes as other possible causes of impaired driving. We expect each impaired driving state to possess distinct characteristics. For example, alcohol influences the function of the cerebellum more than that of the basal ganglia, whereas drowsiness has the opposite effects. Therefore, performance degradation resulting from different driving impairments should manifest different characteristics.

Complex tasks involving higher level cognitive functions (such as knowledge-based tasks) should be addressed in further studies. It is often thought that complex tasks are more sensitive to sleep loss than simple tasks. However, the tasks that most often reveal sleep loss early and profoundly are simple sustained attention RT tasks, which can hardly be considered complex tasks [37]. As there is a confounding effect of additional performance degradation due to tedium and a loss of interest in the task [37], human motivation may significantly affect the performance in complex tasks. This needs to be addressed in further studies.

Collaborations between policy makers and research engineers are essential. For example, we need to have *a priori* information of drivers to utilize the BN paradigm. Drivers' information may systematically be obtained through fatigue-related public policies or regulations. We may consider including fatigue-related tests for commercial drivers along with their training processes. Information obtained from these tests can then provide driver characteristics that are necessary for the implementation of individualized detection systems.

REFERENCES

[1] Ford Motor Company, 2005. [Online]. Available: http://www.media.ford.com
[2] N. Wright, B. Stone, T. Horberry, and N. Reed, "A review of in-vehicle sleepiness detection devices," Trans. Res. Lab., Berkshire, U.K., 2007. Tech. Rep. PPR.
[3] G. Renner and S. Mehring, "Lane departure and drowsiness—Two major accident causes-one safety system," in *4th World Congr. Intell. Transp. Syst.*, Berlin, Germany, Oct. 21–24, 1997.
[4] D. Dinges, "An overview of sleepiness and accidents," *J. Sleep Res.*, vol. 4, pp. 4–14, Dec. 1995. Suppl. 2.
[5] A. Hobson, *Sleep*. New York: Sci. Amer. Library, 1989.
[6] C. Drake, 2006, Detroit, MI: Henry Ford Hospital. private communication.
[7] S. Popkin, "Effect of day length on sleep habits and subjective on duty alertness in irregular work schedules," in *82nd Annu. Meeting Transp. Res. Board*, Washington DC, Jan. 12–16, 2003.
[8] D. McRuer and D. H. Weir, "Theory of manual vehicular control," *Ergonomics*, vol. 12, no. 4, pp. 599–633, Dec. 1969.
[9] V. Neale, T. Dingus, S. Klauer, J. Sudweeks, and M. Goodman, "An overview of the 100-car naturalistic study and findings," Virginia Tech Trans. Inst., Nat. Highway Traffic Safety Admin., Washington DC, Tech. Rep. 05-0400, 2005.
[10] I. D. Brown, "Prospects for technological countermeasures against driver fatigue," *Accident Anal. Prev.*, vol. 29, no. 4, pp. 525–531, Jul. 1997.
[11] T. Dingus, M. Hulse, S. Jahns, J. Alves-Foss, S. Confer, A. Rice, I. Roberts, R. Hanowski, and D. Sorenson, "Development of human factors guidelines for advanced traveler information systems and commercial vehicle operations: Literature review," Federal Highway Admin., Washington DC, Tech. Rep. FHWA-RD-95-153, 1996.
[12] Canada Safety Council, *Safety Canada: Drivers on Pot-Issues and Options,* Tech. Rep.
[13] C. Marberger, M. Dangelmaier, H. Widlroither, and E. Bekiaris, "User centered HMI development in the awake-project," presented at the *23rd Eur. Annu. Conf. Human Decision Making Manual Control*, Hague, The Netherlands, Oct. 11, 2004. 0-7803-8566-7/04.
[14] L. Evans, *Traffic Safety*. Bloomfield Hills, MI: Sci. Serving Soc., 2004.
[15] D. Dinges, M. Mallis, G. Maislin, and J. Powell, "Final report: Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management," NHTSA, Washington DC, Tech. Rep. DOT HS 808762, 1998.
[16] R. Knipling, "PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance," Office Motor Carrier Highway Safety, FHWA, Washington DC, Tech. Rep. HWA-MCRT-98-006, Oct. 1999.
[17] J. Yang, "Analysis and detection of driver fatigue caused by sleep deprivation," Ph.D. dissertation, Dept. Aeronaut. Astronaut., MIT, Cambridge, MA, 2007.
[18] K. P. Strohl, "Drowsy driving and automobile crashes," US Dept. Health Human Service, NHTSA, Washington DC, Tech. Rep. DOT HS 808707, Apr. 1998.
[19] National Sleep Foundation, 2007. [Online]. Available: http://www.sleepfoundation.org
[20] M. Kaneda, H. Iizuka, H. Ueno, M. Hiramatsu, M. Taguchi, and M. Tsukino, "Development of a drowsiness warning system," presented at the 14th Int. Tech. Conf. Enhanced Safety Vehicles, Munich, Germany, 1994, Paper 94-S3-O-08.
[21] Q. Wang, J. Yang, M. Ren, and Y. Zheng, "Driver fatigue detection: A survey," in *Proc. 6th World Congr. Intell. Control Autom.*, Dalian, China, Jun. 21–23, 2006, pp. 8587–8591.
[22] A. Kircher, "Vehicle control and drowsiness," Swedish Nat. Road Trans. Res. Inst., Linköping, Sweden, Tech. Rep. VTI Meddelande 922A, 2002.
[23] J. Arnedt, G. Wilde, P. Munt, and A. Maclean, "Simulated driving performance following prolonged wakefulness and alcohol consumption: Separated and combined contributions to impairment," *J. Sleep Res.*, vol. 9, no. 3, pp. 233–241, Sep. 2000.
[24] R. R. Mackie and C. D. Wylie, "Countermeasures to loss of alertness in motor vehicle drivers: A taxonomy and evaluation," in *Proc. Human Factors Soc. 35th Annu. Meeting*, San Francisco, CA, Sep. 2–6, 1991, pp. 1149–1153.
[25] T. Pilutti and A. G. Ulsoy, "Identification of driver state for lane-keeping tasks," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 29, no. 5, pp. 486–502, Sep. 1999.
[26] G. Tao and P. Kokotovic, *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*. New York: Wiley-Interscience, 1996.
[27] C. Wickens, S. Gordon, and Y. Liu, *An Introduction to Human Factors Engineering*. Reading, MA: Addison-Wesley, 1998.

[28] D. T. McRuer, *Mathematical Models of Human Pilot Behavior*. Brussels, Belgium: Advisory Group Aerosp. Res. Develop., North Atlantic Treaty Org., Dec. 1974.

[29] M. Kutner, *Applied Linear Statistical Models*. New York: McGraw-Hill, 2005.

[30] K. Kozak, R. Curry, J. Greenberg, B. Arts, M. Blommer, and L. Cathey, "Leading indicators of drowsiness in simulated driving," in *Human Factors Ergonomics Soc. 49th Annu. Meeting*, Orlando, FL, Sep. 26–30, 2005, pp. 1917–1921.

[31] D. Dinges, F. Pack, K. Williams, K. Gillen, J. Powell, G. Ott, C. Aptowicz, and A. Pack, "Cumulative sleepiness, mood disturbance, and psychomotor vigilance performance decrements during a week of sleep restricted to 4–5 hours per night," *Sleep*, vol. 20, no. 4, pp. 267–277, Apr. 1997.

[32] C. Wickens and J. Holland, *Engineering Psychology and Human Performance*. Englewood Cliffs, NJ: Prentice-Hall, 1999.

[33] T. Sheridan, *Humans and Automation*. Hoboken, NJ: Wiley, 2002.

[34] Q. Ji, P. Lan, and C. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 36, no. 5, pp. 862–875, Sep. 2006.

[35] F. Jensen, *Bayesian Networks and Decision Graphs*. New York: Springer-Verlag, 2001.

[36] J. K. Kuchar, "A unified methodology for the evaluation of hazard alerting systems," Ph.D. dissertation, Dept. Aeronaut. Astronaut., MIT, Cambridge, MA, 1995.

[37] D. Dinges and N. Kribbs, "Performing while sleepy: Effects of experimentally-induced sleepiness," in *Sleep, Sleepiness, and Performance*. Hoboken, NJ: Wiley, 1991, pp. 97–128.

Louis Tijerina received the M.S. degree in industrial and systems engineering and the Ph.D. degree in experimental psychology from The Ohio State University, Columbus.

He spent four years with the Transportation Research Center, Inc. While he was there, he supported human-factor projects being conducted at the National Highway Traffic Safety Administration's Vehicle Research and Test Center. His other previous experience includes 11 years with Battelle Memorial Institute, Columbus, where he worked on human-factor problems in transportation, defense, and medical products. Since November 1999, he has been with the Ford Motor Company, Dearborn, MI, where his research activities focus on driver distraction and driver interaction with active safety systems.

Tom Pilutti received the B.S. degrees in mechanical engineering from Purdue University, West Lafayette, IN, in 1983 and in electrical engineering from Washington University, St. Louis, MO, in 1987 and the M.S. and Ph.D. degrees in mechanical engineering from the University of Michigan, Ann Arbor, in 1989 and 1997, respectively.

Since 1990, he has been a Researcher with the Ford Research Laboratory, Dearborn, MI. He has worked on energy-management system design and control of hybrid electric vehicles and active safety. His present research involves system design and development of collision warning and avoidance systems using vision, radar, and GPS/Map information.

Joseph F. Coughlin received B.A. degree from the State University of New York, Oswego, in 1982, the A.M. degree from Brown University, Providence, RI, in 1985, and the Ph.D. degree from Boston University, Boston, MA, in 1995.

He is currently with the Massachusetts Institute of Technology, Cambridge, where he is the Director of the AgeLab, conducting research on older drivers, new technology, and safety, teaches transportation policy in the Engineering Systems Division and the Department of Civil and Environmental Engineering, and leads the U.S. Department of Transportation's New England University Transportation Center within the Center for Transportation and Logistics. He is a consultant with the automobile and insurance industries, as well as with governments worldwide, on older adult mobility. His research examines the disruptive demographics of an aging society as a driver of technical, policy, and market innovations.

Ji Hyun Yang received the B.S. degree in mechanical and aerospace engineering from Seoul National University, Seoul, Korea, in 2001 and the S.M. and Ph.D. degrees in aeronautics and astronautics from the Massachusetts Institute of Technology (MIT), Cambridge, in 2003 and 2007, respectively.

Since 2008, she has been a National Research Council Fellow working with the Human Systems Integration Laboratory, Operations Research Department, Naval Postgraduate School, Monterey, CA. Her research focuses on human systems integration, team cognition and performance, active safety enhancement systems, and human–automation interaction.

Eric Feron received the B.S. degree in mathematics from the Ecole Polytechnique, Palaiseau, France, in 1989, the M.S. degree in computer science from the Ecole Normale Supérieure, Paris, France, in 1990, and the Ph.D. degree in aeronautics and astronautics from Stanford University, Stanford, CA, in 1994.

He was a tenured faculty member of the Massachusetts Institute of Technology, Cambridge, a position he held for 12 years. Since 2005, he has been the Dutton/Ducoffe Professor in aerospace software engineering with the Georgia Institute of Technology, Atlanta. His research focuses on the applications of optimization, control, and computer sciences to transportation systems, including autonomous airborne vehicles, air transportation systems, and automotive systems. He has coauthored two books and more than 120 peer-reviewed journal and conference papers. His research interests include the application of mathematical programming to control problems, air traffic control, airport and airline operations, aircraft flight testing, and autonomous aerial vehicles. He is currently an Associate Editor for the *Journal of Field Robotics* and has been an Associate Editor for *the International Journal of Robust and Nonlinear Control*.

Dr. Feron was an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is also an Advisor to the French National Academy of Technologies.

Zhi-Hong Mao (S'96–M'05) received the B.S. degrees in automatic control and applied mathematics and the M.Eng. degree in intelligent control and pattern recognition from Tsinghua University, Beijing, China, in 1995 and 1998, respectively, the S.M. degree in aeronautics and astronautics from the Massachusetts Institute of Technology (MIT), Cambridge, in 2000, and the Ph.D. degree in electrical and medical engineering from the Harvard–MIT Division of Health Sciences and Technology, Cambridge, in 2005.

Since 2005, he has been an Assistant Professor with the Department of Electrical and Computer Engineering and the Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA.