Smart Driver Monitoring:
When Signal Processing Meets Human Factors

Amirhossein S. Aghaei*, Huei-Yen Winnie Chen†, George Liu†, Cheng Liu*, Zohreh Sojoudi†, Dengbo He†, Birsen Donmez†, Konstantinos N. Plataniotis*

* Electrical and Computer Engineering Department, University of Toronto, Canada
† Industrial Engineering Department, University of Toronto, Canada
‡ Faculty of Arts and Science, University of Toronto, Canada

Abstract

This paper provides an interdisciplinary perspective on driver monitoring systems by discussing state of the art signal processing solutions in the context of road safety issues identified in human factors research. Recently, the human factors community has made significant progress in understanding driver behaviours, and assessed the efficacy of various interventions for unsafe driving practices. In parallel, the signal processing community has had significant advancements in developing signal acquisition and processing methods for driver monitoring systems. This paper aims to bridge these efforts and help initiate new collaborations across the two fields. Towards this end, we discuss how vehicle measures, facial/body expressions, and physiological signals can assist in improving driving safety through adaptive interactions with the driver, based on driver’s state and driving environment. Moreover, by highlighting the current human factors research in road safety, we provide insights for building feedback and mitigation technologies, which can act both in real-time and post-drive. We provide insights into areas with great potential to improve driver monitoring systems, which have not yet been extensively studied in the literature, such as affect recognition and data fusion. Finally, we provide a high level discussion on the challenges and possible future directions for driver monitoring systems.

I. INTRODUCTION

Automobile driving is a demanding activity, in which the drivers simultaneously control the vehicle (laterally and longitudinally), manage hazards, and make decisions about navigation and route planning. Driving can also be a dangerous activity, with driver error a major crash risk source. The World Health Organization estimates that motor vehicle crashes kill 3000 people a day [1]. In 2013 alone, the United States reported 32,719 fatalities and estimated 2,313,000 injuries from motor vehicle crashes [2]. The majority of crashes is attributed to driver error.

Corresponding Author is Amirhossein S. Aghaei, email: aghaei@ece.utoronto.ca
An analysis of crashes between 2005 and 2007 in the United States found the driver as the critical reason in 94% of crashes, citing errors such as distraction, sleep, excessive speed, and false assumption of others’ actions [3].

Certain driver populations are at an elevated risk for vehicular crashes. Older drivers are affected by age-related declines in perceptual, cognitive, and motor abilities [4], while younger and particularly novice drivers tend to lack sufficient skills to recognize or anticipate road hazards [5]. Younger drivers may also be risk unaware and engage in potentially risky activities, such as speeding. For example, among U.S. drivers involved in fatal crashes in 2013, 15-24 year old males were the most likely to be speeding at the time of the crash [2]. In addition to driver’s age and gender, there are social-psychological factors that can also contribute to crash risk. For example, self-reported sensation seeking tendency, which is a personality trait, has been associated with risky driving (e.g., records of impaired driving and self-reported speeding [6]). Attitudes (evaluations of the potential outcomes) and perceived social norms (how other people behave, or what societal expectations there are) about unsafe driving behaviours can also play a role in vehicular crashes. For example, [7] found that drivers who drove faster than others also had a more positive attitude toward speeding. The same study also found that drivers who perceived others to drive at excessive speeds were more likely to drive fast compared to those who perceived others to comply with the limits. Medical conditions, such as sleep apnea [8], are also known to increase crash risk. Finally, more transient driver characteristics, i.e., driver state, such as fatigue, drowsiness, distraction, mental workload, and mood can also affect safety by impairing driver’s information-processing abilities as well as risk-taking tendencies.

Infotainment systems and smart personal devices (e.g., smartphones) can add to the already demanding activity of driving, and may increase crash risk. Digital interactions have become an integral part of many drivers’ daily life such that they expect digital content to be available even while driving. Furthermore, it is inevitable that a growing array of technologies find their way into the vehicle given their large economic value. It should be noted that although certain types of technology lead to distraction, technology can also enhance safety. Automobile manufacturers are constantly seeking innovative solutions that leverage emerging sensor technology and increased computational power to support the driver. In the past decade, there has been a rapidly growing interest in both the signal processing and the human factors communities to develop smart driver monitoring systems that can sense and monitor driver’s state, vehicle operation, and changes in the environment to provide drivers with useful and appropriately timed feedback.

A smart driver monitoring system can provide immediate driving assistance (e.g., automated lane departure warnings), support drivers in maintaining situation awareness while driving (e.g., alerts of incidents ahead), and foster positive behavioural changes in the long run (e.g., by providing aggregated information after a trip to help drivers understand potential consequences of unsafe driving behaviours). A smart driver monitoring system can also address some of the risk factors associated with certain driver characteristics. For example, the 2016 Chevy Malibu
Fig. 1: The inverted-U model for driver performance as a function of arousal.

has a feature tailored to teen drivers. This feature warns teen drivers when they exceed a predetermined speed limit, blocks the sound from the stereo until the front seat belts are buckled, and generates a report card on driving safety metrics (e.g., number of over-speed-limit warnings and forward collision alerts). In addition to real-time feedback on speed and seat belt use, the system provides an opportunity for parents to discuss safer driving behaviours with the teen drivers in a post-drive feedback form, thus calibrating or reinforcing a more accurate mental model of safe driving.

While Malibu’s teen driver monitoring system provides useful feedback, it does not directly address the more transient characteristics of the driver. The driver’s state, such as fatigue, anger, or mental overload (i.e., when the driver’s cognitive capacity limit has been reached), can influence driving behaviour, both in terms of information processing capabilities (e.g., reduced reaction times to hazards due to drowsiness) and decisions to perform risky behaviours (e.g., following a vehicle too closely while in a state of aggravation). The ability to monitor driver state can ensure that factors such as current mood, mental workload, and vigilance level of the driver are taken into account and feedback is provided in a timely manner (e.g., based on early detection of drowsiness rather than detection of unsafe driving performance). However, there are many challenges associated with monitoring driver state. The remainder of this article focuses on addressing these challenges, as well as reviewing important design considerations regarding driver feedback. Overall, we aim to (I) provide readers with an overview of current and emerging signal acquisition and processing techniques for driver state monitoring systems, and (II) draw insights from human factors research on driver behaviours and intervention strategies to facilitate discussions about existing techniques for smart driver monitoring systems.

II. DRIVER STATE MONITORING

The state of a driver can have a significant impact on her driving performance, a relationship best captured by the well-known Yerkes-Dodson Law [9]. The Yerkes-Dodson Law, when adapted to driver state [10], suggests an inverted-U function between driver’s arousal and driving performance. As illustrated in Fig. 1, both low arousal and high arousal are associated with performance decrements, and optimal performance occurs when there is an appropriate amount of arousal to keep the driver attentive, but not stressed. On one hand, a driver who is fatigued...
or has lost vigilance due to monotonous conditions (e.g., prolonged driving on a straight road with little traffic) is at an increased risk of losing control of the vehicle or failing to respond to hazards in time. On the other hand, strenuous driving conditions (e.g., reduced visibility or heavy traffic) or additional workload from a secondary task (e.g., having to manipulate the navigation system or engaging in a heated phone conversation) can increase stress, overload the driver, and drastically limit the cognitive resources available to drive safely.

A smart driver monitoring system should provide appropriate feedback or interventions that take into consideration the current condition of the driving environment (e.g., poor weather), the behaviours exhibited by the driver (e.g., hands off the steering wheel), as well as the driver’s physiological state (e.g., excitement or drowsiness). Smart monitoring can also be extended to monitor other in-vehicle activities, including the driver’s manual or auditory interaction with in-vehicle devices, driver’s conversation on the phone or with passengers, and conversation among passengers. Data about the in-vehicle environment and activities may help interpret driver state more accurately. However, due to limited space, this article focuses on driver state monitoring.

While the exact state of a driver may not be directly measurable, it can often be inferred from its manifestations, such as driver’s facial and body expressions (e.g., eye gaze, eye closure, yawning, and seating posture), how the driver controls the vehicle (e.g., speed, lane deviation, and headway distance), and/or driver’s physiological signals (e.g., heart rate and brain activity). Figure 2 provides a schematic overview of some examples for these three measurement categories. The remainder of this section describes the above measurement categories and the technologies associated with them, and concludes with remarks about the current algorithms for fusing the information obtained from these measures.

A. Vehicle-Based Measures

Advanced driver monitoring technologies based on vehicle measures are becoming widely available on the market and are provided by both car manufacturers and aftermarket retrofits. The following list categorizes these measures into three major groups and briefly summarizes some of the recent findings in the literature regarding these measures and their potential relationship with driver’s arousal level. It should be noted that the reported relationships are based on experimental studies. For detailed information about the experimental procedures and participant demographics please refer to the corresponding references.

- Driver input to the vehicle, such as
  - Steering: Drowsiness or low arousal may decrease the frequency of steering reversals, deteriorate steering performance, increase the amplitude of steering-wheel movements, and increase the standard deviation of steering angle. Visual distraction may cause steering neglect and overcompensation, whereas cognitive distraction or overload may lead to under-compensation (e.g., see [11]).
Fig. 2: Illustration of sensors and feedback technologies that can be used in smart driver monitoring systems.

- Braking: Both low and high arousal may cause changes in the brake and acceleration patterns (e.g., see [12]). Cognitive distraction may cause hard (high intensity) braking (e.g., see [11]), and high arousal may increase response time when braking is required (e.g., see [13]).
- Vehicle response to driver input, such as
  - Velocity/acceleration: Fatigue or low arousal may increase the standard deviation of speed. Visual distraction due to working with multimedia systems may decrease speed, increase speed variance, and result in unintended speed changes (e.g., see [11], [13]).
  - Jerk: Both low and high arousal may cause no steering correction for a prolonged period of time followed by a jerky motion to correct steering (e.g., see [14]).
- Vehicle state relative to traffic and the driving environment, such as
  - Headway distance: Distractions may both increase or decrease headway distance, depending on the type of distraction and the overall driving demands. Some studies have shown that working with email systems or using iPods lead to increased headway distance, whereas some studies have shown watching DVD players can result in decreased headway distance (e.g., see [13]).
  - Lane deviation (LD): Drowsiness or low arousal may increase standard deviation of lane position (e.g., see [13]). Distraction due to secondary tasks may also increase lane deviation (e.g., see [15], [16]).
  - Time to lane crossing (TLC): Drowsiness or low arousal may lead to irregularities in the vehicle’s tracking
Fig. 3: Example data of vehicle measures from a simulator study on driver sleepiness conducted by the authors, currently under analysis. Data shown here represent a participant’s one-hour drive, observing a 60 mph speed limit, on a two-lane track with no traffic. At the top, sleepiness is illustrated using a 5-level subjective scale (measured based on [18]), and at the bottom, the changes in various vehicle measures are illustrated. The vertical dotted lines separate the time intervals during which the sleepiness has a constant scale value. The range for each vehicle measure is illustrated on the right side.

and increase the range of deviations. Both low and high arousal may cause driving patterns with lower TLC values (e.g., see [17]).

The first two categories can be directly measured by sensors installed inside the vehicle, whereas the third category requires information regarding the driving environment or the road geometry. In the latter case, the extra information can be obtained from cameras or radars installed on the vehicle, or from the information provided by smart infrastructures (vehicle to infrastructure communication, V2I) or other vehicles (vehicle to vehicle communication, V2V). Figure 3 presents data collected in a driving simulator experiment conducted by the author, which illustrates the changes in a variety of vehicle state measures with respect to driver sleepiness and demonstrates how measures like TLC and LD can be investigated in terms of range (min, max), mean, and standard deviation. It can be seen that when driver’s sleepiness increases, a number of vehicle measures are affected. Over/under speeding happens more frequently at the higher levels of sleepiness, so does large lane-deviation variance. Moreover, during interval G and at the end of interval I, the driver changes lanes very frequently, which suggests a degradation in lane keeping ability for this particular driving scenario. Finally, at higher levels of sleepiness, TLC tends to get closer to zero, which is an indicator of unsafe driving. This effect is reflected in the graphs that indicate the number/fraction of TLC measure going below a certain threshold. Note that experimental setup used for the study in Fig. 3 is described
in Fig. 4. A comprehensive review of the literature on vehicle measures is provided in [13], [19].

Almost all major automobile manufacturers have equipped their newer high-end models with warning technologies, such as lane departure (which tracks the vehicle’s position in the lane and alerts the driver when the vehicle starts to drift laterally) and collision warning (which uses a forward looking camera or a radar to measure the distance to the lead vehicle and alerts the driver when this distance is less than a safe threshold). Furthermore, many manufacturers such as Tesla, Mercedes-Benz, BMW, Lexus, Infinity, and Honda have equipped some of their models with vehicle control intervention technologies. Some examples include lane keeping assistance (which actively applies corrective torque to the steering wheel to keep the vehicle in the lane) and collision avoidance (which activates the brakes if the driver does not respond to the collision warning).

Aside from the aforementioned systems, some manufacturers have introduced smart technologies that utilize vehicle measures to monitor driver states such as drowsiness. For example, Mercedes-Benz’s Attention Assist\(^1\) warns the driver if signs of drowsiness are detected. This system is mainly based on tracking steering wheel movements, while assessing parameters such as speed, longitudinal/lateral acceleration, and indicator and pedal usage. It also takes into account external factors such as crosswinds and the unevenness of the road. Another example is Volkswagen’s Driver Fatigue Detection System\(^2\), which continually measures steering wheel movements together with other vehicle signals as side information, and warns the driver if signs of fatigue are detected.

\(^1\)Ref. http://m.mercedes-benz.ca/en_CAn/attention_assist/detail.html

B. Facial and Body Expression Measures

Observable cues from driver’s facial and body expressions can be used to infer about driver’s arousal level. For example, eye blinks, percentage of eye closure (PERCLOS) [20], and yawning are particularly useful for detecting fatigue and drowsiness, whereas eye movements, gaze direction, and head movements provide information on drivers’ visual attention (e.g., on the road or infotainment system). These measures are heavily reliant on video processing, and thus have been traditionally difficult to implement due to computational limitations. However, recent advancements in computational power and algorithm design have made such implementations possible in real-time.

Figure 4 illustrates an example of using video cameras (dashboard mounted camera on the left and head-mounted eye-tracker on the right) in a research setting for driver state monitoring. A commercial example is the Lexus Driver Monitoring System\(^3\), which utilizes six built-in near-infrared LED sensors as well as a CCD camera facing towards the driver. It performs eye tracking, eyelid detection, and head motion detection to detect the onset of sleepiness from facial expressions and warns the driver.

A typical system for analyzing facial and body expressions is illustrated in Fig. 5. First, visual input is acquired using cameras, one or more. A camera system with an active lighting source (near infrared) is often used to alleviate the problem of extremely large illumination variations during driving. Then, the region of interest or facial/body landmarks need to be located by either direct detection from each frame, or with the help of tracking algorithms. At the next stage, meta-features, such as PERCLOS, head pose, gaze direction (and different statistics associated with them), are extracted. Lastly, the system makes a decision by looking for particular patterns in the time series of the meta-features, which often involves fusion with other signals as will be discussed later in the paper. A comprehensive review of associated algorithms and techniques used in the literature is provided by [11], [21].

The first fundamental task for this system is to successfully locate desired tracking features in each frame in real-time, which can be the face, eyes, or more precise facial landmarks. The ensuing software detection algorithm is highly dependent on the image acquisition hardware used. In earlier studies, structured illumination/camera eye-tracking systems with a near-infrared (NIR) source were widely used to exploit the special optical property of human pupils known as the “dark/bright” pupil effect. When the illumination is coaxial with gaze path, pupils reflect light back to the source and appear bright in the acquired images, while an offset of illumination placement would not create such an effect and would result in dark pupils [22]. For this system, pupils are located by simple subtraction of the alternating dark pupil frame and the bright pupil frame, a method that is computationally cheap [22]. Although this eye tracking technique is accurate in the laboratory environment, its heavy reliance on the hardware setup degrades its robustness for realistic driving scenarios. Based on advancements in computational power and machine learning algorithms, it is possible to develop intelligent computer vision solutions with large

\(^3\)Implemented in some Lexus vehicles since 2006 (e.g., Model GS 450h) and in some Toyota vehicles since 2008 (e.g., Crown Hybrid).
Fig. 5: An example of driver monitoring pipeline using facial and body expression measures, depicting typical processing steps and intermediate outputs from them: (a) video frames acquired, (b) tracking of driver’s face (yellow rectangular box) and facial landmarks (green dots), and (c) example time series plot of two meta-features: blinking (binary variable) and PERCLOS (scales from 0 to 1 representing eyes open to close).

Video data input for real-time applications that do not rely heavily on image acquisition hardware configurations. For example, the boosted cascade classifier, the underlying idea behind the Viola-Jones object recognition algorithm [23], has been incorporated in many recent camera-based driver face or eye detection algorithms (e.g., see [24]). Furthermore, tracking algorithms are commonly utilized for enhancing the processing speed of the overall system. These algorithms allow us to replace the costly processing of a full frame of the input image by the lower cost searching for the target in a more constrained manner. Algorithms such as Kalman filter or particle filter have been commonly used in both eye-trackers with structured lighting [25] and computer vision based systems [24].

Next, discriminant meta-features can be calculated based on the obtained lower-level features. This step can be seen as higher-level feature extraction. Many of the visual indicators for fatigue detection focus on the eye area, using detailed geometric measurements of iris, pupils, eyelids and sometimes the specular point on the iris. These measures are also used for gaze estimation, in which head pose estimation is also necessary. In general, there are two approaches for head movement tracking: rigid 3D head model [24] and front face orientation estimation using facial landmarks [25]. Meta-features can also be calculated based on statistics of the aforementioned visual cues, for example frequency, average intensity, or standard deviation in a certain time window. Using multiple meta-features often improves the system performance, where a fuzzy inference system [22] or Bayesian network [25] can be used for data fusion. Detection can also be performed by learning special patterns from the temporal signals using algorithms such as support vector machines (SVM) [26]. Data fusion and inference methods will be further discussed in a later part of this paper. An alternative approach is the use of holistic feature extraction algorithms.
that do not require any hand-crafted features and extract statistically discriminant features/patterns from the image frame using holistic machine learning and pattern recognition techniques. This approach is widely used in other areas (e.g., face recognition, emotion recognition) and has been shown to have great potential; however, it has not been deployed in driver monitoring yet.

Today’s cameras are empowered with abilities beyond human vision, and thus can capture extremely subtle changes, such as heart-rate using thermal imaging [27]. These technologies potentially enable non-intrusive evaluation of driver physiology. Another example is the use of pupillary dilation as an indicator of cognitive load. In addition to facial features, body expressions such as seating posture, speaking, and head gestures have been studied in the literature to infer human states [24], [28]. For example, activities related to distracted driving such as answering a cell phone, eating, and smoking can be detected by hand tracking [29]. Moreover, vision systems looking at feet and hand positions can directly monitor driver’s interaction with the vehicle, and use this information for maneuver prediction [30].

It is worth mentioning that monitoring driver’s facial and body expressions also provides possible means for determining driver’s emotions. These measures are successfully used in other contexts such as human-computer interaction studies for affect recognition. There is great potential to use facial expressions for affect recognition in the context of driver monitoring, which has not been explored yet. Finally, it is noteworthy that facial/body expressions do not necessarily need to be measured using video recorders. For example, eye movement, gaze direction, and blinks can all be monitored by attaching a set of electrical sensors to the face around the eyes and measuring the electrooculogram (EOG) signals. However, this method is highly invasive and is not suitable for realistic driving conditions. Another example of monitoring body expression without video recorders is the use of sensors embedded in the driver’s seat to measure the body pressure distribution on the seat, from which certain body postures can be detected [11].

C. Physiological Measures

Physiological measures can recognize the changes in driver’s arousal level [31]. Arousal is largely affected by the autonomous nervous system (ANS), which includes sympathetic and parasympathetic branches. The sympathetic branch generates an alerting response in stressful situations, such as a fight-or-flight response in an extreme case. This alert state can be recognized by increased breathing rate, accelerated heart rate, sweaty palms, and dilated pupils [32]. On the other hand, parasympathetic branch is mostly activated during relaxed situations, such as sleep periods. Parasympathetic activation leads to decreased breathing rate, heart rate, and blood pressure. Therefore, breathing rate, heart rate, and skin moisture can be used as indicators of the autonomous nervous system’s activity, which in turn can indicate the driver’s arousal level and alertness. These indicators/responses can be measured using breathing, electrocardiogram (ECG), and galvanic skin response (GSR) sensors. Two other commonly used
Fig. 6: Example physiological measures recorded in a simulator study (setup as shown in Fig. 4.): EEG collected using a 14-channel wireless headset; Respiration activities recorded using an inductive respiratory belt; ECG recorded using three disposable gel-electrodes located on driver’s chest; and GSR recorded using two disposable gel-electrodes located on the sole of driver’s left foot.

Physiological measures will be particularly important with higher levels of automation in emerging self-driving vehicles. When vehicle takes control (i.e., automated driving), there will be no driving performance measurement from the driver. However, physiological measures can still provide information on whether or not the driver is alert. This information together with facial/body expressions and measures of driver’s current activity can be used as a deciding factor on how to transfer back the control from the vehicle to the driver. The rest of this section provides a short review of physiological measures.

Skin moisture/conductance: GSR, also known as electrodermal response, has been commonly used to estimate driver’s physiological state [33]. GSR is a measurement of the skin conductance controlled by the sympathetic nervous system through skin’s sweat glands [34]. Highly demanding mental tasks can increase skin moisture and hence skin conductance. Electrodermal activity consists of two components: general tonic activity, which corresponds to the slower activities and background signal; and phasic activity, which
corresponds to the faster changing signal elements. Parameters such as frequency, amplitude, latency, rise time, and recovery time of the electrodermal activities are among the measures that have been used in the literature. A comprehensive review of GSR measures is provided in [35].

Respiratory activities: The respiratory frequency is generally between 9-21 cycles per minute. During relaxation, this rate can become as low as 6 cycles per minute, whereas performing highly demanding mental tasks cause higher than normal frequencies [36]. Respiratory activities can be measured either through sensors that are connected to the person’s mouth/nose, which are not suitable for driving, or using a breathing belt which measures the changes in thoracic or abdominal circumference during respiration (Fig. 4).

Eye movements: EOG monitors eye movements, hence eye gaze, by placing sensors around the eyes to measure the corneo-retinal potential between the front and the back of the eye. Furthermore, several eye-blink indicators such as blink rate, duration, and latency can be derived from EOG to analyze visual attention, mental workload, and drowsiness [31]. Despite EOG’s promising results, current measurement techniques are highly intrusive during driving and cannot be used in practice.

Heart activities: Heart rate (HR) and heart rate variability (HRV) can be used to monitor the effect of ANS on the heart. Heightened levels of task difficulty result in increased HR, whereas fatigue/drowsiness decreases HR. HRV is defined as the variation in the time interval between heartbeats, i.e., the beat-to-beat interval. In the context of driver monitoring, HRV is usually calculated by taking the Fourier transform of the HR signal. High-frequencies of HRV’s power spectrum (0.15 – 0.40 Hz) have been found to reflect parasympathetic activities, whereas low-frequencies (0.04 – 0.15 Hz) reflect both sympathetic and parasympathetic activities (see [37] and references therein). Studies have shown that changes in mental workload may affect HRV: decreases in mid-frequency bands (0.07-0.14 Hz) have been associated with increases in mental effort (e.g., [39]). Another study also reported reduction in the high-frequency component of HRV and an increase in the low- to high-frequency ratio in more stressful situations [38]. HR can be easily measured using a single sensor that tracks heartbeats from the artery pulsation, or it can be extracted from electrocardiogram (ECG), which is a recording of the projection of the electric activities of the heart on the body surface. In a standard medical ECG device, 10 electrodes (conductive pads) are attached to the person’s chest, which is rather impractical and invasive for realistic driving conditions. Therefore, recently, there has been a growing interest to measure ECG using sensors embedded on the steering wheel (Texas Instruments 4) or the driver seat (Ford5).

Brain activities: There exist two non-invasive portable measurement methods to monitor brain activities: functional near-infrared spectroscopy (fNIR) and electroencephalography (EEG). In the fNIR method, the

4http://www.ti.com/lit/pdf/tidu479
5www.medtees.com/content/ecg_seat_fact_sheet_2.pdf
concentration of (de)oxygenated hemoglobin in different parts of the brain cortex is measured using near-infrared electromagnetic waves, based on which, active parts of the cortex are determined. fNIR is particularly helpful in monitoring increased arousal due to mental workload [40]. A review of the features that can be extracted from fNIR signals, and the required processing steps, is provided in [41].

In the EEG method, two (or three including the active ground) to more than 100 electrodes are located on the scalp, each measuring an aggregate of electric voltage fields from millions of neighboring neurons. EEG signals can be expressed in terms of a number of rhythmic activities, which are usually divided into the following frequency bands: Delta (\( < 4 \) Hz), Theta (\( 4 - 8 \) Hz), Alpha (\( 8 - 12 \) Hz), Beta (\( 12 - 30 \) Hz), and Gamma (\( > 30 \) Hz) [42]. When no major cognitive or motor task is performed, large populations of neurons are synchronized and result in steady rhythmic activities. In contrast, during cognitive or motor tasks, the synchronization of these populations usually decreases (in some cases increases), which results in a decrease (or increase) in the power of corresponding oscillatory rhythms.

The waking EEG is mostly characterized by a desynchronized pattern, causing rapid, high-frequency waves in the beta range. When people are awake in a quiet relaxed state, the increased synchrony of underlying neural activity in non-stimulated brain regions results in patterns of alpha waves particularly at the posterior part of the brain. In driver monitoring systems, one of the most difficult tasks is to detect the onset of drowsiness, often referred to as NREM (non rapid eye movement) sleep stage N1 [43]. However, in EEG signals, this transition can be easily identified by the changes of alpha and theta waves. If the driver passes this stage and enters the next NREM sleep stage (N2) [43], the eye movements will cease, heart rate slows down, and body temperature decreases, preparing the body to enter deep sleep. At this stage, background EEG oscillations decrease below 5 Hz. Furthermore, these slow oscillations will be superimposed by periodic, transient EEG patterns called sleep spindles and K-complexes. These synchronization changes at different regions of the brain and their corresponding effects on the EEG characteristics in time/frequency/space domains can be used for driver state monitoring. Some examples of commonly used methods for extraction of discriminative features from EEG signals include: parametric spectral estimation (e.g., autoregressive/moving-average), non-parametric spectral estimation (e.g., Fourier/Wavelet transform), bandpass filtering, and spatial filtering (e.g., common spatial patterns, independent component analysis, surface Laplacian filtering).

Figure 7 provides an example of the changes in an EEG feature (the ratio between powers in Alpha and Theta bands for T7-O1 bipolar EEG channel) and the driver’s sleepiness level recorded in a simulator experiment. In Fig. 7, both graphs in (a) and (b) illustrate the relationship between the changes in sleepiness level and the EEG signal. Note that the sleepiness scale is a subjective scale with limited granularity (5
Fig. 7: Example of changes in brain activity with respect to sleepiness during a one hour driving simulator experiment: (a) At the top, driver’s sleepiness using a 5 level scale is illustrated, and at the bottom, the changes in the ratio of Alpha band power over Theta band power is illustrated. The vertical dotted lines separate the time intervals during which the level of sleepiness is constant. (b) The box-plot representation of changes in the distribution of Alpha/Theta power ratio for different time intervals. The labels on the horizontal axis denote the time interval and its corresponding sleepiness scale value.

Discrete levels) whereas EEG provides a non-subjective continuous measure of sleepiness. As shown in the lower plot of Fig. 7 (a), corresponding to the 3 min moving average window (dashed line), the brain wave power ratio shows an increasing trend with increasing sleepiness level, and vice versa (except a few locations, such as ‘I’ and ‘J’). A comprehensive review of various EEG processing techniques used to monitor driver’s mental workload, fatigue, and drowsiness is provided in [31].

Finally, it is noteworthy that any facial activity or eye movement can cause interference in the EEG signals, which are commonly referred to as artifacts. For example, the blue highlighted parts in Fig. 6
are the artifacts caused by eye blinks. The common practice in EEG analysis is to remove these artifacts to study the underlying brain activity. However, in the context of driver monitoring, these artifacts can provide side information regarding the facial activities, in particular eye movements. Therefore, one area with great potential for future research is to exploit EEG artifacts for extraction of features, such as blink duration/frequency, eye closures, and even eye gaze direction.

D. Data Fusion

Each of the three measurement categories discussed in the previous sections has its own pros and cons, which have to be taken into account in the design of smart driver monitoring systems. The main limitation of both vehicle-based measures and facial/body expressions is that they depend on factors that tend to manifest themselves at the late stages of fatigue or drowsiness (low arousal) or mental overload (high arousal), whereas physiological measures can provide early indicators of changes in arousal. However, since current technologies for physiological measurement require direct attachment of sensors to driver’s body, these measures might be more intrusive than other categories. Similarly, some drivers may have concerns with having a camera analyzing their facial/body expressions while driving. If the driver’s concerns regarding privacy, ease of use, and non-intrusiveness are taken into account in the design of new sensing technologies such that drivers embrace them, the combination of these three categories can provide a multifaceted driver monitoring system that can benefit from the advantages of each of these categories and provide higher accuracy and performance. Furthermore, information provided by intelligent transportation Systems [44] as well as information about driver’s interactions with his surroundings, such as conversation with other passengers, can also be utilized. Combining multiple sources of information calls for the development of new data fusion techniques. For example, [45] uses a Fuzzy Bayesian framework to combine the information obtained from facial features, ECG, photoplethysmography, temperature, and a three-axis accelerometer to monitor driver drowsiness. Still, fusion techniques for driver monitoring systems are in their early stages and are not fully explored yet. Potential reasons include: (I) relatively recent success in computationally-efficient real-time analysis of video data, (II) relatively recent advancements in non-intrusive measurement techniques for physiological signals, (III) complex nature of sensing technologies in driver monitoring which usually requires multi-disciplinary collaboration between experts from several different areas for development of effective fusion techniques. Considering the extensive body of research on data fusion in the signal processing community, there is great potential for development of new fusion techniques for driver monitoring.

III. Feedback Mechanisms

This section presents a discussion of how different feedback timings can support overall safer driving, and highlight some design considerations for the various types of feedback.
A. Real-Time Feedback

Real-time feedback is provided during driving. For example, a driver may have the habit of driving too closely behind another vehicle (i.e., tailgating), may receive feedback about an unsafe headway (e.g., warning through a collision avoidance system, verbal comment made by a passenger), and may adjust her headway accordingly. Studies have shown that real-time feedback is useful in notifying the driver about a potentially hazardous situation or an improper action that may lead to dangerous events (e.g., rear-end collision warning using time headway [46], and driver distraction feedback based on drivers’ off-road eye glances [47]). However, humans have limited attentional resources and feedback provided during driving, like any other secondary task, may consume a certain amount of attentional resources and subsequently reduce the amount of resources available to driving. For thorough discussions of the important role of attention in everyday driving, readers are referred to [48]. Also, [49] provides a detailed depiction of driver-feedback interactions, discussing how feedback can both help and interfere with the allocation of limited attentional resources.

Overall, real-time feedback design needs to be concise, to communicate a salient message to the driver, and be delivered through an appropriate modality that will be perceived by the driver with minimal additional demand. For example, an auditory alert is likely to capture a driver’s attention immediately even when the driving task is demanding [46]. However, if the driver is already cognitively overloaded, such feedback may become detrimental by interfering with the ongoing driving task. In such cases, delaying feedback, even for a few seconds, can mitigate the potential disruption while still providing some positive effects on driving.

B. Post-drive Feedback

Post-drive feedback is provided after a trip. A driver’s mental model about safe driving can guide her decisions and intentional behaviours. For example, if a driver perceives a significant increase in driving demands, she may allocate more attention to driving by pausing her conversation with a passenger. While the mental model of safe driving may be in part guided by formal learning, it is also updated continually according to driving experience, feedback, and perceived social norms. For example, memory of previous driving experiences may influence a driver’s future behaviours. However, a study has estimated that 80% of near crashes are forgotten after only two weeks [50]. Thus, post-drive feedback provided after driving may facilitate long-term memory formation of critical driving incidents to motivate behavioural changes. The Chevy Malibu teen driver feature is an example for this type of feedback. Another example would be to use video replays of driving incidents (e.g., near crashes) to help strengthen the memory of such events, and thus influencing driver’s mental model of safe driving.

C. Situation Awareness and Anticipatory Driving

Situation awareness (SA) is the perception of critical elements, the comprehension of their meaning, and the projection of their status [51]. High level of SA has been associated with expert operators in complex and dynamic
environments, including aviation, healthcare, and automobile driving [51], [52]. For example, operating at a lower level of SA, a driver may not register the object(s) he sees, e.g., a stalled truck, to be a potential hazard, while another driver, operating at a high level of SA, may not only perceive and understand the stalled truck as a potential hazard, but actively anticipate other drivers’ behaviours around the stalled truck. Furthermore, SA in driving includes an internal awareness of the driver’s own state and behaviours. For example, a driver aware of her drowsiness may decide to take a break from driving. SA may be best supported through real-time feedback, helping drivers track and understand critical elements on the road. A smart driver monitoring system that detects low arousal state of the driver may also encourage drivers to take a rest break or seek out some stimulation, such as music.

High level of SA can be demonstrated through “anticipatory driving,” a cognitive competence in identifying traffic situations through perception of cues, such that the vehicle can be positioned efficiently for potential upcoming changes in traffic [53]. Feedback intended to support anticipatory driving can focus on highlighting potential traffic conflicts ahead, and may be particularly useful for the younger inexperienced drivers, who have been systematically found to exhibit less efficient visual scanning patterns compared to experienced drivers [54], and are less proficient in interpreting the situation correctly [53], [55]. For more experienced drivers, a feedback system may still be helpful in directing their attention back to critical elements in the driving scene if they are distracted. A number of aids have been proposed in the literature to support anticipatory driving, e.g., highlighting of important cues [56], visual depiction of stopping distance [57], and visual feedback using an LED array for modulating brake pedal control [58]. However, interface designs in these studies remain largely at the conceptual stage, and have been evaluated in limited settings such as in driving simulators.

IV. CONCLUDING REMARKS AND FUTURE DIRECTIONS

Recent developments in sensing technologies as well as advancements in our understanding of human factors affecting driving performance provide opportunities for the development of smart driver monitoring systems. These systems can use vehicle measures, facial/body expressions, and physiological signals to continuously monitor driver state and also sense the environment to provide feedback to drivers or take vehicle control if there is a need. In practice, however, only vehicle measures are widely used in commercially available driver monitoring systems. The use of facial/body expressions is limited due to technical challenges in acquisition and real-time processing of these signals. For physiological measures, the most important challenge has been non-intrusive signal acquisition, which has not been possible for signals such as EEG and ECG until recently. In order to fully exploit the information provided by these three measurement categories, the information extracted from different measures can be further processed using data fusion techniques. However, considering the fact that simultaneous real-time acquisition and analysis of vehicle measures, facial/body expressions, and physiological data has not been feasible until recently, data fusion techniques have not yet been studied much in the context of smart driver monitoring.
As vehicles are becoming increasingly automated\(^6\), it is becoming important for the vehicle to be able to monitor the driver’s state to safely take control from the driver and transfer it back to the driver when there is a need. In the former scenario, i.e., the transfer of control from the driver to the vehicle, driver state monitoring can play a crucial role in the timely and accurate engagement of automatic safety systems. In such a scenario, an understanding of driver’s intentions becomes paramount. For example, the activation of the turn signal can temporarily disable lane departure warnings. However, cases such as intentional tailgating are difficult or impossible to detect. Therefore, an open question is: Should there be situations in which autonomous systems are given the authority to override driver input? It is clear that vehicles should warn drivers of drowsiness or distraction, but it is unclear whether safety systems should prevent intentionally risky driving behaviours. In the latter scenario, i.e., the transfer of control from the vehicle to the driver, the role of driver monitoring systems is as important but less obvious. Once the vehicle gains control (e.g., due to driver drowsiness, or at driver’s request), the driver monitoring system can actively monitor the driver by measuring facial/body expressions and physiological signals. Such continuous monitoring of the driver allows the system to determine when it is safe for the driver to regain vehicle control.

A major challenge in the research and development of smart driver monitoring systems is the unique interdisciplinary nature of this area, which requires a close collaboration between researchers and practitioners in both signal processing and human factors communities. The recent advancements in human factors research on driving safety can provide insights for incorporating efficient mitigation technologies into smart monitoring systems, using both real-time and post-drive feedback. This paper presented a high-level interdisciplinary overview of the smart driver monitoring systems to facilitate such collaborations.

ACKNOWLEDGMENT

The authors acknowledge support by the Natural Science and Engineering Research Council (NSERC) through the strategic partnership grant entitled DREAMs: Enhancing Driver Interaction with Digital Media through Cognitive Monitoring.

REFERENCES


