A Data Set of Real World Driving to Assess Driver Workload

Stefan Schneegass, Bastian Pfleging, Nora Broy, Albrecht Schmidt
University of Stuttgart, VIS
Stuttgart
{firstname.lastname}@vis.uni-stuttgart.de

Frederik Heinrich
University of Stuttgart, VIS
Stuttgart
frederik.heinrich@gmx.info

ABSTRACT

Driving a car is becoming increasingly complex. Many new features (e.g., for communication or entertainment) that can be used in addition to the primary task of driving a car increase the driver‘s workload. Assessing the driver‘s workload, however, is still a challenging task. A variety of means are explored which rather focus on experimental conditions than on real world scenarios (e.g., questionnaires). We focus on physiological data that may be assessed in an non-obtrusive way in the future and is therefore applicable in the real world.

Hence, we conducted a real world driving experiment with 10 participants measuring a variety of physiological data as well as a post-hoc video rating session. We use this data to analyze the differences in the workload in terms of road type as well as especially important parts of the route such as exits and on-ramps. Furthermore, we investigate the correlation between the objective assessed and subjective measured data.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation (e.g., HCI)**]: User Interfaces

General Terms

Human Factors

Keywords

Adaptive systems, automotive user interfaces, data set, physiological sensing, real world driving study, workload.

1. INTRODUCTION

Measuring the driver’s workload is an important indicator to estimate the driver’s ability to maneuver a car. Despite the primary task of driving a car, nowadays drivers are engaged in various other tasks. These tasks are not only related to the actual driving task but also comprise secondary and tertiary tasks [3, 7]. While secondary tasks are related to increasing safety of car, driver, and environment, tertiary tasks (e.g., using mobile phones or talking to passengers) are related to infotainment and communication and are factors that influence the driver’s workload.

Various factors are known to affect the driver’s workload (cf., [10]). One of these factors is the context in which the driver is operating the vehicle: For instance, drivers might feel more stressed during heavy rain on a jammed highway than on sunny days driving along an empty road. Another influencing factor is the general condition of the driver that may change because of time pressure, current events, or

Figure 1: Study setup: Two cameras and a smartphone are placed within the car (bottom left). The driver is connected to electrodes measuring the Heart Rate (top left). Skin Conductance Response (top right) and Body Temperature (bottom right) are measured at the drivers left hand.

http://dx.doi.org/10.1145/2516540.2516561

the driver's mood. These factors potentially increase the workload for the driver and interfere with the driving task.

The goal of using non-invasive, easy-to-use sensors to automatically assess the workload is still a challenging task: Physiological data such as the current heart rate can be measured while driving but often require the driver to wear special sensors attached to driver's body. As this might negatively influence the driver's willingness to use such a system, technologies are necessary that do not need to be installed manually: Currently, new methods to asses the physiological condition by using optical technologies (e.g., [18]) are arriving the mass market. For instance, Smartphone apps allow to measure the heart rate by analyzing face videos[1]. Bringing this technology into the car will reveal new methods for taking the driver's workload into account by creating adaptive interfaces and security systems.

The contribution of this paper is twofold. At first, we explore the different factors influencing the workload in detail. Therefore, we conducted a real-world driving study on public roads with 10 participants as we believe that workload cannot easily be assessed in a simulator study. The results show that the workload significantly differs for different road types. Furthermore, we explore the correlation between subjective data (using video rating) and objective data (using physiological sensing). Later, we look into two specific points of interest and explore their influence on the driver‘s workload. A second contribution is the data set itself: We recorded all parameters and release the data set to be publicly available for other research projects.

2. RELATED WORK

Different methods to assess the driver's workload have been explored in the automotive domain. The methods are either subjective (e.g., asking the user) or objective (e.g., measuring the users physiological properties or performance). Driver‘s workload is defined by de Waard as "the amount of information processing capacity that is used for task performance" [23]. The differences between workload and distraction has been discussed by Mehler et al. [10]. They state that distraction can also occur while the driver's workload is very low, e.g. through daydreaming. Thereby, the driver retains enough "capacity" to react appropriately on critical situations, which is not the case if a complex tertiary task induces high a workload.

2.1 Subjective Methods

One of the early approaches using questionnaires to assess the workload for users operating heavy machines is the NASA Task Load Index [4]. This questionnaire was later adopted by Pauzie for the automotive domain known as the Driver Activity Load Index [16]. The Subjective Workload Assessment Technique is another questionnaire that divides the mental workload into three areas [19]: Time Load, Mental Effort Load, and Stress Load. A very simple scale is the Rating Scale of Mental Effort, which constitutes a quick method to assess the subjectively felt effort of the driver [26]. These questionnaires are commonly used assessing the driver‘s workload.

Another approach to assess the workload is showing the user a video of the drive and ask to rate the workload for this situation using a specific scale (e.g., used by Totzke et al. in [22]). These approaches, however, have a temporary delay and may, therefore, reflect rather the perceived workload then the actual workload.

2.2 Objective Methods

In contrast to subjective rating scales, objective methods such as physiological sensing or engaging driver in secondary tasks were used. Physiological Sensing uses the reaction of the body to reflect onto the drivers workload. Therefore values such as the skin conductance response (SCR), heart rate (HR), or skin temperature are measured.

Michaels [12] showed that the SCR is related to the amount of traffic the driver is facing at the moment. The direct relation to the workload is shown by Collet et al. [2] in an experiment with air traffic controllers.

Mittelmann and Wolff [13] found that there is a strong correlation between skin temperature and emotional stress. In contrast, Or and Duffy [15] used a thermal camera as a non-intrusive way. They showed a signifcant correlation between driver's workload and facial skin temperature through a driving simulator and field experiment. However, the work of Anzengruber and Riener [1] indicates that thermal imaging does not work reliably to classify the driver's stress level.

Roscoe [21] found a strong correlation between HR and workload in studies with pilots. Meshkati [11] and Myrtek et al. [14] report that a decreased heart rate variability (HRV) indicates increased workload. HR and HRV are commonly measured by electrodes attached to the user‘s body. For instance, Riener et al. [20] measured the HR in a field study investigating the driver‘s arousal state indicating critical situations in which the driver should be aware of. However, different approaches are investigated to get rid of attaching electrodes. Ford[2] proposed a system by using the seat belt to measure the HR. Wu et al. [24] showed that the HR is observable with a SLR camera. These methods allow the integration in an non obtrusive way that might open up mass market capabilities.

In contrast to physiological sensing approaches, there are methods that engage the user in a secondary task. Then inference is drawn from the performance in this task about the workload. For the "n-back" task a sequence of stimuli is presented and the driver has to react through speaking aloud the same stimulus as the one presented n steps before. Mehler et al. [9] used this technique to evaluate the relationship between HR and SCR with the workload increased by such a task. They showed in a driving simulator study that the HR as well as the SCR are influenced by an increasing workload. The peripheral detection task (PDT) is another approach in which the user has to react on a stimulus in the peripheral field of view as fast as possible. The reaction time as well as the detection rate are measured to give insights into the user's workload level. Jahn et al. showed in a field study that the PDT is a useful method to assess the driver‘s workload [5]. Knappe et al. used small subliminal steering maneuvers to get insights into the drivers workload [8].

To our knowledge, the study presented in this paper is the first study recording workload with a comprehensive set of physiological sensors and context information in a real world driving study with at least 10 participants.

[1] `http://www.whatsmyheartrate.com/` (accessed June 25, 2013)

[2] `http://media.ford.com/article_pdf.cfm?article_id=34669` (accessed June 25, 2013)

3. REAL WORLD DRIVING STUDY

Assessing the drivers‘ workload in a simulator study is hardly possible, because drivers always know that they navigate through a virtual world. Therefore, we conducted a real world driving study consisting of a drive of about 30 minutes and a subsequent video rating session.

3.1 Apparatus and Data Collection

Three different data sets were recorded during the driving session (see Figure 1). At first, the physiological state of the driver is recorded. Hence, three sensors were attached to the participant. The skin conductance and temperature sensors were attached to the participant‘s left hand whereas the ECG was attached to the participant‘s chest. These sensors were connected to the Nexus 4 Biofeedback system[3] that has been used to record the driver's physiological data. At second, context data was collected through an Android Smartphone (Google Nexus S). In particular, the GPS position, brightness level, and acceleration were recorded. At last, two webcams (Logitech QuickCam Pro 9000 and Creative VF0610 Live! Cam Socialize HD) were used to record the driving scenario (passenger view onto the road) and a view of the driver as shown in Figure 2. As all data sets were recorded with different sampling frequencies timestamps were used to synchronize all data post-hoc.

3.2 Participants and Procedure

In total, ten participants (3 female, 7 male) aged between 23 and 57 years ($M = 35.60$, $SD = 9.06$) took part in this study. We recruited the participants within the employees of our institute in order to be covered by insurance. All of them owned a valid driver's license and brought their own car they were used to drive. The participant and the participant's car were first equipped with the different sensors by a researcher. Then, the participant was instructed to drive a specific route (cf. Figure 3) with the researcher guiding them by simple voice commands (e.g., “on the next intersection: please turn left”) from the backseat. After returning from the drive, the participant was guided to our lab and directly performed a *video rating*, evaluating the perceived workload from high to low using a slider. The video shown was a side by side composition of the video recorded while driving (cf. Figure 2).

3.3 Route

The selected route for our study has a total length of 23.6 km and consists of various road types (cf., Figure 3). For the evaluation of our study, we classified five different road types: 30 km/h zone, 50 km/h zone, highway, freeway, and tunnel. The *tunnel* in general is an ordinary road. However, we choose to add it as a special road type, because of the special conditions that may influence the driver (e.g., lighting). Furthermore, we defined different points of interest: 2x on-ramp, 2x freeway exits, 2x roundabouts, 20x traffic lights, and 2x curvy roads. Due to the fact that we conducted a real world driving study, we cannot control the environment (e.g., traffic, weather). However, we strove for a consistent setting among all participants: none of the participant drove during rush hours and the study was only conducted during daylight.

[3]www.mindmedia.nl (accessed June 25, 2013)

Figure 2: The five different road types: 30 km/h zone, 50 km/h zone, highway, freeway, and tunnel. The view of the driver camera is shown on the left side and the front view on the right side. This side by side composition video was shown to the participant during the video rating session.

4. DATA SET

The data set we recorded is publicly available as an archive of comma separated files[4] where each file contains the merged data set of the recordings of one participant. The complete data set has a size of 450 MB and consists of 2.5 million samples. It is anonymized and contains information about GPS, brightness, acceleration, physiological data, and data of the video rating. Note that the number of samples per participant varies due to different traffic conditions and driving behaviors resulting in different driving times. We excluded the video from the data set for privacy reasons.

Beyond the analysis of this paper, this data set may be used for different purposes. A first evaluation of concepts of automotive systems can be done without the need of

[4]The file can be found at: www.hcilab.org/automotive/

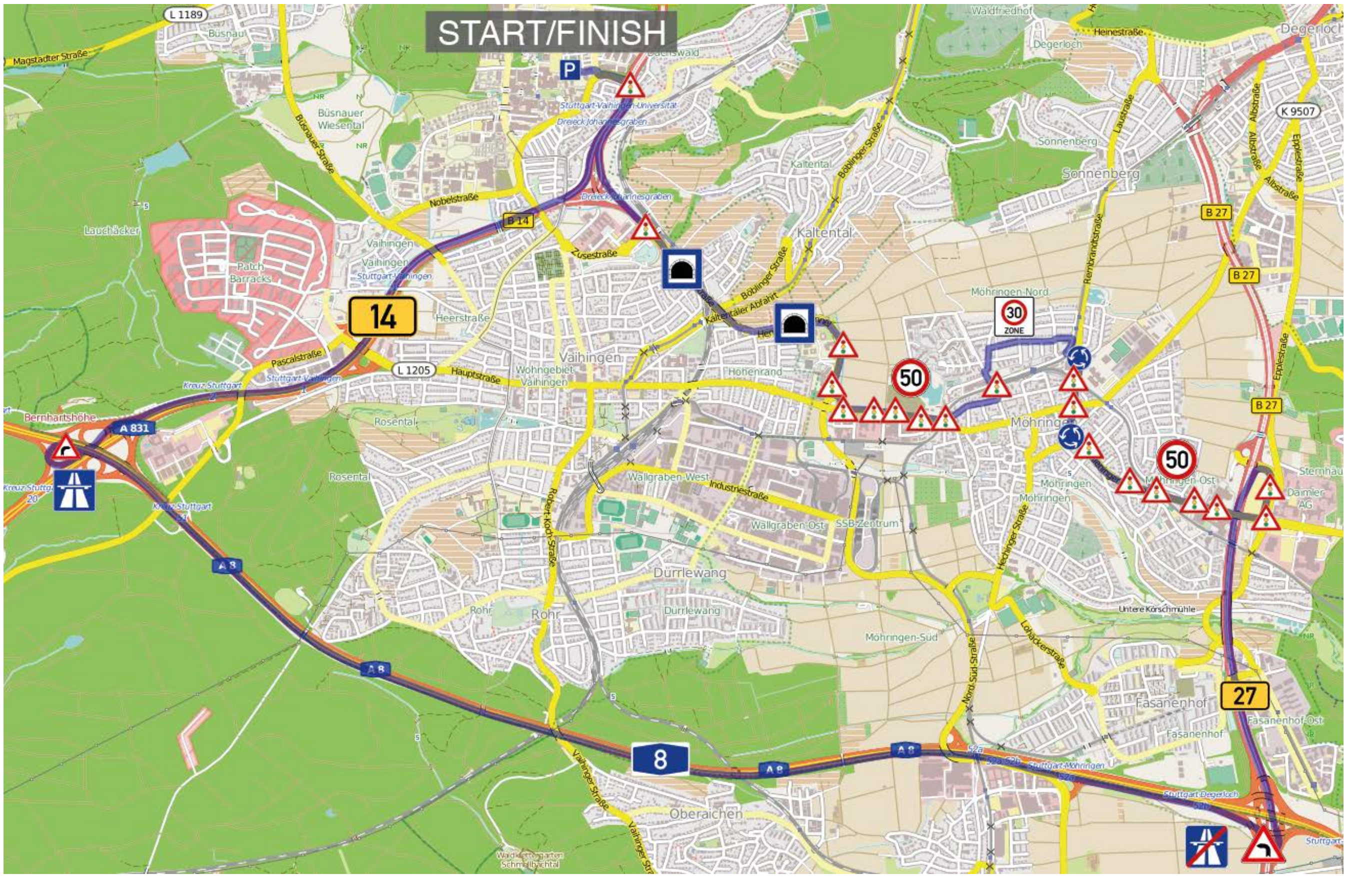

Figure 3: Map of the route each participant drove during the study. Each type of road is marked accordingly (A8: freeway, B14/B27: highway, ordinary streets (50 km/h), 30 km/h zone. All points of interest (freeway on-ramp/exit, roundabout, traffic lights, roundabouts, tunnel entry/exit) are shown with the respective symbols. Map ©OpenStreetMap contributors, tiles CC-BY-SA 2.0.

conducting an own study. We believe the data set has a value for the research community and we believe this could be a starting point for building a larger collection of data sets.

4.1 Location - GPS

The GPS information is recorded via the mobile phone. On the one hand, the GPS data consists of the longitude and latitude values (in degree) that define the position of the car. On the other hand, it contains further information about accuracy (in meter), altitude (in meter), speed (in meter per second), and bearing (in degree). In order to map the GPS data to the other data, a timestamp has been recorded as well. The information has been recorded at a frequency of 1 Hz.

4.2 Brightness and Acceleration

Further information recorded on the mobile phone were brightness as well as acceleration as perceived by the phone's sensors. The brightness is a single value measured in lumen. In contrast, the acceleration consists of three values for the X, Y, and Z direction. Both sensors have their own timestamps and are recorded at frequencies between 8 Hz and 12 Hz.

4.3 Physiological Data

Several physiological sensors are used to record information about the driver. We used a Nexus 4 physiological sensing system. The electrocardiogram (ECG, in μV) is recorded at 1024 Hz and is used to calculate the heart-rate (beats per minute) and heart rate variance at 128 Hz. Furthermore, the skin-conductance (in μS) and body temperature (in degree Celsius) are recorded at 128 Hz, as well. Again, a timestamp is recorded for all physiological data.

4.4 Videorating

The score of the post-hoc video rating is recorded between 0 (no workload) to 128 (maximum workload). In addition, the frame the user saw during the rating is recorded. It is ascending numbered starting from 0 with a frequency equivalent to the video frame-rate of 29 frames per seconds.

4.5 Data Extrapolation

The data is recorded at different sample rates. Hence, some data needs to be extrapolated to create a uniformed data set. We chose to extrapolate the data to the highest frequency keeping all available information of the sensor with the highest sample rate (ECG).

5. ANALYSIS AND DISCUSSION

In the following we present the results of the study. At first, the correlation between the objective and subjective measures is investigated. Afterwards, the statistical differences between the road types are shown as well as the statistical differences between points of interest and road type.

5.1 Data Preparation

Before evaluating the recorded data, it needs to be prepared to remove noise effects as well as to normalize the physiological properties of each participant. We modified the data in several steps. At first, we sampled the data up to one sample per second, taking the mean of each value. We used the acceleration values to create a force vector. This vector is used rather than the force values for each dimension. Next, we normalized the physiological data as well as the video rating results to achieve comparable values between all participants in the range of 0 to 1.

In this evaluation we focus on two physiological values (Skin Conductance Response (SCR) and Body Temperature (BTemp) as suggested by related work (cf., [12, 13])), the results of the Video Rating (VR), and the actual driving speed.

5.2 Comparing Subjective and Objective Data

At first, we compared the subjective measurement (VR - cf. Figure 4) with the objective measurements (SCR and BTemp). Hence, we conducted correlational research comparing the VR to the physiological values using Pearson‘s correlation coefficient. The SCR and VR, $r(17725) = .202$, $p < .001$, as well as the BTemp and VR, $r(17725) = .128$, $p < .001$, are positively correlated. The correlations are both statistically significant, however, the effect size is small.

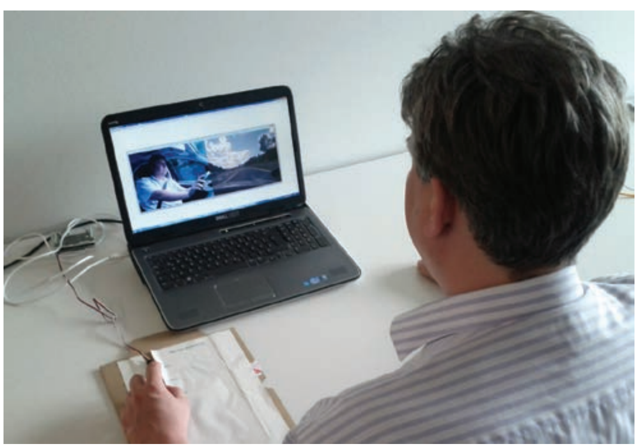

Figure 4: A participant is performing the video rating task.

Evaluating all participants individually, we see a high variance in the the data. For instance, the data of participant # 10 shows high correlation for VR with SCR, $r(1903) = .689$, $p < .001$, and VR with BTemp, $r(1903) = .449$, $p < .001$ (cf., Figure 5). In contrast, the data of participant # 6 shows only a significant correlation for BTemp, $r(1710) = .072$, $p < .01$, and a no significance for the SCR, $r(1710) = .043$, $p = .078$. Thus, the data highly differs from driver to driver. This needs to be taken into account when using these values for assessing the workload.

5.3 Impact of Road Types

Next, we evaluated the differences between the five road types. Since the values highly depend on each other and the different road types are not equally distributed within our sample (cf., Figure 5), we chose to use the mean values of each participant on each type of road. Thus, we eliminate most of the dependencies in the data and create an equal distribution.

The results show that the physiological data (SCR and BTemp), and hence the workload, is influenced by the road type. The variance in the data is high (cf., Figure 6), which indicates that all types of roads have situations in which the workload is high. Furthermore, we used a repeated measures analysis of variances (ANOVA) to investigate statistically significant differences. A Shapiro-Wilk test shows for all cases that the assumption of normal distribution is not violated.

Table 1: Overview of the mean and standard deviation of the normalized skin conductance response (SCR) and body temperature (BTemp) on the different road types.

Road Type	M_{SCR}	SD_{SCR}	M_{BTemp}	SD_{BTemp}
30 km/h zone	.482	.178	.357	.152
50 km/h zone	.423	.152	.484	.137
Highway	.343	.110	.487	.156
Freeway	.271	.121	.522	.155
Tunnel	.394	.223	.468	.266

Skin Conductance Response.

The SCR is lowest for the freeway and highest for the 30 km/h zone (cf. Table 1). Mauchly‘s test indicates that the assumption of sphericity had been violated, $\chi^2(9) = 17.890$, $p = .041$, therefore, degrees of freedom were corrected using Greenhouse-Geisser estimation of sphericity, $\epsilon = .529$. The ANOVA reveals statistically significant differences within the five road types, $F(2.116, 19.042) = 6.756$, $p < .05$, $\eta^2 = .429$. A Least Significant Difference (LSD) post-hoc test reveals a statistically difference between all road types, $p < .05$, except for Tunnel with 50 km/h zone, $p < .438$, and highway, $p < .439$. This can be explained by the fact that the Tunnel in our route is at a highway with a speed limit (50 km/h).

Body Temperature.

The BTemp is lowest for the 30 km/h and highest for the freeway (cf. Table 1) indicating that the workload is highest for the 30 km/h zone and lowest for the freeway. Again, Mauchly‘s test indicated that the assumption of sphericity had been violated, $\chi^2(9) = 27.069$, $p = .002$, therefore, degrees of freedom were corrected using Greenhouse-Geisser estimation of sphericity, $\epsilon = .357$. After the correction, the ANOVA does not reveal statistically significant differences within the road types $F(1.427, 12.842) = 1.305$, $p = .292$, $\eta^2 = .127$. Even with the ANOVA not revealing significant results, the data indicates that at least the 30 km/h zone is different from the other road types (see Figure 6).

Driving Speed.

On all five different road types, the speed limit is different. Furthermore, the driving situation (e.g., traffic, weather) has an influence on driving speed. Again, Mauchly‘s test is significant. Hence, the assumption of sphericity had been

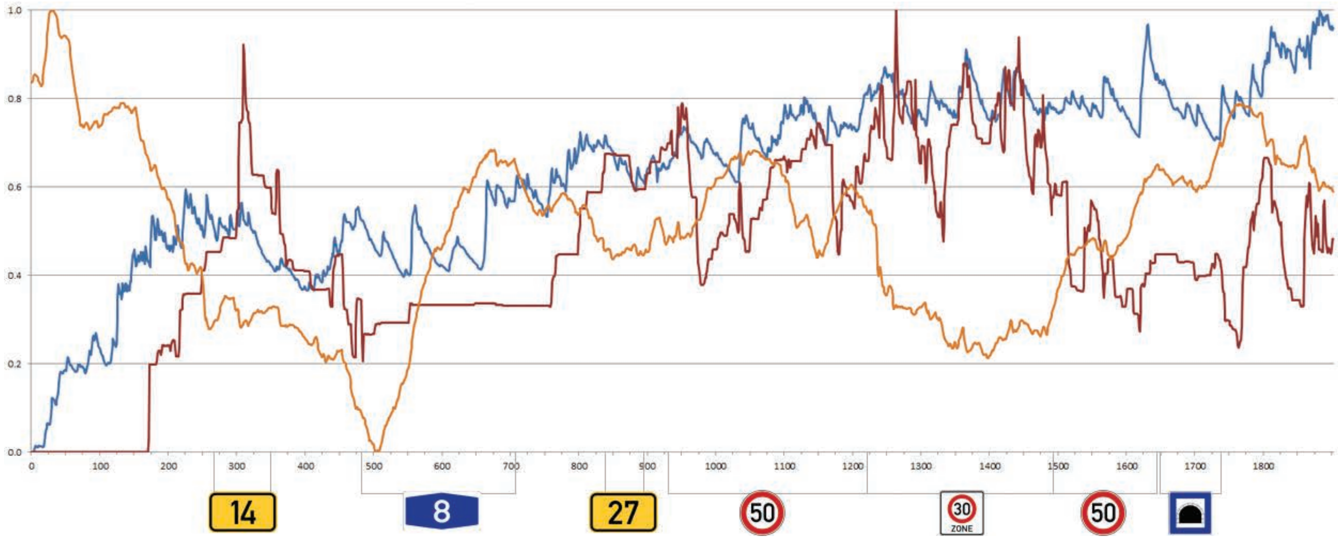

Figure 5: The Graph shows the normalized Skin Conductance Response (blue), normalized Body Temperature (orange), and the normalized result from the Video Rating (red) of a single user (User #10).

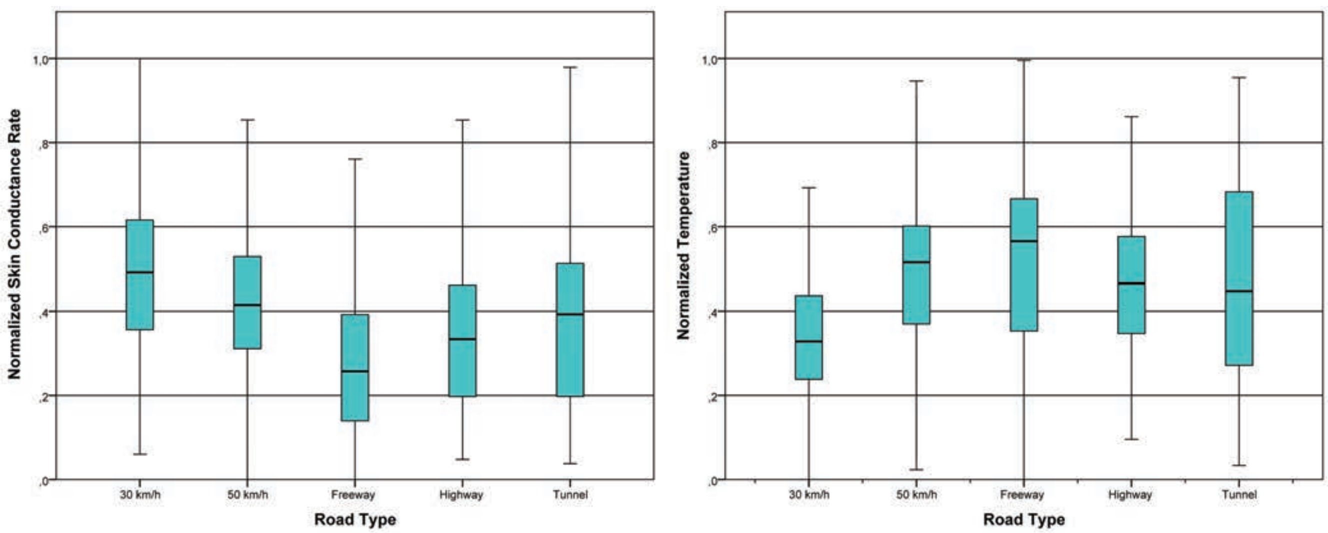

Figure 6: Boxplots for Skin Conductance Response and Body Temperature for each of the five road types.

violated , $\chi^2(9) = 37.846$, $p = .000$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimation of sphericity, $\epsilon = .407$. Nevertheless, the ANOVA shows statistically significant results, $F(1.628, 14.649) = 444.505$, $p < .05$. A Least Significant Difference (LSD) post-hoc test reveals a statistically difference between all road types, $p < .05$, except for highway with freeway, $p < .728$, because both road types have similar amount of traffic and roughly the same speed limits.

Video Rating.

In the VR session, the participant rated the highway lowest and the 30 km/h zone highest (cf., Table 1). The assumption of sphericity had been violated, shown by Mauchly‘s test of sphericity, $\chi^2(9) = 20.589$, $p = .017$, thus, degrees of freedom were corrected using Greenhouse-Geisser estimation of sphericity, $\epsilon = .601$. Between the road types, the ANOVA does not reveal any statistically significant difference, $F(2.405, 21.647) = 1.249$, $p = .312$, $\eta^2 = .122$. Again, the highest difference is between the 30 km/h zone and the other road types.

Discussion.

Interpreting the physiological data, the road type has an influence on the driver‘s workload. The workload seems to be high especially in the 30 km/h zone (low BTemp and high SCR and VR) that contains spots in which the driver has to decide who has the right of way that may increase the workload. Furthermore, there are a many parked cars that are potentially sources for unexpected events such as pedestrians crossing the street, playing children, or car doors that are carelessly opened. In contrast, the freeway (high

BTemp and low SCR) is very predictable and does not need that much attention due to larger distances between the cars. These results match the results from Micheals et al. [12] as well as from Mittelmann and Wolff [13].

5.4 Points of Interest

We identified five different categories of points of interest (POI): on-ramps, exits, roundabouts, traffic lights, and very curvy road segments. In this evaluation we focus on on the freeway on-ramp and exit. Hence, we compare the SCR, BTemp, and VR of these POI with the average of the freeway by using a series of t tests.

The SCR increases at both POI (on-ramp: $M = .409$, $SD = .095$; exit: $M = .328$, $SD = .152$) compared to the average of the freeway ($M = .271$, $SD = .122$). A dependent t test shows that the difference between on-ramp and freeway average is statistically significant, $t(9) = -3.546$, $p < .05$. However, the difference between exit and freeway average is not statistically significant, $t(9) = -1.624$, $p > .05$.

Investigating the BTemp, we see a reduced BTemp on the on-ramp ($M = .437$, $SD = .210$) compared to the average of the freeway ($M = .522$, $SD = .155$) but an increased BTemp on the freeway‘s exit ($M = .561$, $SD = .145$). A dependent t test shows no statistically significant results comparing the average freeway‘s BTemp with on-ramp, $t(9) = 1.668$, $p > .05$, and exits, $t(9) = -1.176$, $p > .05$.

In addition to the objective methods described above, the subjective VR shows similar results. Both, on-ramp ($M = .463$, $SD = .285$) and exit ($M = .384$, $SD = .239$) are increased compared to the average freeway ($M = .302$, $SD = .171$). A dependent t test shows statistically differences for the on-ramp, $t(9) = -2.643$, $p < .05$, but no difference for the exit, $t(9) = -1.895$, $p > .05$.

Summing these results up, the objective as well as the subjective methods indicate that the POI result in a different workload compared to the average freeway. Especially the on-ramp show statistically significant increased driver‘s workload.

5.5 Limitations

There are some limitations in the presented study. At first, we used the data directly from starting the car till the end. Physiological data can be biased due to nervousness. However, we think that this effect is rather low because we tried to create a comfortable environment (e.g., using the participants car). Second, the number of streets used is very limited. We strive using a representative set of streets but the characteristics of streets can be very different and can depend on the region (e.g., more stressful in alpine regions). Third, even though we measured many different physiological values in this study, we could only present a limited number of results in this work. Many parts of the data set might be evaluated using specific algorithms or more sophisticated approaches. Thus, we published the collected data set to encourage further investigations.

6. TOWARDS WORKLOAD-ADAPTIVE SYSTEMS

Adapting the in-car interface to the driver‘s state can be used keeping the complexity at a suitable level. Information that is not necessary at the moment or the shortcut of a system that is potentially troublesome could be hidden during high workload. In the following we describe three different areas of application in which such a system is useful.

Advanced Driver Assistance Systems.

As soon as the workload of the user is high, the distraction of the user is increased as well. Therefore, the reaction to unexpected events (e.g., a breaking car in front) is lowered. A system that uses different settings for the driver assistance systems can react on the different workload and increase its involvement into the driving task (e.g., pre-load the brake pressure).

Automotive User Interfaces.

The information the driver can assimilate easily depends on the current workload. If the driver‘s workload is low, a complex visualization of the user interface may reduce the risk of an increased distraction due to daydreaming [10]. On the other hand, if the driver is in a high workload situation, a complex visualization would demand too much of the driver and thus increase distraction. Hence, an interface with a reduced complexity would be easier and more safe to operate. For instance, a workload adaptive navigation system could be easier to operate in high-workload conditions. Ziegler et al. propose a system that shows routing information depending on the driver‘s knowledge of the route [25]. They show that the adaptive route information is preferred. Such systems could easily take the workload into account to be more effective. Another idea could be to make the driver aware of the current workload. This could for instance be done using haptic feedback [6].

Communication Systems.

Nowadays, communication is ubiquitous. Although it is prohibited in many countries, drivers use mobile phones while driving causing an increased workload and level of distraction. In many cases this increased level does not result into situations that are dangerous for the driver. However, in cases in which the workload is already high, this increase can lead into dangerous situation. One can imagine a system that postpones incoming communication (e.g., phone calls) until the level of workload is appropriate. Similarly, the caller could be informed about the current situation by transmitting certain context information before or during a phone call [17].

7. CONCLUSION

In this paper we present a real world driving study assessing different physiological and contextual information. We created a data set containing about 2.500.000 samples measured in a real world driving study including a video rating session taking part afterwards. This data set is publicly available.

The results from an initial evaluation show significantly different physiological values for different road types. Furthermore, we show a correlation between video rating and the physiological values. We present three concepts of adaptive systems in the automotive domain that will benefit from a workload adaption.

8. ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme

([FP7/2007-2013]) under grant agreement no. 323849. Furthermore, this project was partly funded from the German Research Foundation within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

9. REFERENCES

[1] B. Anzengruber and A. Riener. "facelight": potentials and drawbacks of thermal imaging to infer driver stress. In *Proc. AutomotiveUI '12*, pages 209–216, New York, NY, USA, 2012. ACM.

[2] C. Collet, G. Delhomme, and A. Dittmar. Subjective aspects of mental workload in air-traffic control, 2003.

[3] G. Geiser. Man Machine Interaction in Vehicles. *ATZ*, 87:74–77, 1985.

[4] S. G. Hart and L. E. Stavenland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati, editors, *Human Mental Workload*, chapter 7, pages 139–183. Elsevier, 1988.

[5] G. Jahn, A. Oehme, J. F. Krems, and C. Gelau. Peripheral detection as a workload measure in driving: Effects of traffic complexity and route guidance system use in a driving study. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(3):255 – 275, 2005.

[6] D. Kern and B. Pfleging. Supporting interaction through haptic feedback in automotive user interfaces. *interactions*, 20(2):16–21, Mar. 2013.

[7] D. Kern and A. Schmidt. Design space for driver-based automotive user interfaces. In *Proc. AutomotiveUI '09*, pages 3–10, New York, NY, USA, 2009. ACM.

[8] G. Knappe, A. Keinath, and C. Meinecke. Die Sensitivitaet verschiedener Maße zur Fahrzeugquerregelung im Vergleich. In J. Schade and A. Engeln, editors, *Fortschritte der Verkehrspsychologie*, pages 237–256. VS Verlag fuer Sozialwissenschaften, 2008.

[9] B. Mehler, B. Reimer, J. F. Coughlin, and J. A. Dusek. The impact of incremental increases in cognitive workload on physiological arousal and performance in young adult drivers. *Transportation Research Record: Journal of the Transportation Research Board*, (2138):6–12, 2009.

[10] B. Mehler, B. Reimer, and M. Zec. Defining workload in the context of driver state detection and hmi evaluation. In *Proc. AutomotiveUI '12*, pages 187–191, New York, NY, USA, 2012. ACM.

[11] N. Meshkati. Heart rate variability and mental workload assessment. In P. A. Hancock and N. Meshkati, editors, *Human Mental Workload*, volume 52 of *Advances in Psychology*, pages 101 – 115. North-Holland, 1988.

[12] R. M. Michaels. Effect of expressway design on driver tension responses. *Highway Research Board Bulletin*, 1962.

[13] B. Mittelmann and H. G. Wolff. Affective states and skin temperature: Experimental study of subjects with "cold hands" and raynaud's syndrome. *Psychosomatic Medicine*, 1(2):271–292, 1939.

[14] M. Myrtek, D. Weber, G. Brügner, and W. Müller. Occupational stress and strain of female students: results of physiological, behavioral, and psychological monitoring. *Biological Psychology*, 42(3):379 – 391, 1996. Psychophysiology of Workload.

[15] C. K. Or and V. G. Duffy. Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement. *Occupational Ergonomics*, 7(2):83–94, 2007.

[16] A. Pauzie. A method to assess the driver mental workload : The driving activity load index (DALI). *Humanist*, 2(April):315–322, 2008.

[17] B. Pfleging, S. Schneegass, and A. Schmidt. Exploring user expectations for context and road video sharing while calling and drivings. In *Proc. Automotive UI '13*, New York, NY, USA, 2013. ACM.

[18] M.-Z. Poh, D. McDuff, and R. Picard. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering*, 58(1):7–11, 2011.

[19] G. B. Reid, F. T. Eggemeier, and C. A. Shingledecker. Subjective workload assessment technique. Technical report, DTIC Document, 1982.

[20] A. Riener, A. Ferscha, and M. Aly. Heart on the road: Hrv analysis for monitoring a driver's affective state. In *Proc. AutomotiveUI '09*, pages 99–106, New York, NY, USA, 2009. ACM.

[21] A. H. ROSCOE. Heart rate as a psychophysiological measure for in-flight workload assessment. *Ergonomics*, 36(9):1055–1062, 1993. PMID: 8404834.

[22] I. Totzke, N. Rauch, E. Ufer, H.-P. Krüger, and S. Rothe. Workload-Management im Verkehr: Prädiktion von Fahrerbeanspruchung durch Informationen in digitalen Karten. In J. Schade and A. Engeln, editors, *Fortschritte der Verkehrspsychologie*, pages 159–182. VS Verlag fuer Sozialwissenschaften, 2008.

[23] D. d. Waard. *The measurement of drivers' mental workload*. PhD thesis, University of Groningen, 2005.

[24] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, 31(4), 2012.

[25] J. Ziegler, T. Hussein, D. Münter, J. Hofmann, and T. Linder. Generating route instructions with varying levels of detail. In *Proc. AutomotiveUI '11*, pages 31–38, New York, NY, USA, 2011. ACM.

[26] F. R. H. Zijlstra. *Efficiency in work behaviour. A design approach for modern tools*. PhD thesis, Technische Universiteit Delft, 1993.