Self-Assessing Visual Attention While Driving: Implications for Driver Monitoring Systems

Max Harding[1], Catherine Harvey[1], David R. Large[1], Gary Burnett[2], Chrisminder Hare[3] & Karl Proctor[3]

[1]University of Nottingham, [2]Loughborough University, [3]Jaguar Land Rover

SUMMARY

Driver Monitoring Systems (DMS) are now a requirement in all new vehicles. DMS aim to reduce crashes and improve driver attention by providing warnings or interventions to the driver. However, driver acceptance is crucial to ensuring their effectiveness. Various methods exist to detect inattention, but if these do not align with a driver's mental model, acceptance issues may arise regarding the warnings provided. The study outlined in this paper examines how drivers assess their own visual attention by comparing self-reported ratings to a visual attention algorithm. Using an on-road experiment with six participants, initial results highlight the importance of contextual information for accurately assessing visual attention and providing effective warning strategies.

KEYWORDS

Driver Monitoring, Driver Attention, Visual Inattention, Driver Acceptance

Introduction

Systems monitoring driver attention are a requirement for all new vehicles via the General Safety Regulations (GSR) of the European Union (European Commission, 2023). Similarly, Euro NCAP details Driver Monitoring Systems (DMS) requirements and guidelines for using DMS with driving assistance features (Euro NCAP, 2023, 2024). The key aim of DMS is to reduce the number of crashes related to visual inattention by providing warnings to redirect a driver's attention back to the driving task, or through other interventions, such as breaking or disabling driver assistance features. It is a GSR requirement that the driver can deactivate these systems. Driver acceptance could be affected if the system is intrusive or prone to false alarms, so it should only warn drivers when genuine inattention is detected (Smyth *et al.*, 2021). There are various methods for estimating driver attention with strong anchors in driver attention theory (Ahlström, Georgoulas and Kircher, 2022). However, attempting to define a ground truth to compare visual attention algorithms against is a difficult task. Nevertheless, the correspondence between DMS outputs and a driver's perception of their attention will likely impact acceptance.

The study looks at a novel on-road method comparing how drivers self-assess their visual attention against the output of the visual attention algorithm AttenD (Kircher and Ahlström, 2013). The output of the AttenD algorithm is a buffer value between 0 and 2, which changes over time depending on where the driver is looking. The buffer increases in value with glances to the road and decreases with glances away from the road or when the driver is visually distracted. Drivers are given additional time when performing driving related tasks such as checking mirrors or speed, before the buffer begins to decrease. When the value of the AttenD buffer drops to 0, the driver is considered visually inattentive. During the study, we examined cases in on-road driving where ratings of inattention by the AttenD algorithm and the driver differed. These discrepancies helped explore factors that influence a driver's assessment of their attention and examine scenarios where DMS acceptance may be impacted.

Methods

Six participants (3 male, 3 female) were recruited for an on-road study comprising a 20-minute drive. To encourage naturalistic driving behaviour, participants could follow a familiar route or use a satnav for navigation if unfamiliar with the local area. Participants wore SMI eye-tracking glasses and drove in the research institute's Nissan Leaf Test Vehicle. Before the drive, participants were briefed on the definition of visual attention and were given examples of attentive and inattentive visual behaviour while driving. Participants were instructed to drive as they normally would. At 30-second intervals, a short audible tone prompted participants to rate their visual attention in the period since the last prompt. Participants were instructed to verbally respond with "yes" if they felt they had been visually attentive since the last prompt, or "no" if they felt they had been visually inattentive, followed by a short reason if it was safe. Other rating scales, such as Likert scales, were piloted. However, the complexity of the scale made it difficult for drivers to distinguish between visual and cognitive attention when giving ratings.

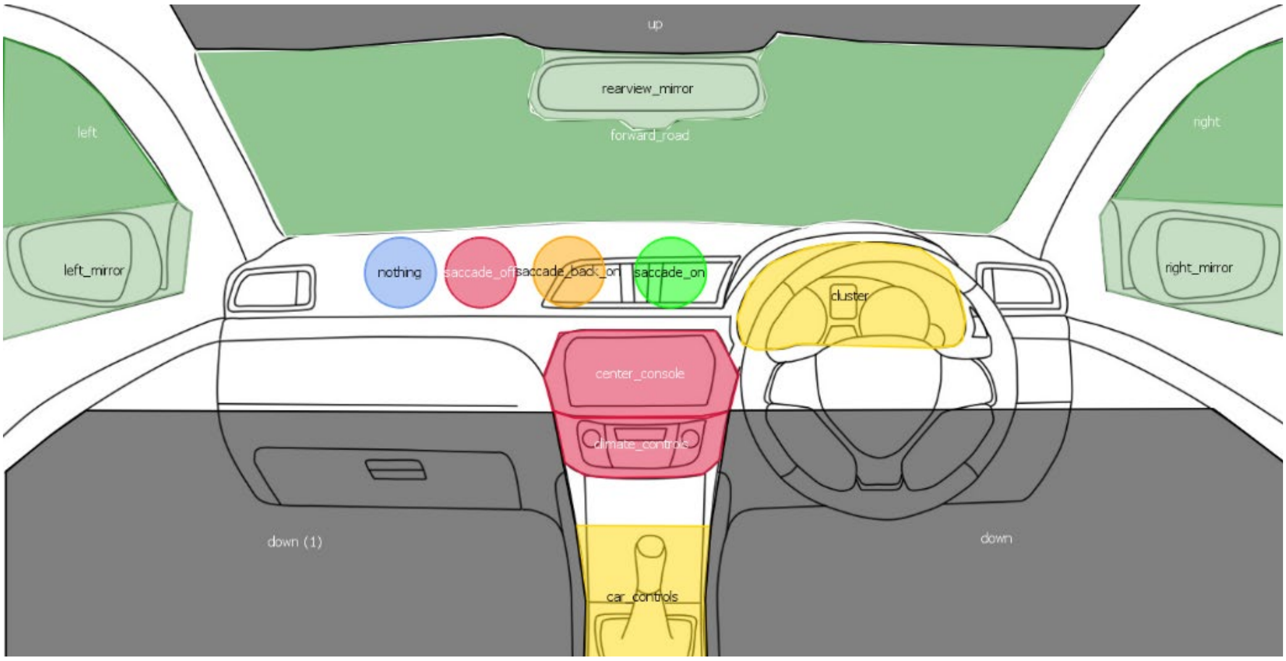

Figure 1: Areas of Interest for semantic gaze mapping

Following each session, the eye tracking data was analysed using semantic gaze mapping to map fixations and saccades to areas of interest defined by AttenD as shown in Figure 1. The AttenD algorithm was retrospectively applied to the gaze data to generate an attention profile across each drive. To compare the participant's self-assessments to the AttenD algorithm, the sections across which drivers were prompted to rate their visual attention, were overlaid on top of the output of the AttenD algorithm. If the AttenD buffer reached zero in any of these sections, then the AttenD rating for that section was marked as inattentive. This simplified output of the AttenD buffer can be easily compared to the driver's binary rating scale ("yes/no").

After the analysis of the gaze data, and processing of the data using AttenD, participants were invited back to review clips of their eye tracking from the drive. An example frame from a clip is shown in Figure 2. Participants were presented with clips where there was either a disagreement between their own rating and the output of the AttenD algorithm, or where both they and the AttenD algorithm had given a rating of inattentive. Participants did not have knowledge of their initial ratings or the output of the AttenD algorithm while watching the clips. Participants were asked to reassess their visual attention based on the clips to explore whether their post-drive responses aligned with their responses during the drive. The time frame between the initial drive and the participant's interviews was kept to a minimum (typically a few days) so that the drive was still familiar to the participant. After the participants reassessed each clip, they were informed of their initial rating and the AttenD algorithm's rating. Participants were then asked whether they would have been receptive to an 'intervention' or prompt at that point in time, alerting them that their visual attention was low in that scenario. Participants were also asked what type of warning they would prefer/expect in that scenario.

Figure 2: Frame from the recording of a participant's eye tracking.

Results and Discussion

Comparison of Participant Ratings

In total, there were 256 ratings of attention across the 6 participants. The results of the study are summarised in Table 1. For the purpose of this discussion, AttenD is taken as the ground truth (i.e. AttenD is a true measure of the participant's visual attention. The data in Table 1 are the signal detection theory comparisons of 'True Positive' or 'Hit', 'True Negative' or 'Correct Rejection', 'False Negative' or 'Miss' and 'False Positive' or 'False Alarm' (Stanton *et al.*, 2009). Table 1 shows that some participant's assessments changed between the drive and post-drive interview. Out of the ratings given during the drive 79% matched the algorithm, this increased to 92% post-drive. Two of the six participants did not change their responses between the drive and the interview. Changes in false positive outcomes were due to participants amending their initial assessment of inattentive to attentive upon post-drive review (note that 70% of these changes are attributed to one participant, who stated their ratings were overly cautious during the drive). For the majority of changes to False Positive ratings, participants could not see a reason for their original rating. For two of the False Positive ratings remaining post-drive, the participant stated they had been staring ahead and daydreaming. As they were looking ahead, the algorithm would not have been able to pick up on this. The other two remaining false positives were when the algorithm got very low but not quite to zero and a case where a driver felt they were still visually inattentive.

Table 1 - Participant Assessment vs AttenD Ratings

Outcomes	Drive	Post-Drive	Description
True Positive	17	24	Both AttenD and Participants Rated Inattentive
True Negative	181	207	Both AttenD and Participants Rated Attentive
False Positive	30	4	AttenD Rated Attentive, Participant Rated Inattentive
False Negative	28	21	AttenD Rated Inattentive, Participant Rated Attentive

Similarly, the post-drive reduction in False Negative outcomes is due to participants changing their attentive ratings to inattentive ratings on review. Of the remaining False Negative ratings after the post-drive interview, 66% of these occurred while participants were stopped at or approaching traffic lights. This suggests that there is a change in what the participants perceived should be classed as inattentive behaviour depending on the scenario. The remaining 34% of disagreements were due to drivers simply disagreeing that they were inattentive while driving. The discrepancy with AttenD and changes in their own assessment upon review highlights drivers' difficulty in assessing their visual attention in the moment, or differences in understanding of what visual attention is. There were also 5 cases due to known issues with the AttenD algorithm, where a false rating of inattentive was given. The false rating occurred when drivers looked away from the forward roadway for extended periods for genuine reasons, such as checking right to see if a roundabout or junction was clear. Without contextual information about where the driver's glances should be prioritised, the algorithm cannot accurately assess the driver's visual attention in all scenarios.

Participants Responses to Warnings

The study also looked at participants' opinions on warnings in response to a DMS detecting visual inattention. The False Negatives due to known issues with AttenD, were excluded from the analysis. Naturally, participants said they would not be open to warnings for the False Negative outcomes where they had rated themselves as visually attentive. Within the True Positive outcomes after the interview, participants said they would not be receptive to warnings in 30% of cases. The times when participants said they would not be receptive to interventions within the True Positive Outcomes, were similar to the cases they had disagreed with AttenD in that they mainly occurred when the participants were stopped or stopping.

Overall, participants said they would not be receptive to warnings in 57.5% of the cases where AttenD rated inattentive (irrespective of how the participant rated). These responses demonstrate the importance of the output of a driver attention algorithm aligning with the drivers' mental models for acceptance of DMS. Looking closely at the clips where participants were not receptive to warnings, 57% occurred while driving, and 43% occurred when participants were stopped or approaching traffic lights. The clips where drivers stopped at traffic lights point to a difference in what the participant and AttenD consider inattention. These cases showcase that contextual information is important for accurately detecting inattention (Kircher and Ahlström, 2024) and acceptance of warnings. Participants also said they would be more accepting of specific warnings in different scenarios and would want the modality or severity of the warning to change depending on different factors. The results highlight the need for contextual information to be accounted for future DMSs in determining driver attention and the modality that warnings about their attention are given.

Conclusion

The study explored the differences between driver self-assessments of visual attention and the AttenD algorithm's outputs. The findings of the study highlight the importance of contextual information for accurately assessing visual attention and raise questions about how visual attention should be assessed in different scenarios. Participants often revised their self-assessments on review post-drive, revealing difficulties in accurately assessing their visual attention in real time, which could further contribute to acceptance issues. The study also highlighted that participants said they would be less receptive to warnings in scenarios where they did not perceive themselves as being visually inattentive. Even when participants agreed they were visually inattentive, they would still not have been receptive to some warnings in specific scenarios, such as being stopped at traffic lights. The discrepancies between AttenD outputs and driver self-assessments suggest that DMS should integrate contextual awareness to better align with the driver's mental models. Further research will explore the impact on driver acceptance of different modalities and forms of warnings from DMS in different scenarios.

References

Ahlström, C., Georgoulas, G. and Kircher, K. (2022) 'Towards a Context-Dependent Multi-Buffer Driver Distraction Detection Algorithm', *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, 25(5).

Euro NCAP (2023) 'Assessment Protocol – Safety Assist Safe Driving'.

Euro NCAP (2024) 'Assisted Driving - Test and Assessment Protocol'.

European Commission (2023) 'Commission Delegated Regulation (EU) 2023/2590 of 13 July 2023', *Official Journal of the European Union*, (L 312), pp. 1–15.

Kircher, K. and Ahlström, C. (2013) 'Chapter 19 The Driver Distraction Detection Algorithm AttenD', in *Driver Distraction and Inattention: Advances in Research and Countermeasures, Volume 1*. SAE, pp. 327–348.

Kircher, K. and Ahlström, C. (2024) 'A comparison of glance coding approaches for driver attention assessment', *Transportation Research Part F: Traffic Psychology and Behaviour*, 100, pp. 243–253. Available at: https://doi.org/10.1016/j.trf.2023.12.003.

Smyth, J. *et al.* (2021) 'Public acceptance of driver state monitoring for automated vehicles: Applying the UTAUT framework', *Transportation Research Part F: Traffic Psychology and Behaviour*, 83, pp. 179–191. Available at: https://doi.org/10.1016/j.trf.2021.10.003.

Stanton, N.A. *et al.* (2009) 'Predicting pilot error: Testing a new methodology and a multi-methods and analysts approach', *Applied Ergonomics*, 40(3), pp. 464–471. Available at: https://doi.org/10.1016/j.apergo.2008.10.005.