Towards Audio-based Distraction Estimation in the Car

Svenja Borchers, Denis Martin, Sarah Mieskes, Stefan Rieger, Cristóbal Curio, Victor Fäßler
TWT GmbH Science & Innovation
Stuttgart, Germany
email: Svenja.Borchers@twt-gmbh.de, Denis.Martin@twt-gmbh.de, Sarah.Mieskes@twt-gmbh.de, Stefan.Rieger@twt-gmbh.de, Cristobal.Curio@reutlingen-university.de, Victor.Faessler@twt-gmbh.de

***Abstract*— Distraction of the driver is one of the most frequent causes for car accidents. We aim for a computational cognitive model predicting the driver's degree of distraction during driving while performing a secondary task, such as talking with co-passengers. The secondary task might cognitively involve the driver to differing degrees depending on the topic of the conversation or the number of co-passengers. In order to detect these subtle differences in everyday driving situations, we aim to analyse in-car audio signals and combine this information with head pose and face tracking information. In the first step, we will assess driving, video and audio parameters reliably predicting cognitive distraction of the driver. These parameters will be used to train the cognitive model in estimating the degree of the driver's distraction. In the second step, we will train and test the cognitive model during conversations of the driver with co-passengers during active driving. This paper describes the work in progress of our first experiment with preliminary results concerning driving parameters corresponding to the driver's degree of distraction. In addition, the technical implementation of our experiment combining driving, video and audio data and first methodological results concerning the auditory analysis will be presented. The overall aim for the application of the cognitive distraction model is the development of a mobile user profile computing the individual distraction degree and being applicable also to other systems.**

Keywords-distraction; auditory; automotive; driver; cognitive model.

I. INTRODUCTION

Distraction during driving leads to a delay in recognition of information that is necessary to safely perform the driving task [1]. Thus, distraction is one of the most frequent causes for car accidents [2][3]. Four different forms of distraction are distinguished, although not mutually exclusive: visual, auditory, bio-mechanical (physical), and cognitive. Human attention is selective and not all sensory information is processed (consciously). When people perform two complex tasks simultaneously, such as driving and having a demanding conversation, there is an attention shift. This kind of attention shifting might also occur unconsciously. Driving performance can thus be impaired when filtered information is not encoded into working memory and thus critical warnings and safety hazards can be missed [4]. Sources for distraction of the driver can be located within and outside of the car. The continuous identification of the driver's degree of distraction could enhance safety by allowing adaptive and cooperative task automation using, e.g., advanced driver assistance systems.

Here, we will focus on in-vehicle information. This includes, but it is not limited to, in-car audio recordings and behavioural data from the driver. Multimodal data integration and synchronization is mandatory for the tool to produce meaningful results. Acoustic scene analysis comprising the detection of the number of speakers, the degree of emotional content, information about the driver's involvement in the conversation (e.g., whether the driver himself is speaking), is to be employed for the prediction of the driver's degree of distraction. In addition, eye-tracking signals, such as eye gaze direction and blink frequency, and face movement information, such as mouth movements and emotional reactions, can be exploited to increase the reliability of distraction prediction. A computational and empirical cognitive distraction model is developed for analysing the different signals, with the aim of computing a 'distraction degree' of the driver.

The effect of cognitive distraction on driving performance is empirically tested in a parallel task in order to assess the impact of auditory stimuli on distraction (cf. Figure 1). In a first experiment, we induce a continuous distraction condition and compare the driving parameters and in-car measurements with a control condition of focused, un-distracted driving. Analysing these results, we assess parameters responding reliably to cognitive distraction. These parameters are used as input for the cognitive model computing the degree of the driver's distraction. In a second experiment, we then induce a more naturalistic conversation condition leading to varying degrees of driver distraction. Our computational and empirical cognitive model is trained and tested in the course of this experiment. An acoustic analysis including the detection of the number of speakers, the degree of emotional content, information about the driver's involvement in the conversation (e.g., whether the

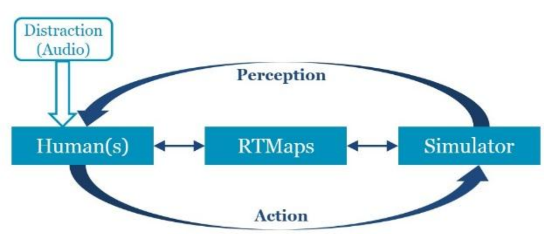

Figure 1. Perception-action loop and the influence of distraction.

driver himself is speaking), is used for the prediction of the driver's degree of distraction.

In Section II, the experiment is described addressing the experimental design and the features being analysed. Section III presents preliminary results of the first experiment. Section IV gives an outlook of the following analysis steps and further experiments.

II. EXPERIMENT

A. Experimental Design

In our first simulator experiment, we used a car following paradigm with the driver's task to keep the same distance to the pace car by ensuring readability of the number on the back end of the pace car. Subjects performed a practice session of three minutes driving without distraction in order to get used to the experiment and the driving simulator. The pace car drives with varying speeds between 30 and 100 km/h and brakes or accelerates 39 times during a 10 minute drive at randomly distributed locations. Some of the subjects started with the control condition, i.e., driving without distraction, of ten minutes, while other subjects started with the distraction condition. After the first condition, subjects continued with the other condition, so that each subject performed once the control and once the distraction condition. During the distraction condition, subjects were presented with simple mathematical tasks (e.g., 22+46 or 9-5) via headphones and subjects were asked to respond verbally [5][6][7]. The inter-trial interval was chosen to eight seconds. All responses were recorded.

Subjects had normal or corrected-to-normal vision and several years of driving experience each. They sat as driver in front of a large screen using a Logitech G27 game controller steering wheel with pedals (cf. Figure 2). The simulator allowed the driver to control an automatic car with the steering wheel, the gas pedal, and the brake pedal (the clutch pedal was not used). As driving simulation software, OpenDS [10] was used. Besides a custom driving task definition, minor modifications of the simulator were necessary to show brake lights of the pace car and to remotely control a software for recording videos from two web-cameras. The cameras were used to record the subject's face. One of the cameras was positioned directly in front of the subject and the other to the side front.

Figure 2. Driving simulator setup.

Synchronisation of the camera streams was guaranteed by RTMaps [9], which was remotely controlled by the OpenDS driving simulator. Facial features, mouth movements and head pose of the subject are automatically extracted to increase the reliability of distraction predictions in further analyses.

Simulator sound and the audio task were presented through a headset. Its microphone was used to record the verbal responses during the distracted condition.

B. Features

Parameters being indicators for driving performance are extracted from the driving experiments. These parameters include: distance to pace car, reaction times (both for braking and speed recovery), steering wheel jitter, and lateral position jitter. Further parameters will be evaluated for their potential use as features in the cognitive model and will be included step-by-step, e.g., head orientation (which will be relevant in conversation tasks), eye blink, and facial expressions (for emotion recognition). For conversation tasks, audio analysis will be included in the feature set of the cognitive model. In this context, features used in voice and speech recognition, such as pitch and Mel-Frequency Cepstral Coefficients (MFCC) are suitable candidates as well as derived features, such as emotional content of the utterances.

Since features used for our cognitive model will eventually come from different sources (car data, video, audio), synchronisation plays an important role. One tool allowing acquisition of multi-modal sensory data is RTMaps [9], which will be used as platform for implementing our auditory driver distraction estimation component.

III. PRELIMINARY RESULTS

Preliminary results of the driving parameters of six subjects are shown in Figure 3. All subjects showed a tendency of a larger mean distance to the pace car during distracted driving despite the explicit assignment of keeping a predefined distance determined by the readability of large numbers on the rear of the pace car. In addition, a larger variance of the distance to the pace car indicates longer reaction times for adapting to the speed of the pace car. Thus, subjects were less constantly able to keep the same distance to the pace car while they were simultaneously solving mathematical problems.

The longer reaction times are especially reflected in the deceleration case (braking instances of pace car): All subjects needed longer reaction times between the occurrence of the breaking lights of the pace car and decelerating of their own car while distracted. The variance of acceleration (including deceleration) indicates a smoother driving behavior during distracted driving (smaller spikes of the acceleration value). Together with the longer reaction times and the increased distance to the pace car, this generally shows that the subjects are more likely to drive in a safer style when cognitive workload is increased.

In conclusion, these driving parameters indicate the effectiveness of the induced distraction through the mathematical problem solving task. During our upcoming experiments, we will use these parameters to evaluate the contribution of specific auditory and facial features.

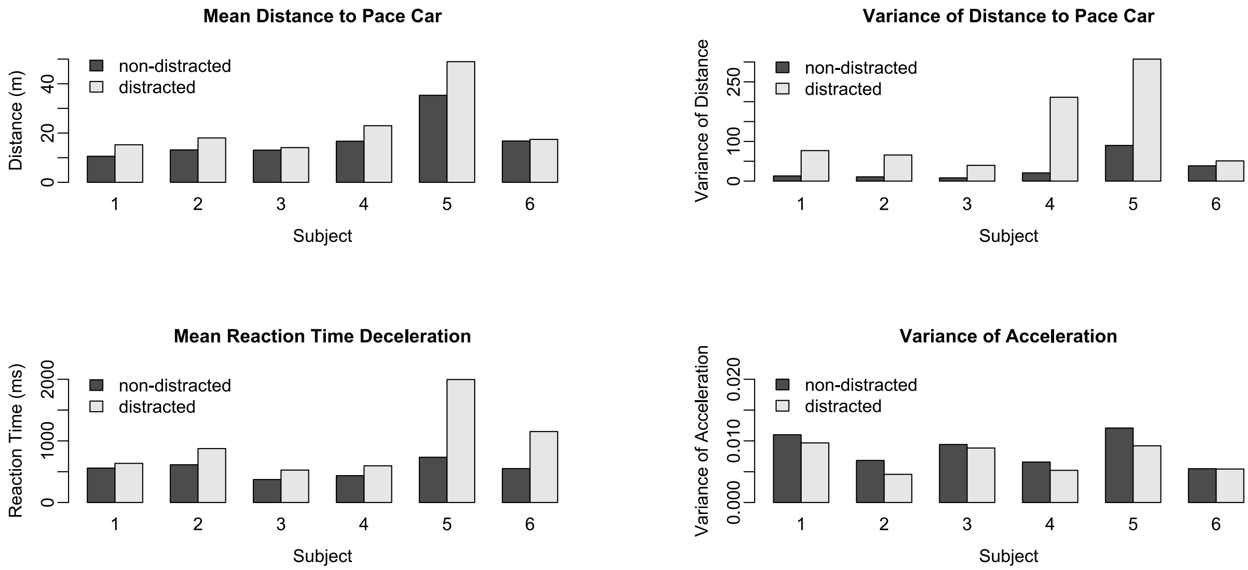

Figure 3. Driving parameters of six subjects: mean distance of the driver to the pace car, variance of distance to the pace car, mean reaction time of deceleration, and variance of acceleration.

IV. CONCLUSION AND FUTURE WORK

We have designed an experimental driving simulator setup that enables the study of behavioural and perceptual manipulations during driving, as shown by first promising quantitative results. In future work, a special focus will be on the audio scene analysis in the car interior. For this, we will extend the experimental paradigm. This will involve the driver under controlled conditions during a conversation, while monitoring her/his emotional states (i.e., through audio and facial signature analysis). First technology studies suggest that auditory features, such as pitch and MFCC are suitable candidates. An analysis of mouth movements will add information to the audio segmentation helping to identify active speakers. As platform for synchronized processing of the different data sources (audio, video, and driving performance parameters from the car), RTMaps [9] will be used (cf. Figure 4).

Several models for cue integration have been suggested for cognitive modelling of distraction. The recent dynamic Bayesian model by Liang and Lee [8] consists of a combined supervised and unsupervised learning approach. It would be interesting to extend this model with higher-level conversational cues, like the degree of estimated conversational interaction as a likely distraction measure.

Besides considering further multimodal observational cues of car passengers, especially the driver, the system should be tested and calibrated in more complex driving situations, like overtaking. Modelling current driver's task difficulty through an artificial driving model will be a further interesting research direction.

Finally, investigating strategies to support the driver by presenting and using the estimated distraction level, e.g., through visual feedback modalities or active interventions, will also be of further interest. The design of future autonomous driving systems will call for functionalities to access driver states at various automation levels.

The final technical implementation of the developed cognitive model as a mobile user profile will be further investigated throughout the project. It is likely that adaptation and life-long learning of the cognitive model will be a key feature for which a mobile application communicating with on-board car systems would be an appropriate choice.

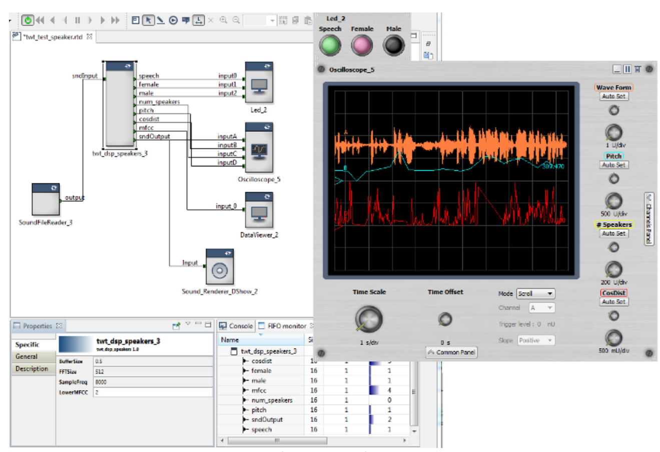

Figure 4. RTMaps for audio feature extraction.

ACKNOWLEDGMENT

This research has been performed with support from the EU ARTEMIS JU project HoliDes (http://www.holides.eu/) SP-8, GA No.: 332933. Any contents herein reflect only the authors' views. The ARTEMIS JU is not liable for any use that may be made of the information contained herein.

REFERENCES

[1] M.A. Regan and K. L. Young, “Driver distraction: a review of the literature and recommendations for countermeasure development,” Proc. Australas. Road Safety Res. Policing Educ. Conf. 7 (v1), pp. 220–227, 2003.

[2] J. Artho, S. Schneider, and C. Boss, “Inattention and distractioin: how does the driver behave in the car?”, original title: Unaufmerksamkeit und Ablenkung: Was macht der Mensch am Steuer?,” Transport Research International Documentation, 2012, Online: http://trid.trb.org/view.aspx?id=1244037. Retrieved: October, 2014.

[3] T. Horberry, J. Anderson, M. A. Regan, T. J. Triggs, and J. Brown, “Driver distraction: the effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance,” Accid Anal Prev 38 (1), pp. 185–191, 2006.

[4] L. M. Trick, J. T. Enns, J. Mills, and J. Vavrik, “Paying attention behind the wheel: a framework for studying the role of attention in driving,” Theoretical Issues in Ergonomics Science 5 (5), pp. 385–424, 2004.

[5] M. Kutila, G. Jokela, G. Markkula, and M. Rue, “Driver distraction detection with a camera vision system,” IEEE International Conference on Image Processing (ICIP 2007), vol. 6, San Antonio, Texas, USA, pp. 201-204, 2007.

[6] F. Putze, J.-P. Jarvis, and T. Schultz, “Multimodal recognition of cognitive workload for multitasking in the car,” International Conference on Pattern Recognition (ICPR 2010), Istanbul, Turkey, August 2010, pp. 3748-3751.

[7] J. Harbluk, Y. Noy, P. Trbovich, and M. Eizenmann, „An on-road assessment of cognitive distraction: Impacts on drivers’ visual behaviour and braking performance,” Accident Analysis and Prevention, vol. 39, no.2, pp. 372-379, 2007.

[8] OpenDS, EU FP7 Project GetHomeSafe, FP7-ICT-2011-7, 288667 STREP. Online: http://opends.eu/. Retrieved: October, 2014.

[9] RTMaps, Real-Time Multimodal Applications, Intempora S.A. Online: http://intempora.com/. Retrieved: October, 2014.

[10] Y. Liang and J. D. Lee, “A hybrid Bayesian network apporach to detect driver cognitive distraction,” Transportation Research Part C 38, pp. 146-155, 2014.