Chapter 11
Using Perceptual Evaluation to Quantify Cognitive and Visual Driver Distractions

Nanxiang Li and Carlos Busso

Abstract Developing feedback systems that can detect the attention level of the driver can play a key role in preventing accidents by alerting the driver about possible hazardous situations. Monitoring drivers' distraction is an important research problem, especially with new forms of technology that are made available to drivers. An important question is how to define reference labels that can be used as ground truth to train machine-learning algorithms to detect distracted drivers. The answer to this question is not simple since drivers are affected by visual, cognitive, auditory, psychological, and physical distractions. This chapter proposes to define reference labels with perceptual evaluations from external evaluators. We describe the consistency and effectiveness of using a visual-cognitive space for subjective evaluations. The analysis shows that this approach captures the multidimensional nature of distractions. The representation also defines natural modes to characterize driving behaviors.

Keywords Driver distraction • Active safety • Driver perception • Subjective evaluation • Driving performance

11.1 Introduction

The development of new in-vehicle technology for communication, navigation, and infotainment has significantly changed the drivers' experience. However, these new systems can negatively affect the drivers' attention, exposing them to hazardous situations leading to motor-vehicle accidents [17]. According to the study reported by *The National Highway Traffic Safety Administration* (NHTSA), over 25 % of police-reported crashes involved inattentive drivers [28]. This finding is not

---

N. Li • C. Busso (✉)

The University of Texas at Dallas, Richardson, TX 75080, USA

e-mail: nxl056000@utdallas.edu; busso@utdallas.edu

surprising since it is estimated that about 30 % of the time that drivers are in a moving vehicle, they are engaged in secondary tasks [27]. Therefore, it is important to develop active safety systems able to detect distracted drivers. A key step in this research direction is the definition of reference metrics or criteria to assess the attention level of the drivers. These reference labels can be used as ground truth to train machine-learning algorithms to detect distracted drivers.

A challenge in defining driver distraction measure methods is the multidimensional nature of the distractions caused by different tasks. Performing secondary tasks while driving affects the primary driving task by inducing visual, cognitive, auditory, psychological, and physical distractions. Each of these distractions has distinct effects on the primary driving performance [9]. During visual distractions, the drivers have their eyes off the road, compromising their situation awareness. During cognitive distractions, the drivers have their mind off the road, impairing their decision making process and their peripheral vision [27] (*looking but not seeing* [32]). A driver distraction measure should capture these facets to reflect the potential risks induced by new in-vehicle systems.

Some studies have used direct measurements derived from the driving activity. These measures include lateral control measures (e.g., lane-related measures), longitudinal control measures (e.g., accelerator-related measures, brake, and deceleration-related measures), obstacle and event detection (e.g., probability of detection measures), driver response measures (e.g., stimulus-response measures), vision-related measures (e.g., visual allocation to roadway), and manual-related measures (e.g., hands-on-wheel frequency) [4, 19, 36–38]. Other studies have used measurements from the drivers including *electroencephalography* (EEG), size of eye pupils and eye movement [4, 22, 23, 26]. Unfortunately, not all these metrics can be directly used to define labels to train machine-learning algorithms to predict distracted drivers. Some of these metrics can only be estimated in simulated conditions (e.g., event detection tasks) while others require intrusive sensors to reliably estimate their values (e.g., bio-signals).

The study addresses the problem of describing driver distraction through perceptual assessments. While common subjective evaluations such as the *NASA task load index* (NASA-TLX), *driving activity load index* (DALI), *subjective workload assessment technique* (SWAT), and *modified Cooper Harper* (MCH) scale rely on self-evaluations [39], we propose the use of external observers to separately evaluate the perceived visual and cognitive distractions—a two-dimensional space to characterize distractions. Subjects, who were not involved in the driving experiments, are invited to observe randomly selected video segments showing both the driver and the road. After watching the videos, they rate the distraction level based on their judgment. Notice that the external observers are required to have driving experience such that they can properly relate to the drivers' actions. The study uses a database recorded in real driving conditions collected with the UTDrive platform—a car equipped with multiple nonintrusive sensors [2]. The recordings include drivers conducting common secondary tasks such as interacting with another passenger, operating a phone, GPS, or radio [6, 15, 16].

Building upon our previous work [16], the chapter analyzes the consistency and effectiveness of using the proposed visual-cognitive space in subjective evaluations to characterize driver distraction. First, the scores are analyzed in terms of the secondary tasks considered in the recordings. The analysis shows high consistency with previous findings describing the detrimental effect of certain secondary tasks. The visual-cognitive space captures the multidimensional nature of driver distractions. Then, the scores provided by different external observers are compared. The inter-evaluator agreement shows very strong correlation for both visual and cognitive distraction scores. The evaluations from external observers are also compared with self-evaluations provided by the drivers. The comparison reveals that both subjective assessments provide consistent descriptions of the distractions induced by secondary tasks. Likewise, the scores from the subjective evaluation are compared with eye glance metrics. The recordings in which the drivers have their eye off the road are consistently perceived with higher visual and cognitive distraction levels. Finally, we highlight the benefits of using the visual-cognitive space for subjective evaluations. This approach defines natural distraction modes to characterize driving behaviors.

The chapter is organized as follows. Section 11.2 summarizes previous work describing metrics to characterize distracted drivers. Section 11.3 describes the experiment framework used to record the audiovisual database and the protocol to obtain the subjective evaluations. Section 11.4 analyzes the subjective evaluation in terms of secondary tasks, and the consistency in the evaluations between external raters. The section also compares the subjective evaluations of external observers with the ones collected from the drivers (e.g., self-evaluations). Section 11.5 studies the deviations observed in eye glance metrics when the driver is engaged in secondary tasks. The section discusses the consistency between perceptual evaluations and eye glance features. Section 11.6 highlights the benefits of using the proposed visual-cognitive space for subjective evaluations to characterize distraction modes. Section 11.7 concludes the chapter with discussion, future directions, and final remarks.

11.2   Related Work

Several studies have proposed and evaluated measurements to characterize driver distractions. This section summarizes some of the proposed metrics.

11.2.1   Secondary Task Performance

A common distraction metric is to measure secondary task performance [4]. In some studies, the recordings in which the driver was performing secondary tasks are directly labeled as distracted while the controlled recordings are labeled

as normal [3, 20, 40]. In other studies, the drivers are asked to complete artificial detection tasks not related to the primary driving task, such as identifying objects or events, and solving mathematical problems. The performance is measured as the effectiveness (accuracy) and efficiency (required time) to complete the task. There are various approaches that fall under this category. Examples include *peripheral detecting task* (PDT), *visual detection task* (VDT), *tactile detection task* (TDT), and *signal detection task* (SDT) [9, 23, 26, 36]. Most of the studies are conducted using car simulators, in which the stimulus can be controlled.

11.2.2   Surrogate Distraction Measurements

Studies have proposed surrogate schemes to evaluate the distraction level when the driver operates an in-vehicle technology. These methods are particularly suitable for early stages in the product design cycle of a device that is intended to be used inside the car. The *lane change test* (LCT) is one example [21]. Using a car simulator, the driver is asked to change lanes according to signals on the road while operating a particular device. The distraction level is measured by analyzing the driving performance. Another example is the *visual occlusion* approach, which has been used by automotive human factor experts as a measure of the visual demand of a particular task [11]. In this approach, the field of view is temporally occluded mimicking the eye off the road patterns for visual or visual-manual tasks. During the occlusion interval (usually set equal to 1.5 s), the subject can manipulate the controls of the device, but cannot see the interface or the control values. The time to complete the task provides an estimation of the required visual demand. However, these metrics are not suitable for our goal of defining ground truth labels to describe the distraction level of recordings collected in real traffic conditions.

11.2.3   Direct Driving Performance

Another type of attention measurement corresponds to primary task performance metrics [10, 14, 20, 22, 23, 33, 35]. They determine the attention level of the driver by directly measuring the car response [4]. These measures include *lateral control* such as lane excursions, and steering wheel pattern, *longitudinal control*, such as speed maintenance and brake pedal pattern, and *car following performance*, such as the distance to the leading car. Notice that these measurements may only capture distractions produced by visual intense tasks, since studies have shown that metrics such as lane keeping performance are not affected by cognitive load [9]. Lee et al. [19] suggested that it is important to study the entire brake response process. In this direction, they considered the *accelerator release time* (i.e., the time between the leading car brakes and the accelerator is released), the *accelerator to brake*

(i.e., the movement time from accelerator release to initial brake depress), and the *brake to maximum brake* (i.e., the time from the initial brake depress to maximum deceleration). From these measurements, they found that the accelerator release time was the most sensitive metric of braking performance.

11.2.4   Eye Glance Behavior

Movement of the eyes usually indicates where the attention is allocated [36]. Therefore, studies have proposed eye glance behavior to characterize inattentive drivers [4, 22]. This is an important aspect since tasks with visual demand require foveal vision, which forces the driver to take the eyes off the road [36]. The proposed metrics range from detailed eye-control metrics, such as within-fixation metrics, saccade profiles, pupil control, and eye closure pattern, to coarse visual behavior metrics, such as head movement [36]. The total eye off the road to complete a task is accepted as a measure of visual demand associated to secondary tasks. It is correlated with the number of lane excursions committed during the task [38]. The farther away from the road that a driver fixes his/her eyes, the higher the detrimental effect on his/her driver performance [36]. Also, longer glances have higher repercussions than few short glances [38]. In fact, when the eye off the road duration is greater than 2 s, the chances of accidents increases [4, 17]. Another interesting metric is the *percent road center* (PRC), which is defined as the percentage of time within 1 min that the gaze falls in the 8° radius circle centered at the center of the road. While visual distraction is the prominent factor that forces drivers to take their eye off the road, cognitive distractions can also have an impact on eye glance behavior. As the cognitive load increases, drivers tend to fix their eye on the road center, decreasing their peripheral visual awareness [27, 29, 30]. Therefore, lack of eye glances may also signal driver distractions.

One important aspect that needs to be defined in many of the aforementioned driver distraction measurements is the corresponding values or thresholds that are considered acceptable for safe driving [39]. In some cases, organizations have defined those values. For example, the *Alliance of Automobile Manufacturers* (AAM) stated that the total duration required to complete a visual-manual task should be less than 20 s. Metrics such as total glance duration, glance frequency, and mean single glance duration have been standardized by the *International Organization for Standardization* (ISO). In other cases, a secondary task such as manual radio tuning is used as a reference task. When a new in-vehicle task is evaluated, the driving behaviors are compared with the ones observed when the driver is performing the reference task. To be considered as an acceptable, safe task, the deviation in driving performance should be lower than the one induced by the reference task.

11.2.5 Physiological Measurements

Physiological measurements provide useful information about the internal response of the drivers' body when they are conducting secondary tasks. Although the information is collected with intrusive sensors, they provide objective, consistent, and continuous measurements describing drivers' attention (e.g., increased mental workload) [9, 23, 26]. Engstrom et al. [9] used cardiac activity and skin conductance as the physiological measurements for their study on visual and cognitive load. They showed that secondary tasks have an impact on physiological signals. Mehler et al. [23] used physiological measurements including heart rate, skin conductance, and respiration rate to study young adult drivers in a simulator. They found physiological measurements are sensitive to mental workload. Putze et al. [26] considered labeling the workload using subjective evaluation, secondary task performance and multiple physiological measurements (skin conductance, pulse, respiration, and EEG). The results suggested a strong correlation between the three measurements. If these physiological metrics are used to label whether a driver is distracted, appropriate thresholds need to be established to determine acceptable driving behaviors. The challenge is that these thresholds may vary across drivers.

11.2.6 Subjective Assessments

Subjective assessments have been proposed to measure driver distraction. The most common techniques are self-evaluation scales for subjective mental workload. Examples include the *NASA task load index* (NASA-TLX), *driving activity load index* (DALI), *subjective workload assessment technique* (SWAT), *Modified Cooper Harper scale* (MCH), and *rating scale mental effort* (RSME) [39]. For assessment of fatigue, studies have used the *Karolinska sleepiness scale* (KSS) [7]. The NASA-TLX is commonly used to rate self-perceived workload [1, 14, 18, 26]. It includes rating on six different subscales: *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort*, and *frustration*. In addition to the six NASA-TLX scales, Lee et al. [18] included a modified version of the questionnaires to assess situation awareness and perceived distraction. These self-reported evaluations were used to evaluate the workload introduced by a speech-based system to read email. Some studies use a subset of these subscales. For example, Aguilo [1] included only the *mental demand*, *temporal demand,* and *frustration scales* as part of the guidelines in designing *in-vehicle information system*s (IVISs). Harbluk et al. [14] combined eye glance behavior, braking performance, and subjective evaluations (NASA-TLX scales) to study cognitive distraction. They concluded that the drivers' ratings were closely related to the task demands.

Along with self-evaluations, subjective evaluations by external observers have also been used to characterize driver distractions [25, 31]. Sathyanarayana et al. [31] relied on perceptual evaluations to label the videos of the drivers as either distracted or not distracted. Four raters were asked to observe video recordings, and the consensus labels were used as labels for pattern recognition experiments. Piechulla et al. [25] used objective and subjective methods to assess the drivers' distraction. Their study proposed an adaptive interface to reduce the drivers' workload.

This study analyzes the consistency and effectiveness of perceptual assessments of visual and cognitive distractions provided by external evaluators. We demonstrate that the use of subjective evaluations is a valid approach that can overcome the limitations of other measurements to characterize driving behaviors.

11.3 Methodology

11.3.1 UTDrive Platform

To collect a corpus in real driving conditions, this study relies on the UTDrive car (Fig. 11.1a). This is a research platform developed at *The Center for Robust Speech Systems* (CRSS) at *The University of Texas at Dallas* (UT Dallas) [2]. Its goal is to serve as a research platform to develop driver behavior models that can be deployed into human-centric active safety systems. The UTDrive car has been custom fit with data acquisition systems comprising various modalities. It has a frontal facing video camera (PBC-700H), which is mounted on the dashboard facing the driver (see Fig. 11.1b). The placement and small size of the camera are suitable for recording frontal views of the driver without obstructing his/her field of vision. The resolution of the camera is set to 320×240 pixels and records at 30 fps. Another camera is
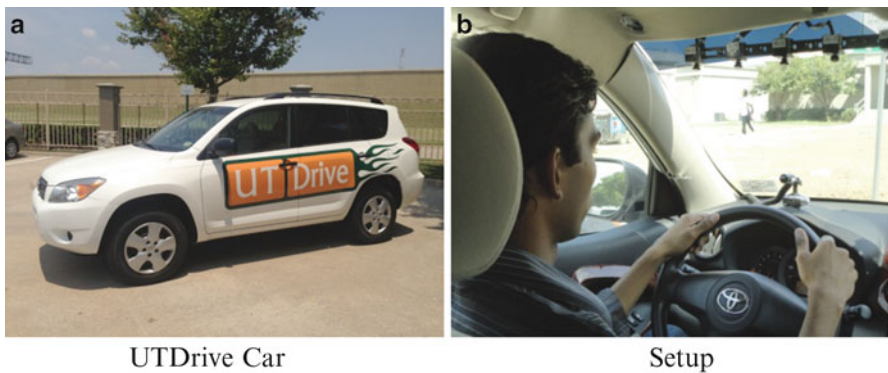

UTDrive Car   Setup

Fig. 11.1 Car platform used for the recording. (**a**) Picture of the UTDrive car (**b**) Placement of the frontal camera and the microphone array

Fig. 11.2 Dewesoft software used for recording and exporting the data. The figure shows the frontal and road videos. It also shows the instantaneous values of various CAN-bus signals

placed facing the road, which records at 15 fps at 320×240 resolution. The video from this camera can be used for lane tracking. Likewise, the UTDrive car has a microphone array placed on top of the windshield next to the sunlight visors (see Fig. 11.1b). The array has five omnidirectional microphones to capture the audio inside the car. We can also extract and record various CAN-bus signals, including vehicle speed, steering wheel angle, brake value, and acceleration. A sensor is separately placed on the gas pedal to record the gas pedal pressure.

The modalities are simultaneously recorded into a Dewetron computer, which is placed behind the driver's seat. A Dewesoft software is used to retrieve synchronized information across modalities. Figure 11.2 shows the interface of the Dewesoft software, which displays the frontal and road videos and various CAN-bus signals. For further details about the UTDrive car, readers are referred to [2].

11.3.2 Database and Protocol

A multimodal database was recorded for this study, using the UTDrive car. Twenty students or employees of the university were asked to drive while performing a number of common secondary tasks. They were required to be at least 18 years old and have a valid driving license. The average and standard deviation of the age of the participants are 25.4 and 7.03, respectively. The recordings were conducted during dry days with good light condition to reduce the impact of the environment variables. Although wet weather can lead to different challenges for the driver,
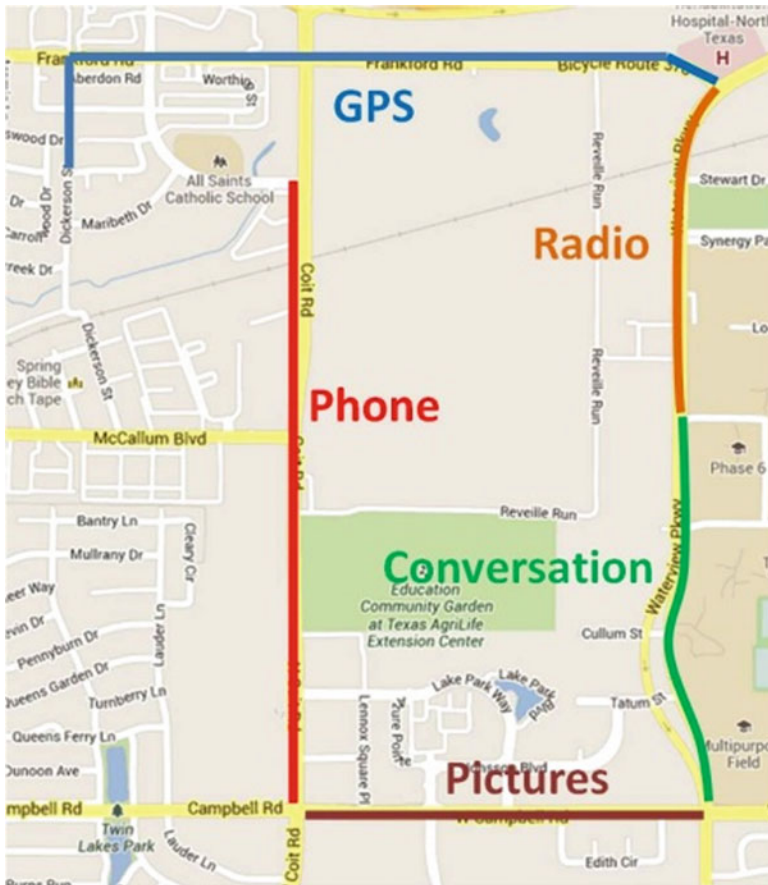
Fig. 11.3 Route used for the collection of the data. The subjects drove this 5.6 miles-long route twice. First, they were asked to perform a series of tasks starting with operating the radio and ending with a conversation with a passenger. Then, they drove the route without any in-vehicle distractions

studies have shown that crashes related to distractions are more likely to occur during dry days with less traffic congestion [13]. By collecting the data during dry days, we have relevant information for the study. The subjects were advised to take their time while performing the tasks to prevent potential accidents.

We defined a 5.6 mile route in the vicinities of the university (see Fig. 11.3). The route includes traffic signals, heavy and low traffic zones, residential areas, and a school zone. We decided to exclude streets with high speed limit (e.g., highways or freeways) from the analysis to minimize the risks in the recording. The participants took between 13 and 17 min to complete the route.

The drivers drove this route twice. During the first run, the participants were asked to perform a number of secondary tasks while driving. Among the tasks mentioned by Stutts et al. [34] and Glaze and Ellis [12], we selected the following tasks: tuning the built-in car radio, operating and following a GPS, dialing and using a cellphone, describing pictures, and interacting with a passenger. Some dangerous tasks such as text messaging, grooming, and eating were not included for security reasons. The details of the selected seven tasks are given below.

Radio: The driver is asked to tune the built-in car radio to some predetermined stations. The radio is in its standard place, on the right side of the driver.

GPS—Operating: A predefined address is given to the driver who is asked to enter the information in the GPS. The device is mounted in the middle of the windshield. The driver is allowed to adjust it before starting the recording.

GPS—Following: After entering the address in the GPS, the driver is asked to follow the instructions to the destination.

Phone—Operating: The driver dials the number of an airlines automatic flight information system (toll-free). A regular cellphone is used for this task. Hands-free cellphones are not used to include the inherent mechanical distraction.

Phone—Talking: After dialing the number, the driver has to retrieve the flight information between two given US cities.

Pictures: The driver has to look and describe randomly selected pictures, which are displayed by another passenger sitting in the front passenger seat. The purpose of this task is to collect representative samples of distractions induced when the driver is looking at billboards, sign boards, shops, or any object inside or outside the car.

Conversation: A passenger in the car asks general questions to establish a spontaneous conversation.

According to the driver resources-based taxonomy defined by Wierwille et al. [37], the selected secondary tasks include visual-manual tasks (e.g., *GPS—Operating* and *Phone—Operating*), visual-only tasks (e.g., *GPS—Following* and *Pictures*), and manual primarily task (e.g., *Radio*). The set also includes tasks characterized by cognitive demand (e.g., *Phone—Talking*) and auditory/verbal demands (e.g., *Conversation*). Therefore, they span a wide spectrum of distractions, meeting the requirements imposed by this study.

During the second run, the drivers were asked to drive the same route without performing any of the aforementioned tasks. This data is collected as a normal reference to compare the deviation observed in the driver behaviors when he/she is engaged in secondary tasks. Since the same route is used to compare normal and task conditions, the analysis is less dependent on the selected road. Overall, the database for this study consists of over 12 h of real driving recordings. More details about this corpus are provided in [6, 15].

11.3.3 Perceived Driver Distraction Using Subjective Evaluations

This study evaluates the use of subjective evaluations to quantify the level of distraction perceived from the driver. The underlying assumption is that the previous driving experience of the external evaluators will allow them to accurately identify and rank the distracting scenarios or actions, as observed in the video recordings showing the driver and the road. One advantage of this approach is that a number quantifying the perceived distraction level is assigned to localized segments in the recording. Therefore, it is possible to identify various multimodal features that correlate with this distraction metric. Using these features, regression models can be designed to directly identify inattentive drivers [16]. Another advantage is that many raters can assess the videos so the aggregated values are more accurate (see Sect. 11.4.2).

As described in Sect. 11.3.2, the database contains over 12 h of data. However, only a portion of the corpus was considered for the study to limit the evaluation time. The corpus was split into 10 s, nonoverlapped recordings. Each set contains synchronized audio and videos showing the driver and the road. For each driver, three videos were randomly selected for each of the seven secondary tasks (Sect. 11.3.2). Three videos from normal condition were also randomly selected. Therefore, 24 videos per driver are considered, which give altogether 480 unique videos (3 videos × 8 conditions × 20 drivers = 480). Eighteen students at UT Dallas with valid driver's license were invited to participate in the subjective assessment. None of the evaluators participated as drivers in the recording of the corpus. A *graphical user interface* (GUI) was built for the subjective evaluation with a sliding bar that takes continuous values between 0 and 1 (see Fig. 11.4). The extreme values are defined as *less distracted* and *more distracted*. In our previous work, we used a single, general metric to describe distraction using a similar GUI [16]. The study concluded that using a single metric for distraction was not enough to properly characterize tasks that increase the driver's cognitive load (e.g., *Phone—Talking*). To overcome this limitation, this study proposes a two-dimensional space to explicitly describe visual and cognitive distractions, separately. First, the evaluators assessed the perceived visual distraction of 80 video segments. In average, the evaluation lasted for 15 min. After a break, they assessed the perceived cognitive distraction of a different set of 80 video segments (nonoverlapped set of videos from the visual distraction evaluation). The average duration of the evaluation was 25 min.

The evaluators were instructed to relate themselves to the scenarios observed in the videos before assigning the perceived metric. We carefully instructed the evaluators with the definition of cognitive and visual distractions to unify their understanding. We follow the description given by Ranney et al. [27]. Visual distraction is defined as eye off the road—drivers looking away from the roadway. The evaluators were asked to rate the visual distraction level based on the glance behavior of the drivers. The road camera was included to help the evaluators to assess whether the observed head motions or eye glances were related to the
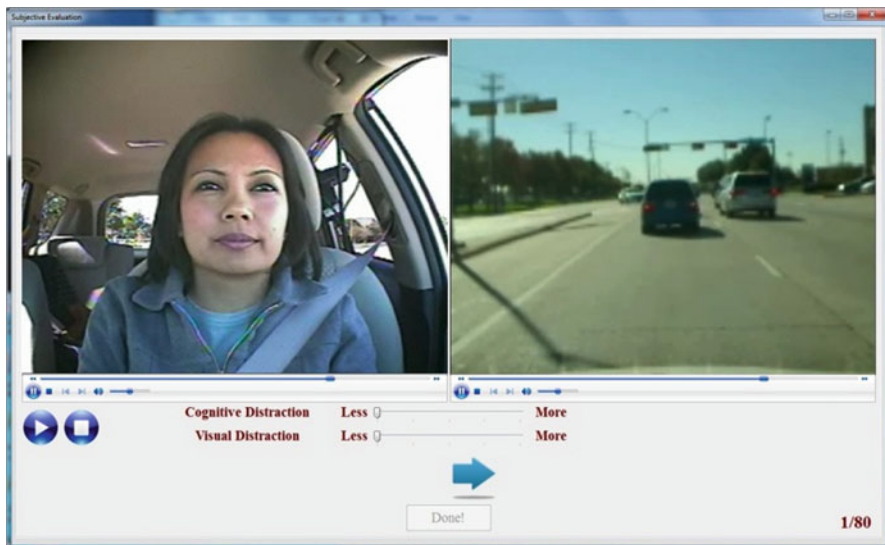
Fig. 11.4 GUI for the subjective evaluation of cognitive and visual distractions (0—*less distracted*, to 1—*more distracted*)

primary driving task. Cognitive distraction is defined as mind off the road—drivers being lost/busy in thought. For cognitive distraction, the evaluators were asked to give ratings based on his/her own judgment. However, we highlighted that facial expressions (stress level, eye pupil size, eye movements), secondary task performance (talking speed, phone dialing speed), and driving performance (vehicle in-lane position, driving speed, distance to front vehicle) can all be used to assess the cognitive distraction level. In total, each video was assessed by six independent evaluators, three for visual distractions and three for cognitive distractions.

11.4 Reliability and Consistency of Subjective Evaluations

This section validates the use of perceptual evaluations to characterize driver behaviors. We argue that employing the perceived visual and cognitive distraction assessments is a valid approach to characterize distractions. This scheme is particularly useful for cognitive distractions. While internal physiological measures can provide consistent indication of the driver's cognitive workload [23], observable driver's behaviors can only provide indirect cues [40]. We expect that evaluators can infer the expected cognitive load of the driver after observing and judging these external behaviors. First, we analyze the results of the perceptual evaluation in terms of secondary tasks (Sect. 11.4.1). Then, we study the consistency of the subjective evaluations by estimating the inter-evaluator agreement (Sect. 11.4.2). The results of the subjective evaluation are compared with self-reports from the drivers that participated in the recording (Sect. 11.4.3).
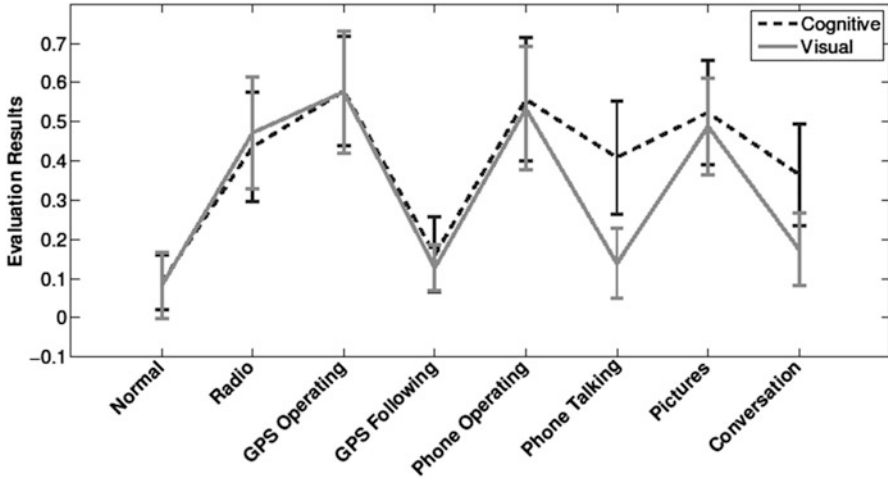
Fig. 11.5 Means and standard deviations of the perceived visual and cognitive distraction scores across secondary tasks collected with subjective evaluations

11.4.1 Analysis of Subjective Evaluations

Figure 11.5 shows the means and standard deviations of the perceived visual (solid line) and cognitive (dashed line) distractions across secondary tasks and normal conditions. The result suggests that secondary tasks identified as visually intensive activities such as *GPS—Operating*, *Phone—Operating* and *Pictures* received the highest scores for visual distractions. The cognitive distraction scores for secondary tasks that are known to increase the cognitive workload of the driver (e.g., *Phone— Talking* and *Conversation*) are higher than the corresponding visual distraction scores. These results are consistent with previous studies reporting that conversation is intrinsically a cognitive task [24]. The perceptual evaluations also agree with Bach et al. [4] who suggested that the cognitive distraction induced by using a cellphone is more detrimental than the mechanical distraction associated with operating the device.

Although Fig. 11.5 suggests that the recordings received similar cognitive and visual distraction scores for most of the secondary tasks, a closer look at the evaluation reveals that the proposed two-dimensional space captures their distinction. Figure 11.6 shows a scattering plot of the subjective evaluation across tasks and normal conditions in the visual-cognitive space. The figure shows samples covering much of the two-dimensional space. The only empty area corresponds to recording with low cognitive distractions but with high visual distractions. Notice that visual demanding tasks also induce cognitive demands. Therefore, this finding is expected. These results suggest that the subjective evaluation is effective in capturing both visual and cognitive distractions. A further discussion about the scattering plot defined by the visual-cognitive space is given in Sect. 11.6.
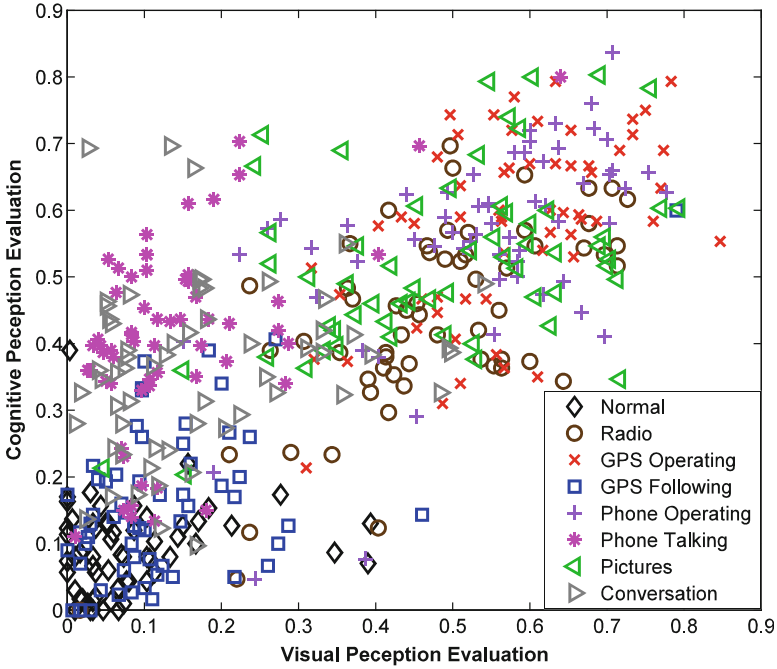
Fig. 11.6 Scattering plot of the subjective evaluation across secondary tasks in the visual-cognitive space

11.4.2 Inter-Evaluator Agreement

Since each video segment is separately assessed by three different evaluators for cognitive and visual distractions, the agreement between raters is a useful indicator of the reliability of these metrics. Stronger agreement suggests higher consistency among the evaluators, which validates the proposed approach. The analysis consists in measuring the correlation between the provided scores. For each evaluator, we calculated the average scores provided by the remaining two raters. Then, we estimate the Pearson correlation between his/her scores and the average scores. We repeat this approach for each of the three evaluators. The average correlation across evaluators is $\rho^v = 0.75$ for visual distractions, and $\rho^c = 0.70$ for cognitive distractions. These correlation values represent very strong positive relationship between the scores provided by raters. Figure 11.7 gives the correlation values for cognitive and visual distractions for each of the 18 evaluators. The correlation values are always above $\rho = 0.5$. These findings reveal high consistency for both visual and cognitive distraction evaluations. In general, visual distraction scores have higher values than cognitive distraction scores. This result and the fact that the duration of the cognitive distraction evaluation was in average 10 min longer than visual distraction evaluations (Sect. 11.3.3) suggest that assessing cognitive distractions is harder than assessing visual distractions.
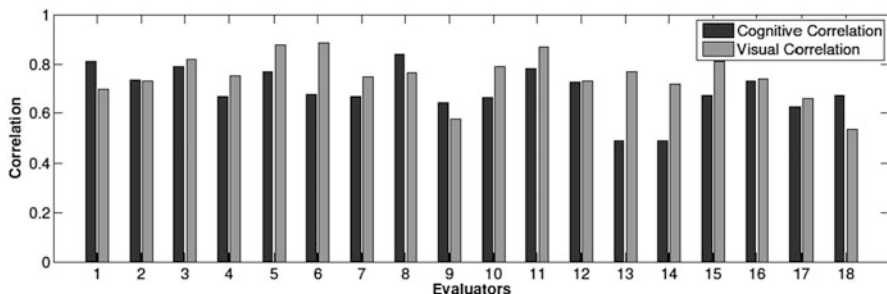
Fig. 11.7 Correlation values for cognitive and visual distractions. The results are given for each of the 18 evaluators

11.4.3   Self-Evaluations Versus External Evaluations

The most common questionnaires used to assess mental workload are based on self-evaluations conducted after the experiments [1, 14, 18, 26]. The underlying assumption in self-evaluations is that drivers are aware of the distraction level felt when they were performing secondary tasks. Therefore, they can rank the tasks that were more distracting to them. This section compares self-evaluations with the assessments provided by external observers.

A self-evaluation was collected from the drivers after recording the data to rate how distracted they felt while performing each of the secondary tasks. Unfortunately, the subjects participating in the driving recordings were not available to provide detailed assessments over small video segments. Therefore, we use a simplified methodology for this self-evaluation. First, the drivers self-evaluated their perceived distraction, without distinguishing between cognitive and visual distractions. Second, instead of evaluating several localized segments in the recording, the drivers provided a single coarse value for each secondary task without watching videos of the recordings. They used a Likert scale with extreme values corresponding to 1—*less distracted*, and 5—*more distracted*. Figure 11.8 presents the average and standard deviation values of the perceived distraction scores. The result suggests that, on average, *GPS—Operating* is regarded as the most distracted task, while *Conversation* is considered as the least distracted task. The fact that *Phone—Talking* is perceived as more distracting than *Conversation* is consistent with the conclusions by Drews et al. [8]. They claimed that the situational awareness of being in the same vehicle makes conversation with a passenger a less distracted task than a conversation with someone who is unaware of the surrounding traffic (e.g., avoiding increasing the driver's cognitive demands during decision making times).

Although the setting for the drivers' self-evaluation differs from the one used to collect evaluations from external observers (Sect. 11.3.3), the global patterns can be compared. Figures 11.5 and 11.8 show consistent patterns across secondary tasks.
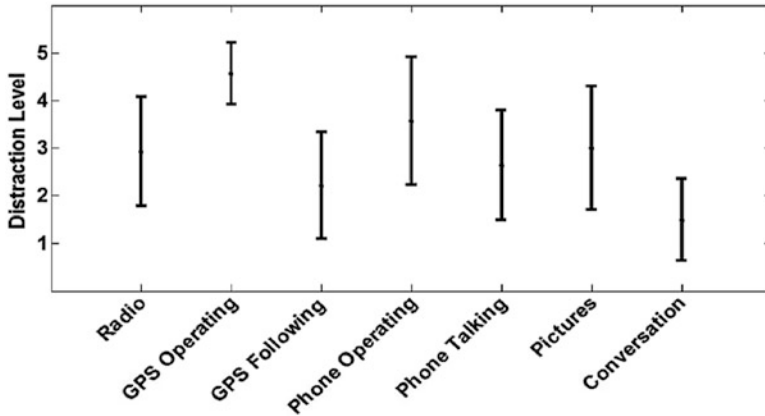
Fig. 11.8 Average distraction levels based on self-evaluations across the drivers. The figure shows the mean and standard deviation of the values assigned to each task

The ranked order of the four tasks that are perceived as the most distracting for self-evaluations are exactly the same as the corresponding ones in the cognitive and visual evaluations from external observers: *GPS—Operating*, *Phone—Operating*, *Picture*, and *Radio*. The main differences are observed in the cognitive evaluations for the tasks *Conversation* and *Phone—Talking*, which received higher values by external observers than by the drivers. Since we requested the external evaluators to specifically assess the perceived cognitive load of the driver, higher values for these tasks are expected.

The average values of self-evaluations provide coarse indicators to represent the distraction level induced by the corresponding task. Depending on the scenario, certain actions associated with secondary tasks can be more distracting than others (e.g., having a conversation in a busy traffic intersection). Self-evaluations fail to capture this inherent within-task variability. Also, drivers may fail to notice the adjustments made to complete secondary tasks (e.g., jittery steering wheel behavior, reduced speed). We believe that perceptual evaluations collected by multiple external evaluators over small segments of driving recordings can overcome these limitations.

11.5   Subjective Evaluations and Eye Glance Behavior

As discussed in Sect. 11.2.4, eye glance behaviors provide useful metrics to characterize distractions [4, 22]. This approach gives unbiased metrics to describe driver behaviors. This section compares perceptual evaluation scores provided by external observers with eye glance behavior measurements. The analysis shows that both approaches provide consistent patterns. First, we describe the eye glance

metrics used in the analysis, which are automatically extracted from the videos (Sect. 11.5.1). Then, we compare the cognitive and visual distraction scores from recordings with extreme eye glance behaviors (Sect. 11.5.2).

11.5.1 Metrics Describing Eye Glance Behavior

The drivers' glance is a reliable indicator of attention. This study relies on two glance metrics that have been previously used to characterize driver distraction: the *total eye off the road* duration (EOR), and the *longest eye off the road duration* (LEOR). These features are automatically estimated over the videos evaluated by the external observers. Given that evaluators assessed 10 s videos, we set the window analysis accordingly. EOR measures the total time within 10 s in which the drivers' glance is not on the road. As mentioned in Sect. 11.2.4, this is an important metric that is considered to assess the visual demand of IVIS. LEOR captures tasks that require longer glances, which are known to increase the chances of accidents [38].

The glance metrics are automatically extracted from the frontal camera facing the drivers using the *computer expression recognition toolbox* (CERT) [5]. CERT is a robust system that extracts facial expression features and head pose. Given the challenges in recognizing the driver's gaze in real recordings, we approximate glance behavior with the drivers' head pose, parameterized with three rotation angles (yaw, pitch, and roll). Certain videos present adverse illumination conditions or occluded faces due to the driver's hands. In these cases, CERT fails to recognize the face producing empty values. If this problem was observed over half of the duration of a video (5 s), the recording was discarded from the analysis. Otherwise, we approximate the head pose by interpolating missing values.

Head yaw (horizontal rotation) and pitch (vertical rotation) are used for eye off the road detection. We define thresholds on these angles to decide whether the driver is looking at the road. Due to the differences in the drivers' height and in their sitting preference, the thresholds are separately calculated for each individual from his/her normal driving recordings. The thresholds for head yaw and head pitch are set at their mean plus/minus two times their standard deviation, defining in average a $16° \times 16°$ rectangular region. This approach aims to replicate the $8°$ radius circle defined in the *percent road center* (PRC) calculation [36]. The frames detected as eye of the road are accumulated over the video sequence to estimate EOR. LEOR is calculated by counting the longest consecutive eye off the road frames. Both measurements are divided by the video frame rate to convert the metrics into seconds. Notice that this approach may detect as eye off the road action glances associated to the primarily driving task (e.g., checking mirrors).

Figure 11.9 shows the average values for EOR and LEOR for normal and task conditions across 20 drivers. The task condition includes the recordings from the
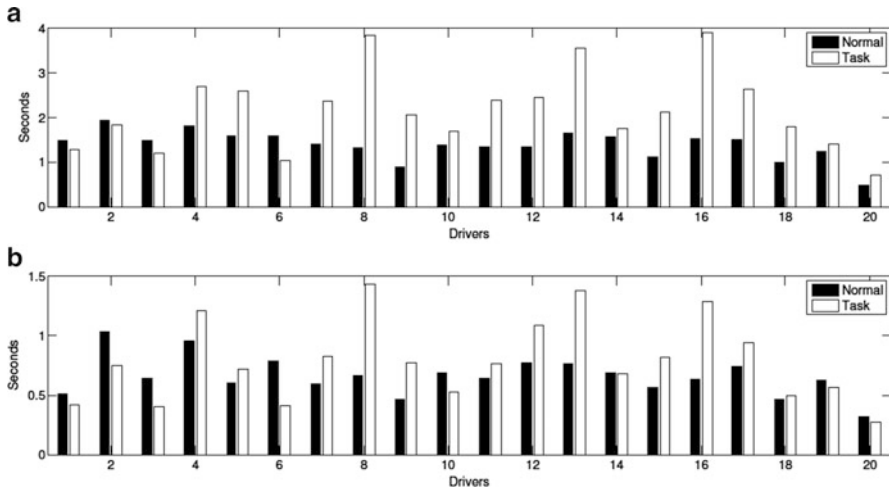
Fig. 11.9 Eye glance features extracted from 20 drivers for normal and task driving conditions. (**a**) Total eye off-road (EOR) duration in 10 s. (**b**) Longest eye off-road (LEOR) duration in 10 s

seven secondary tasks considered in this study (Sect. 11.3.2). The figure reveals that most of the drivers glance longer and more frequently when they are involved in secondary tasks. Therefore, these glance metrics are appropriate to evaluate the effectiveness of subjective evaluations.

11.5.2 Eye Glance Metrics and Subjective Evaluations

Given that EOR and LEOR have been used to characterize drivers' distractions, we expect to observe agreement between extreme values of these metrics and the subjective evaluations. We follow the approach presented by Liang et al. [20], which defined distracted recordings when the considered metrics have higher values (e.g., the upper quartile of steering error values). For each glance metric, we select a subset of the video recordings to form two extreme groups: driving recordings with low EOR or LEOR values (e.g., "normal" class), and driving recordings with high EOR or LEOR values (e.g., "distracted" class). Figure 11.10 shows the distributions for EOR and LEOR values estimated from the 10 s videos used for the subjective evaluation. The vertical lines are the thresholds defined to create the two groups, which are set so that each group has at least 72 samples to estimate reliable distributions (Figs. 11.11 and 11.12). For EOR, a recording is considered as "normal" if the EOR duration is less than 1 s, and as "distracted" if its value is more than 3 s. For LEOR, a recording is considered as "normal" if the LEOR duration is less than 1 s. Otherwise, it is considered as "distracted."
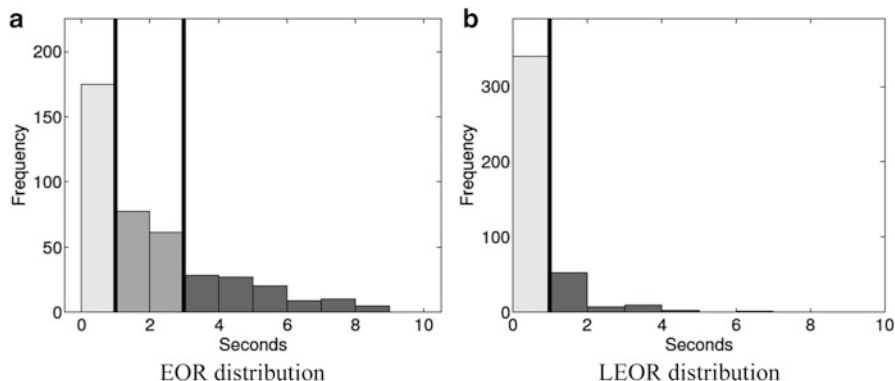
Fig. 11.10 Distributions for EOR and LEOR. The *vertical lines* give the thresholds defining the classes "normal" (*light gray*) and "distracted" (*dark gray*)
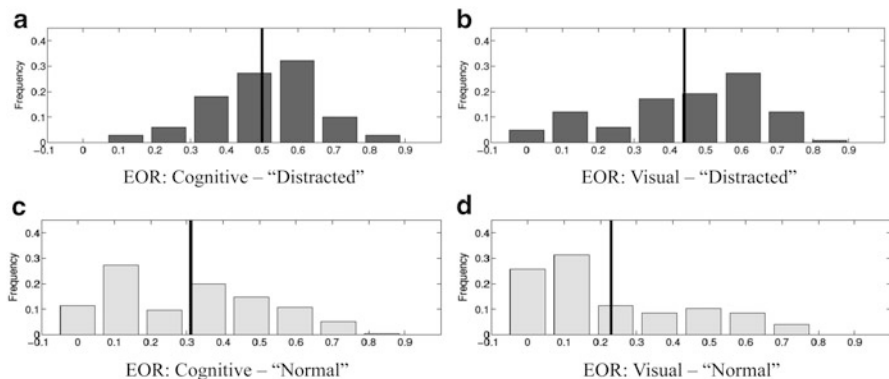

Fig. 11.11 Distribution of perceptual evaluation for extreme EOR values. (**a**) and (**b**) correspond to "distracted" class and (**c**) and (**d**) to "normal" class

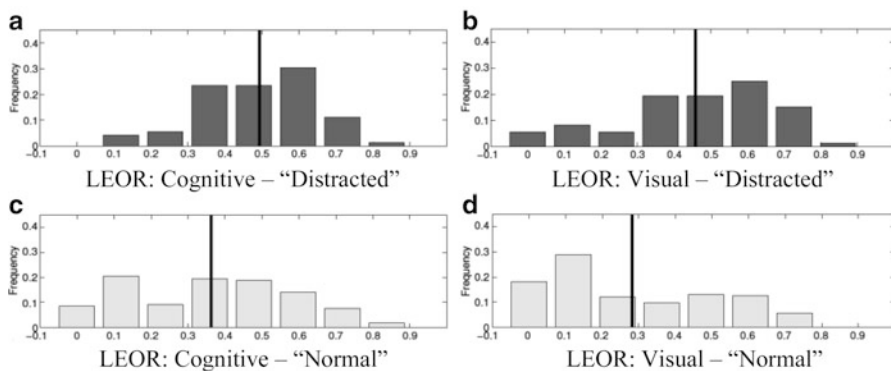**Fig. 11.12** Distribution of perceptual evaluation for extreme LEOR values. (**a**) and (**b**) correspond to "distracted" class and (**c**) and (**d**) to "normal" class

The analysis aims to identity whether the subjective evaluations capture the differences between the extreme video groups. We expect that the recordings with high EOR or LEOR values are perceived with higher distraction levels. We address this question by studying the distributions of visual and cognitive distraction scores assigned to the recordings labeled as "normal" and "distracted." Figures 11.11 and 11.12 report the results for EOR and LEOR, respectively. The vertical lines represent the means values. The distributions for the subjective evaluation are consistently skewed toward higher values for the "distracted" classes. For EOR, the mean values for both cognitive ($\mu^c_{distracted} = 0.50$) and visual ($\mu^v_{distracted} = 0.44$) distractions for the "distracted" class are significantly higher than the corresponding values for "normal" class ($\mu^c_{normal} = 0.23$, and $\mu^v_{normal} = 0.31$, respectively). The same results are observed for LEOR values.

Figure 11.11b presents a peak at 0.1. This peak may correspond to eye off the road actions associated to the primary driving task. While the EOR duration is above 3 s, the external observers may recognize that these actions do not represent distractions. Figures 11.11c and 11.12c show peaks at 0.4. These results suggest that the evaluators assigned moderate cognitive scores to recordings in which the drivers were looking at the road. These results may indicate that eye glance behaviors may provide an incomplete description of driver behaviors. As mentioned in Sect. 11.2.4, cognitive distracted drivers may have reduced peripheral visual awareness [27, 29, 30]. External observers may recognize the lack of eye glance movements as a signal of distraction.

The results reveal that subjective evaluations and eye glance behavior metrics provide consistent assessments of driver distractions (especially for visual distractions). Notice that certain eye off the road actions do not represent distractions (e.g., checking mirrors). External observers can distinguish between actions associated with primary driving tasks or secondary tasks after watching multiple cues in the road and driver videos. In these cases, the proposed two-dimensional space for perceptual evaluations can give a better representation of driver distractions.

11.6   Distraction Modes Defined by Subjective Evaluation

The final analysis in this chapter aims to highlight the benefits of using the visual-cognitive space for subjective evaluations. The results in Sect. 11.4.1 show important differences in the visual and cognitive distraction scores for certain tasks. An active safety system focusing only on visual distraction cannot provide a complete picture of the driver behaviors. These differences are captured by the proposed two-dimensional evaluation space, which defines natural distraction modes (see Fig. 11.6). The distraction modes can be automatically derived from the data by clustering the evaluations scores. The resulting modes can give a useful representation of driver distractions.
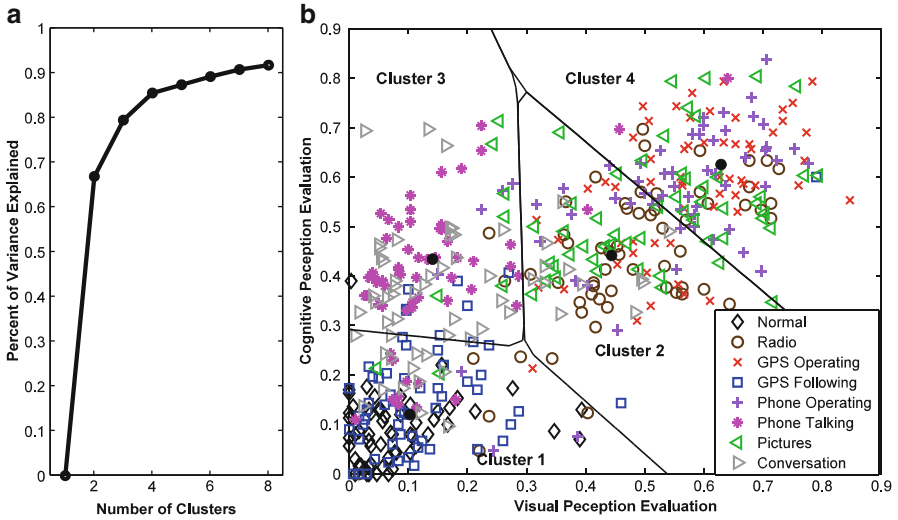
Fig. 11.13 Distraction modes defined by subjective evaluations (**a**) number of cluster defined by elbow analysis and (**b**) K-Means clusters in visual cognitive space

The clustering analysis relies on the K-means algorithm. An important aspect of the algorithm is the number of clusters, which is defined with the elbow criterion. In this approach, the number of clusters is increased, recording the percentage of variance explained by the corresponding clustering. Figure 11.13a shows that increasing the number of cluster above four does not reduce significantly the percentage of variance. Therefore, we set the number of clusters accordingly. Figure 11.13b shows the resulting clustering. The locations of the centroids suggest that drivers' distractions can be divided into (the most representative secondary tasks are given in brackets):

- *Cluster 1*—low visual and low cognitive distractions (*Normal* and *GPS— Following*).
- *Cluster 2*—medium visual and medium cognitive distractions (*Radio* and *Picture*);
- *Cluster 3*—low visual and medium cognitive distractions (*Phone—Talking* and *Conversation*);
- *Cluster 4*—high visual and high cognitive distractions (*GPS—Operating*, and *Phone—Operating*);

The proposed modes provide a new, useful representation space to characterize driving behaviors. It can be argued that clusters 3 and 4 are the most dangerous distraction modes. When a new IVIS is evaluated, multimodal features from the car and from the driver can be estimated to determine the underlying distraction mode. We are currently studying multiclass recognition problems (four class problem) and binary classification problems (one cluster versus the rest). Our preliminary analysis shows promising results in this area.

11.7 Discussion and Conclusions

This study explored the use of subjective evaluations from external observers to characterize driver behaviors. The goal is to define reference labels that can be used to train human-centric active safety systems. We conducted subjective evaluations to assess the perceived visual and cognitive distractions in randomly selected videos showing the driver and road. The analysis suggests that this two-dimensional space captures the multidimensional nature of distractions. The inter-evaluator agreement analysis shows very strong correlation for visual and cognitive assessments. The scores from external evaluators are consistent with self-evaluations collected from the drivers, and with eye glance metrics (videos with higher EOR and LEOR values are perceived more distracted).

The study suggests that perceptual evaluations from external observers have important advantages over other alternative approaches. First, multiple evaluators can provide reliable scores over short video recordings. This approach facilitates the study of relevant multimodal cues describing cognitive and visual distractions. Second, external evaluators can perceive important actions or cues that may be ignored by the drivers. For example, previous studies show the detrimental effects of the task *Phone—Talking* on the primary driving task [32, 33]. Drivers using cellphone may experience inattention blindness or selective withdrawal of attention, failing to see objects even though they are in front of them [32]. While this task was identified as the least distracted task by the self-evaluations, the external observers assigned higher scores. Third, external observers can capture the underlying driving dynamics, providing more reliable insights than metrics describing eye glance behaviors. For example, cognitive tasks reduce the drivers' peripheral visual awareness [27, 29, 30]. Therefore, lack of eye glances can signal cognitive distraction. While metrics such as eye off the road duration fail to capture these cues, external evaluators can complement their judgment by looking the driver's facial expressions.

The analysis suggests natural distraction modes to describe driver behaviors. These modes are estimated by clustering the evaluations in the visual-cognitive space. Some of these distraction modes can have a higher detrimental effect on the primary driving task (e.g., *clusters* 3 and 4). Our current research direction is to use these labels to build machine-learning algorithms to recognize the corresponding clusters. We are also planning to extend our database to include other secondary tasks, providing a better coverage of common distractions observed in real scenarios. The intended driver behavior monitoring system will provide feedbacks to inattentive drivers, preventing accidents, and increasing the security on the roads.

References

1. M. C. F. Aguilo, Development of guidelines for in-vehicle information presentation: text vs. speech, Master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, May 2004
2. P. Angkititrakul, D. Kwak, S. Choi, J. Kim, A. Phucphan, A. Sathyanarayana, and J.H.L. Hansen, Getting start with UTDrive: driver-behavior modeling and assessment of distraction for in-vehicle speech systems, in *Interspeech 2007*, Antwerp, Belgium, August 2007, pp. 1334–1337
3. A. Azman, Q. Meng, E. Edirisinghe, Non intrusive physiological measurement for driver cognitive distraction detection: Eye and mouth movements. In *International Conference on Advanced Computer Theory and Engineering (ICACTE 2010)*, vol. 3, Chengdu, China, August 2010
4. K. M. Bach, M.G. Jaeger, M.B. Skov, and N.G. Thomassen, Interacting with in-vehicle systems: understanding, measuring, and evaluating attention. In *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, Cambridge, United Kingdom, September 2009
5. M.S. Bartlett, G.C. Littlewort, M.G. Frank, C. Lainscsek, I. Fasel, J.R. Movellan, Automatic recognition of facial actions in spontaneous expressions. J. Multimedia **1**, 22–35 (2006)
6. C. Busso, J. Jain, Advances in multimodal tracking of driver distraction, in *Digital Signal Processing for In-Vehicle Systems and Safety*, ed. by J. Hansen, P. Boyraz, K. Takeda, H. Abut (Springer, New York, NY, 2011), pp. 253–270
7. Y. Dong, Z. Hu, K. Uchimura, N. Murayama, Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans. Intel. Trans. Syst. **12**(2), 596–614 (2011)
8. F.A. Drews, M. Pasupathi, D.L. Strayer, Passenger and cell phone conversations in simulated driving. J. Exp. Psychol. Appl. **14**(4), 392–400 (2008)
9. J. Engström, E. Johansson, J. Östlund, Effects of visual and cognitive load in real and simulated motorway driving. Trans. Res. Part F Traffic Psychol. Behav. **8**(2), 97–120 (2005)
10. T. Ersal, H.J.A. Fuller, O. Tsimhoni, J.L. Stein, H.K. Fathy, Model-based analysis and classification of driver distraction under secondary tasks. IEEE Trans. Intel. Trans. Syst. **11**(3), 692–701 (2010)
11. J.P. Foley, Now you see it, now you dont: visual occlusion as a surrogate distraction measurement technique, in *Driver Distraction: Theory Effects, and Mitigation*, ed. by M.A. Regan, J.D. Lee, K.L. Young (CRC Press, Boca Raton, FL, 2008), pp. 123–134
12. A. L. Glaze, J.M. Ellis, Pilot study of distracted drivers. Technical report, Transportation and Safety Training Center, Virginia Commonwealth University, Richmond, VA, USA, January 2003
13. P. Green, The 15-second rule for driver information systems. In *Intelligent Transportation Society (ITS) America Ninth Annual Meeting*, Washington, DC, USA, April 1999
14. J.L. Harbluk, Y.I. Noy, P.L. Trbovich, M. Eizenman, An on-road assessment of cognitive distraction: impacts on drivers' visual behavior and braking performance. Accid. Anal. Prev. **39**(2), 372–379 (2007)
15. J. Jain, C. Busso. Analysis of driver behaviors during common tasks using frontal video camera and CAN-Bus information. In *IEEE International Conference on Multi media and Expo (ICME 2011)*, Barcelona, Spain, July 2011
16. J.J. Jain, C. Busso. Assessment of driver's distraction using perceptual evaluations, self assessments and multimodal feature analysis. In *5th Biennial Workshop on DSP for In-Vehicle Systems*, Kiel, Germany, September 2011
17. S.G. Klauer, T.A. Dingus, V.L. Neale, J.D. Sudweeks, D.J. Ramsey, The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data. Technical Report DOT HS 810 594, National Highway Traffic Safety Administration, Blacksburg, VA, USA, April 2006

18. J.D. Lee, B. Caven, S. Haake, T.L. Brown, Speech-based interaction with in-vehicle computers: the effect of speech-based e-mail on drivers' attention to the road-way. Hum. Factors **43**(4), 631–640 (Winter 2001)
19. J.D. Lee, D.V. McGehee, T.L. Brown, M.L. Reyes, Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Hum. Factors **44**, 314–334 (Summer 2002)
20. Y. Liang, M.L. Reyes, J.D. Lee, Real-time detection of driver cognitive distraction using support vector machines. IEEE Trans. Intel. Trans. Syst. **8**(2), 340–350 (2007)
21. S. Mattes, A. Hallén, Surrogate distraction measurement techniques: the lane change test, in *Driver Distraction: Theory, Effects, and Mitigation*, ed. by M.A. Regan, J.D. Lee, K.L. Young (CRC Press, Boca Raton, FL, 2008), pp. 107–122
22. J.C. McCall, M.M. Trivedi, Driver behavior and situation aware brake assistance for intelligent vehicles. Proc. IEEE **95**(2), 374–387 (2007)
23. B. Mehler, B. Reimer, J.F. Coughlin, J.A. Dusek, Impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. Trans. Res. Record **2138**, 6–12 (2009)
24. V. Neale, T. Dingus, S. Klauer, J. Sudweeks, M. Goodman, An overview of the 100-car naturalistic study and findings. Technical Report Paper No. 05-0400, National Highway Traffic Safety Administration, June 2005
25. W. Piechulla, C. Mayser, H. Gehrke, W. Koenig, Reducing drivers' mental work-load by means of an adaptive man–machine interface. Trans. Res. Part F Traffic Psychol. Behav. **6**(4), 233–248 (2003)
26. F. Putze, J.-P. Jarvis, T. Schultz, Multimodal recognition of cognitive workload for multitasking in the car. In *International Conference on Pattern Recognition (ICPR 2010)*, Istanbul, Turkey, August 2010
27. T.A. Ranney, Driver distraction: a review of the current state-of-knowledge. Technical Report DOT HS 810 787, National Highway Traffic Safety Administration, April 2008
28. T.A. Ranney, W.R. Garrott, M.J. Goodman, NHTSA driver distraction research: past, present, and future. Technical Report Paper No. 2001-06-0177, National High- way Traffic Safety Administration, June 2001
29. E.M. Rantanen, J.H. Goldberg, The effect of mental workload on the visual field size and shape. Ergonomics **42**(6), 816–834 (1999)
30. M.A. Recarte, L.M. Nunes, Mental workload while driving: effects on visual search, discrimination, and decision making. J. Exp. Psychol. Appl. **9**(2), 119–137 (2003)
31. A. Sathyanarayana, S. Nageswaren, H. Ghasemzadeh, R. Jafari, J.H.L. Hansen, Body sensor networks for driver distraction identification. In *IEEE International Conference on Vehicular Electronics and Safety (ICVES 2008)*, Columbus, OH, USA, September 2008
32. D.L. Strayer, J.M. Cooper, F.A. Drews. What do drivers fail to see when conversing on a cell phone? In *Proceedings of Human Factors and Ergonomics Society Annual Meeting*, volume 48, New Orleans, LA, USA, September 2004
33. D.L. Strayer, J.M. Watson, F.A. Drews, Cognitive distraction while multitasking in the automobile, in *The Psychology of Learning and Motivation*, ed. by B.H. Ross, vol. 54 (Academic, Burlington, MA, 2011), pp. 29–58
34. J.C. Stutts, D.W. Reinfurt, L. Staplin, E.A. Rodgman, The role of driver distraction in traffic crashes. Technical report, AAA Foundation for Traffic Safety, Washington, DC, USA, May 2001
35. F. Tango, M. Botta, Evaluation of distraction in a driver-vehicle-environment framework: an application of different data-mining techniques, in *Advances in Data Mining. Applications and Theoretical Aspects*, ed. by P. Perner. Lecture Notes in Computer Science, vol. 5633 (Springer, Berlin, 2009), pp. 176–190
36. T.W. Victor, J. Engstroem, J.L. Harbluk, Distraction assessment methods based on visual behavior and event detection, in *Driver Distraction: Theory, Effects, and Mitigation*, ed. by M.A. Regan, J.D. Lee, K.L. Young (CRC Press, Boca Raton, FL, 2008), pp. 135–165

37. W. Wierwille, L. Tijerina, S. Kiger, T. Rockwell, E. Lauber, A. Bittner Jr, Final report supplement—task 4: review of workload and related research. Technical Report DOT HS 808 467, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, October 1996
38. Q. Wu, An overview of driving distraction measure methods. In *IEEE 10th International Conference on Computer-Aided Industrial Design & Conceptual Design (CAID CD 2009)*, Wenzhou, China, November 2009
39. K.L. Young, M.A. Regan, J.D. Lee, Measuring the effects of driver distraction: direct driving performance methods and measures, in *Driver Distraction: Theory Effects, and Mitigation*, ed. by M.A. Regan, J.D. Lee, K.L. Young (CRC Press, Boca Raton, FL, 2008), pp. 85–105
40. Y. Zhang, Y. Owechko, J. Zhang. Driver cognitive workload estimation: a data-driven perspective. In *IEEE Intelligent Transportation Systems*, Washington, DC, USA, October 2004, pp. 642–647