Modelling driver styles based on driving data

Peter Mörtl, Andreas Festl, Peter Wimmer, Christian Kaiser, & Alexander Stocker
VIRTUAL VEHICLE Research Center
Austria

Abstract

Driving styles are habitual ways of driving that are characteristic for groups of drivers and represent an important topic of research for advanced automation in the vehicle of the future. Relatively little knowledge exists concerning the connection between driving styles and the underlying cognitive-psychological aspects of the driver. To better understand this connection, we investigate driving style indicators in a driving simulator and create a cognitive model of the underlying cognitive-psychological processes that we compare with the empirical data. The cognitive model produces steering behaviour that approximates the lateral deviations of human drivers while also producing similar rates of steering wheel direction reversals. These results confirm the utility of this approach for representing individual driving styles and states for advanced vehicle automation.

Introduction

Research in driving styles has a long tradition and recently experienced a new focus of interest because modern automotive technologies have become smart enough for personalised in-vehicle interventions. The term driving style has been defined by (Sagberg, Selpi, Piccinini, & Engström, 2015) as a “habitual way of driving, which is characteristic for a driver or a group of drivers”. Driving styles also stay constant for a given driver across different driving contexts. Global driving styles combine multiple driving indicators (such as aggressive, calm, or careful driving) and specific driving styles are measured by one or two indicators.

While there are many driving metrics, there are currently few models that tie them to their underlying psychological processes. Such models would be important because of several reasons. First, they may allow tying multi-sensory observational data streams together to inform driver state inferences. Second, such integrative models could be used in virtual safety assessments to represent the human driver. Such virtual safety assessments are needed to determine whether higher levels of automated driving are safe because repeated extensive real-world studies would take too long and be too cost-intensive. And third, such models could be used to better personalise automated driving so that the self-driving car has knowledge of the human driver and adjust to his or her style. Also, personalised in-vehicle support could detect a need for intervention based on observing the driving behaviour for fatigue or distraction. The detection of such driver states depends on representing the driver’s individual driving style which is the focus of this paper.

In this paper we first briefly describe some driving style indicators. Then we describe an empirical study to collect driving data in a driving simulator and investigate four driving style metrics: steering, speed, acceleration, and distance keeping. We then describe the development of a psychological driving process model to predict lateral lane deviations and compare them to human participants. We validate the model for one of the four driving metrics, the steering metric, and describe the results.

Specific driving styles

Steering behaviour represents a thoroughly investigated driving style metric. Li et al. (2017) use measures of entropy and variation in timely sequences of steering wheel angles to estimate driver drowsiness. Fairclough & Graham (1999) measured steering wheel reversal (SWR) differences between a control group and partially and fully sleep-deprived drivers in a simulator study. They find a reduction of SWR from about 15 to 11 per minute for the sleep-deprived group compared to the control group (see also McLean & Hoffmann, 1975). Thiffault & Bergeron (2003) examined the influence of fatigue on various steering measures including the mean steering angle amplitude, frequency of larger steering wheel movements, and their standard deviation. Otmani et al. (2005) investigated the influence of fatigue on the mean steer wheel angle changes and found that they amount to about between 0.5 and 5 degrees, averaged over 1 to 10 min driving periods. Similarly, Yan, Radwan, & Guo (2007) and Ungoren & Peng (2005) report that individual drivers differ in their steering behaviour.

Another well investigated specific driving style metric are driving speed and accelerations. Ericsson (2000) reports on traffic dependent as well as individual differences in driving speed, for example between males and females (see also Brundell-Freij & Ericsson, 2005; Ericsson, 2001 who also investigated driving acceleration). Af Wåhlberg (2007) investigated the variability and amount of driving accelerations and found that they could serve as precursors of accidents (see also Af Wåhlberg, 2008; af Wåhlberg & Dorn, 2007). Bagdadi & Várhelyi (2011) investigated jerky driving as predictor of accidents, see also (Murphey, Milton, & Kiliaris, 2009). Desai & Haque (2006) investigated pressure on the acceleration pedal as individual parameters for driver alertness.

Another field of specific driving style metrics is the headway distance to the vehicle ahead (see, e.g. Shinar & Schechtman, 2002). Taieb-Maimon (2007) and Taieb-Maimon & Shinar (2001) investigated the impact of training on improving inter-vehicular distance.

Data Collection

We collected driving data in a non-motion based driving simulator. After an initial warm-up, 16 participants drove a car with automatic transmission on a curvy road of 10 km length. The road segment was extracted from satellite imagery recreated in the simulator and eight metres wide. Participants were between 20 to 60 years old, 12 were male and four female. All had driver licences and drove between 5,000 and 20,000 km per year. Participants completed two scenarios on two different roads.
In the first scenario (“driving and passing scenario”), they could select their driving speed and encountered oncoming traffic approximately every 25 seconds. They encountered slower vehicles that they could pass if desired. Participants were encouraged to drive as close as possible to how they drove in the real world. When passing a vehicle they were asked to indicate their intent to pass by actuating a lever on the left side of the steering wheel. In a second scenario (“steering only scenario”), participants only steered their vehicle that drove at a constant speed of 90 km/h. There were no opportunities to pass other vehicles. The order of the scenarios was randomized and counterbalanced so that 50% of the participants experienced scenario 1 before scenario 2 and vice versa to avoid order effects.

Before the simulation started, participants completed a questionnaire (see Table 1) that assessed some aspects of their general driving style. After completion of both driving scenarios they completed a short questionnaire about their driving. The response scale to all questions was a 7-item Likert scale.

Table 1. Questionnaire items

<table>
<thead>
<tr>
<th>General Driving Style</th>
<th>Scenario Specific Driving</th>
</tr>
</thead>
<tbody>
<tr>
<td>I frequently pass cars.</td>
<td>I frequently passed cars.</td>
</tr>
<tr>
<td>I usually do not have to brake prior to curves.</td>
<td>I usually did not have to brake prior to curves.</td>
</tr>
<tr>
<td>I sometimes cross red lights.</td>
<td>I think that I drove safely.</td>
</tr>
<tr>
<td>People say that I am a safe driver.</td>
<td>I think that I drove very carefully.</td>
</tr>
<tr>
<td>I usually drive very carefully.</td>
<td>I drove as fast as possible.</td>
</tr>
<tr>
<td>I sometimes drive as fast as possible.</td>
<td>I was braking hard at least once.</td>
</tr>
<tr>
<td>I never “chase” yellow lights.</td>
<td>I drove “sportily”.</td>
</tr>
<tr>
<td>I often have to brake.</td>
<td>When driving, cars often pass me.</td>
</tr>
<tr>
<td>People say that I am a “sporty” driver.</td>
<td></td>
</tr>
</tbody>
</table>

Driving style characterisation

We first present the results of the observed driving style metrics for scenario 1 where participants could freely choose their speed. Driving data during periods of free driving (i.e. without a vehicle ahead) were separately analysed from periods when they had to adjust their speed because of a vehicle ahead. The road of 10,000 m was divided into 100 segments that were each 100 m long. The first segment was removed because all drivers accelerated the vehicle. Each of the remaining segments was classified for each participant either as following another car (if it came within 50 m of the vehicle ahead), as driving freely, or overtaking a car. Only free and following driving segments were considered further in the analysis. Following driving metrics were investigated:
1. Count of the steering wheel direction reversals per second over “free” and “following” segments.
2. Mean vehicle speed in the “free” driving segments.
3. Mean following distance in the “following” driving segments.
4. Count of accelerations in all “free” and “following” segments.

Figure 1 shows the identified variability of the four driving metrics that differed considerably between them.

Several of the metrics correlate with each other: The faster drivers accelerated more frequently, however only when driving freely (i.e. $r=0.23$, $p < 0.001$ when driving freely but only $r=0.09$, $p > 0.1$ when following). Drivers who drove faster also more frequently reversed the steering wheel direction ($r=0.21$, $p < 0.001$) regardless whether they followed or drove freely. When following a vehicle, people who kept more distance to the vehicle tended to reverse the steering wheel direction more frequently ($r = 0.18$, $p < 0.001$) and tended to drive faster ($r=0.29$, $p < 0.001$). No other correlation among the driving metrics reached significance. Also, we found no correlation between driving metrics and the participants’ responses on the questionnaires.
To determine to what extent these variations reflect just random noise or differed statistically between drivers and road segments, we utilised a two-way random effects model (Searle, Casella, & McCulloch, 1992), which is a special case of the linear mixed effects model:

\[
Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij},
\]


where the indices \(i\) and \(j\) indicate the drivers and road segments and the variables are defined in the following way:

- \(Y_{ij}\) is the response variable, i.e. the specific driving metric.
- \(\mu\) is the overall mean of the driving metric.
- \(\alpha_i\) is the random effect of the driver which is normally distributed around 0 with a variance of \(\sigma^2_\alpha\) i.e. \(\alpha_i \sim \mathcal{N}(0, \sigma^2_\alpha)\).
- \(\beta_j\) is the random effect of the road segment which is normally distributed around 0 with a variance of \(\sigma^2_\beta\) i.e. \(\beta_j \sim \mathcal{N}(0, \sigma^2_\beta)\).
- \(\epsilon_{ij}\) is the random noise that cannot be attributed to the driver or road section. The variation of possible other factors of influence are also captured in this term.

No interaction \((\alpha\beta)_{ij}\) is included because each driver drove each road segment only once. That means we had to assume that the drivers responded equally to each road segment.

The results of this analysis are shown in Table 2. The variances of all random effects are highly significant, indicating statistically significant differences in the driving metrics among the drivers and the road segments. Knowing the driver and the road segment reduces the overall random noise as indicated Table 2. For example, the
variance for the driving metric “median distance to the next vehicle” reduces from 15.06 m to 6.6 m when subtracting driver and road segment effects.

Table 2. Driving style metric variability

<table>
<thead>
<tr>
<th></th>
<th>Median Distance to Next Vehicle (m)</th>
<th>Mean Vehicle Speed / sec</th>
<th>Steering Wheel Reversals / sec</th>
<th>Acceleration Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver</td>
<td>5.5510*</td>
<td>2.4920*</td>
<td>0.2033*</td>
<td>0.3270*</td>
</tr>
<tr>
<td>Road Section</td>
<td>2.8630**</td>
<td>2.4700*</td>
<td>0.2126*</td>
<td>0.4254*</td>
</tr>
<tr>
<td>Residual</td>
<td>6.6450</td>
<td>2.7960</td>
<td>0.5173</td>
<td>1.3250</td>
</tr>
<tr>
<td>Total</td>
<td>15.0590</td>
<td>7.7580</td>
<td>0.9332</td>
<td>2.0774</td>
</tr>
</tbody>
</table>

* p < 0.001; ** p < 0.01

In Figure 3 the proportions of the explained variances are compared with each other. We find that the greatest amount of overall variation is explained by the driver’s median distance to the next vehicle (36.86%). This is somewhat expected because distance keeping should more or less reflect a conscious decision by the driver. Road characteristics also contribute significant variability though the driver influence is almost double of the road segments. Road and drivers explain about equal variances of the mean vehicle speed, similar is the case for SWRs. In terms of number of accelerations, knowing the road segment has a bigger impact on reducing the overall variance than knowing the driver.

Figure 3. Relative contribution of the variances within each driving metric
Cognitive Modelling Architecture

There are many cognitive modelling approaches that have been applied to driving (e.g. Anderson et al., 2004; Bubb, Bengler, Grünen, & Vollrath, 2015; Deml, Neumann, Müller, & Wünsche, 2008; Kieras & Meyer, 1997; Laird, Newell, & Rosenbloom, 1987; Lewandowsky & Farrell, 2011; Liu, Feyen, & Tsimhoni, 2006; Salvucci, 2006) and it is beyond the scope of this paper to provide an overview. Primarily we were searching for a general purpose, modular architecture that would help to represent human psychological processes for engineering tasks. For this we came to utilise a cognitive modelling architecture that we describe in more detail in Moertl, Wimmer, & Rudigier (2017). In this approach we adopted the basic elements of the human cognitive architecture by Card, Moran, & Newell (1986), the Model Human Processor (MHP) and adapted it to specific driving tasks, see figure 4.

![Figure 4. Cognitive modelling architecture.](image)

In the context of the driving task, the MHP cognitive processor receives information from the perceptual processor about the external world and executes motoric tasks to control the vehicle. Each process takes a certain execution time and, dependent on serial or parallel processing, determines the overall task duration and timing of interactions. The central cognitive task is to determine the next driving action. In the model that we utilised in this study all tasks were executed serially but could also be executed in parallel. The transfer of information into memory and from memory involves buffers. However, in the simple model that we utilise in this study all the information was visually available in the environment and therefore did not require the use of explicit buffers.

**Steering Modelling Method**

With this model we implemented Salvucchi’s steering model (see Salvucci, 2006; Salvucci & Gray, 2004) where steering is a direct result of perceptual fixations of certain points on the road ahead and the amount of time between subsequent environmental scans. The amount of time between scans is directly proportional to the number of steering corrections and inversely proportional to their size. Therefore, the more time is dedicated to steering and perceiving, the higher the number of control actions and the smaller their size. On the other hand, if only
limited time is available for perceiving and steering such as when multitasking and searching for signs or talking on the cell phone, the less frequent should be the steering control actions and the larger their size.

The model that we implement based on Salvucchi’s work uses perceptual information that is available to human drivers when driving. The model has received empirical validation from visual occlusion experiments where human drivers drove in a driving simulator while some areas of the visual scenery were obscured (Land & Horwood, 1995; Land & Lee, 1994). This steering model utilises both a far point and a near point that are both ahead of the driver’s own vehicle: the near point is a constant distance ahead and is located in the middle of the driving lane. The far point is further ahead and consists alternatively of the tangent point of an upcoming curve, the vanishing line of a straight road segment, or a vehicle that is driving ahead. The far point is intended to allow steering the vehicle into and out of curves whereas the near point helps to centre the vehicle on the driven lane.

We implemented this steering model in our cognitive architecture by only considering three cognitive processes that are executed in turn: a perception, a cognitive, and a motoric process. Each process was assumed to take 50 ms see e.g. (Card et al., 1986), so that one full cycle of steering updates takes 150 ms. We also updated the main parameters of the model: Whereas Salvucci (2006) suggested the three model parameters as $k_{far} = 16$, $k_{near} = 4.0$, and $k_l = 3.0$ we adjusted them to $k_{far} = 1.6$, $k_{near} = 0.4$, and $k_l = 0.09$ to achieve better performance.

Results

Comparison lateral performance

We compared the steering quality of human participants with the psychological driver model in scenario 2. The red solid line in Figure 5 shows the lateral deviations of the model versus the 90 percentile of human steering over the whole 10 km. The model was 79.4 % of the time within the human driving boundaries and correlated on average with $r = 0.36$ with the human drivers.

![Figure 5. Comparison of lateral deviation between humans and psychological steering model.](image-url)
Comparison steering wheel reversals

Given that the psychological model steered similar to human drivers, we compared the number of SWRs between humans and the model. It is important to note that we did not specifically attempt to fit SWR: we only fitted the model to the lateral lane deviations by adjusting the above mentioned three model parameters. Figure 6 shows the results for both scenarios 1 and 2. The bars indicate the mean number of steering wheel reversals per minute for our 16 participants. The blue line indicates the steering wheel reversals of our psychological driving models. On the left, drivers controlled speed and were passing other cars (scenario 1) whereas in the right figure, they only steered the vehicle (scenario 2). The roads were different between scenarios 1 and 2.

The results indicate that the psychological model produced very similar amounts of SWR to our human participants in both scenarios. All their counts fell within one standard deviation of the human participants. The model was similarly accurate in its steering wheel reversals in both scenarios as it produced, similar to human drivers, relatively more SWRs on scenario 1 (84.3 SWR/min for the model versus 85.7 SWR/min for the human) than on scenario 2 (71.0 SWR/min for the model versus 77.0 SWR/min for the human drivers). This indicates that by modelling the underlying cognitive processes the model was able to match human performance on two different dimensions despite having been fit only on lateral lane deviations.

Conclusions

The results of our study indicate that knowledge of the individual driver and road segment could explain up to 64% of the overall variability of the investigated driving metrics. The remainder of the variance apparently represents random fluctuations that would exceed driver modelling. This knowledge helps bound the expectations of how well driving models can approach real human driving.

After review of relevant literature we derived a relatively simple and modular cognitive modelling architecture in which we implemented as first step a psychologically plausible steering algorithm by Salvucci (2006). With some
moderate amount of fitting that basically consisted of adjusting the model’s three critical steering parameters, the model not only resembled the lateral driving deviations of our 16 human participants but also approximated their rate of steering wheel direction reversals. This demonstrates the principal benefit of models that not only represent outcomes but underlying structure for the applications in the vehicle of the future: new and valid behavioural predictions can be derived from the model structure rather than having to base each prediction on an extensive learning process of stimulus-response. Such power of generalization is essentially missing in pure machine learning algorithms but seems crucial to better fit the contextual needs of the driver and help allow for inferential processes to ascertain whether a system is safe.

Much remains to be done to establish psychological driver modelling as a standard tool for human-centred automotive developments. First we will need to confirm that our psychological models are not only valid for simulation studies but also for real world driving. Then we need to test how the psychological model can be adapted to capture individual driver styles and states, such as, for example, driving distraction. Finally, we will extend our modelling to other driving aspects, specifically braking, distance keeping, and speed selections.

Acknowledgement

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 732189. The paper expresses the views of the authors.

References

modeling driving style


