Method for Inferential Continuous Assessment of Driver's Situational Awareness

Kristina Stojmenova Pečečnik[1], Marko Kofol[1] and Jaka Sodnik[1]

[1] University of Ljubljana, Faculty of Electrical Engineering, Tržaška cesta 25, Ljubljana, 1000, Slovenia

Abstract

This paper proposes a new method for inferential continuous assessment of driver situational awareness (ICA-DSA) that provides the level of knowledge needed to make effective decisions and take appropriate actions. The paper presents the development of the method, which combines eye-tracking and driving performance data to provide a comprehensive situational awareness (SA) assessment of all three SA levels. The approach provides a continuous and non-intrusive assessment that can be applied in simulated and vehicle-based studies. It also presents a user study conducted to collect data for developing the eye-tracking model for assessment of SA level 1, and the rationale behind the selection of the driving performance indicators for SA levels 2 and 3. Finally, it presents the application of the method to the user study results as an example of how it can be used to evaluate new user interfaces.

Keywords

Driver situational awareness, eye gaze, eye tracking, automated assessment

1. Introduction

Situational awareness (SA) plays an important role in any dynamic process of human decision making, as it provides the level of knowledge required to make effective decisions and take appropriate actions [1]. There are several SA definitions, that view and interpret SA from different angles and standpoints [1], [2], [3], [4], [5]. However, closer examination reveals that they all attempt to capture similar key elements about the operator's ability to perceive, understand, and project system status.

The core of understanding and defining situational awareness is the idea of a clear separation between the operator's comprehension of the system status and the actual system status [6]. Consequently, it is expected that better alignment between them should also lead to more successful human-machine interaction between the system and its operator and vice versa. In this regard, assessment of SA has attracted much attention over the last three decades, as it has been found to provide a lot of significant information about the operator, the human-machine interface of the system, and the overall complexity of the system that requires human decision-making in dynamic environments. Although it was primarily observed in aircraft, it has now spread to many other dynamic domains where the environment may be constantly changing. This includes the automotive domain - particularly with the introduction of automation, where the role of the driver, including the scope of situational awareness, changes with each level of automation.

According to the theory SA [1], to achieve SA it is necessary to perceive the elements of the environment (SA level 1), understand their meaning (SA level 2) and be able to project their status in the near future (SA level 3). The three levels are arranged in hierarchical order, with SA level 3 being the highest. The first level is about the perception of all relevant elements in the environment, their status, properties and dynamics. This is followed by the second level, which

HCI SI 2023: Proceedings of the 8th Human-Computer Interaction Slovenia conference 2023, Human-Computer Interaction Slovenia 2023, January 26, 2024, Maribor, Slovenia
kristina.stojmenova@fe.uni-lj.si (K. S. Pečečnik); kofolmarko00@gmail.com (M. Kofol); jaka.sodnik@fe.uni-lj.si (J. Sodnik)

0000-0001-6584-7147 (K. S. Pečečnik); 0000-0002-8915-9493 (J. Sodnik)

focuses on understanding the environment and the meaning of the perceived elements and their properties. The third and highest level of SA reflects the ability to anticipate and predict the actions of the elements in the environment in the near future. Various methods have been developed to assess SA, which generally fall into three categories: query, self-assessment, and inference techniques. Query techniques require operators to self-report information about the system that points to their SA. In this approach, operators are asked questions about their perceptions and understanding of the operated system at a particular point in time. Their answers are then compared to an established (predefined) ground truth. Self-rating techniques, on the other hand, do not interfere with the operation process because they are always presented to the operator after the task is completed. Instead of asking questions about system operation, the operator is asked to provide a (numerical) subjective evaluation of their SA for a given period of time or during the execution of a given task. This type of evaluation is usually based on questionnaires and rating scales that attempt to capture subjective indications of the operator's SA by eliciting the individuals' self-perceptions of the system. Lastly, SA has also been evaluated using inferential or external procedures that seek implicit evidence of the operator's SA using observable and measurable correlates. There is no single format for conducting inferential SA assessment; the individual's performance and behavior are observed using various techniques as indirect evidence of the presence or absence of appropriate SA. This can be done by expert observation of the operator and completion of behaviorally anchored rating scales developed for performance assessment.

1.1. Our contribution

In this paper, we present a new method for inferential continuous assessment of driver's situational awareness (ICA-DSA) that draws inspiration from the greatest strengths of currently available solutions while attempting to overcome their greatest limitations. It combines eye tracking and driving performance data to provide a comprehensive SA assessment of all three SA levels. This approach also provides a continuous and non-intrusive assessment that can be applied to both simulated and vehicle-based studies. Because it does not involve self-rating, it is language-independent, allowing for broad and relatively easy application on a global scale.

In continuation, the paper presents the development of the ICA-DSA method. First, the user study that was used to collect data for the development of the eye tracking model for the assessment of SA level 1 is presented. Then, the rationale behind the selection of the driving performance-based indicators for SA level 2 and SA level 3 is explained. Finally, the application of the method to the user study results is presented as an example of how it can be used to evaluate new user interfaces.

2. Related work

The most widely known methods for assessment of SA are using *performance assessment* with query. The first and still most commonly used method is the Situation Awareness Global Assessment Technique (SAGAT) [7]. Originally, SAGAT was used for assessment of SA of operators operating industrial machines, however it was later used and adapted for numerous other fields. In driving, it has been used to correlate SA ability with driver's age, showing that older adults are less attentive to important information cues compared to younger drivers [8]. It was further used to correlate the working memory with SA, with the results indicating that visual-spatial and auditory cues interfere with the spatial SA of drivers [9]. Beukel & Vort [10] used it to investigate the correlation of headways and response times of distracted drivers in a semi-automated vehicle, finding a positive relationship between advanced warning time and rates of successfully avoided collisions. It was also used for evaluation of novel approaches and interfaces for increasing SA [11], [12]. Based on SAGAT, a similar frame-freeze query method was used to develop a mathematical model intended to describe the dynamic process of building SA after a take-over request in a semi-automated vehicle, showing an exponential relationship between driver's SA and the traffic density, and SA and the time spent under automated mode of driving [13].

Since the process of driving is primarily a visual-manual task, the wide use of eye tracking is somewhat expected. Eye-trackers can be used from monitoring of where the driver is directing their visual attention to [14], [15], [16], up to differentiation of different levels of driver's cognitive load [17]. More specifically, it has been used to observe if the driver detects important cues or are nonessential cues drawing away their attention [18], [19]. Furthermore, it has been used to explore how long it take the driver to regain visual attention or, as referred by Gold et al., environmental SA [20]. From the point of view of the SA theoretical level, eye-tracking is mostly related to assessment of SA level 1, which deals with perception of the environment [21].

Due to its ease and cost-effective nature, another common technique for assessment of driver's SA is with self-assessment and use of questionnaires. Self-assessed data has revealed that SA positively affects trust in automated vehicles [22], [23] and that driver states such as anger can negatively affect SA and driving performance [24].

Lastly, behavior assessment has also been used to evaluate operator's SA. Driver behavior assessment has also been used in driving, but the purpose of the expert evaluation has been mainly to obtain ground truth for initial weighing of neural networks [25] rather than as a standardized behavioral metric. As for assessment of driver's SA, there is no uniform or established process of how and which driving behavior data should be observed. For example, observing the driver's behavior in critical situations was used to correlate shorter response times, headway control and time to collision to SA [26]. Furthermore, Ma & Kaber [27] revealed a significant negative linear association between decreased SA level 3 scores and driving navigation errors, which is in line with Matthews et al. theoretical linkage of SA level 3 with the strategic level of driving behavior [28]. Increased SA due to auditory cues informing the driver of slow traffic ahead resulted in smoother deceleration [29], which is a performance task related to SA Level 2.

3. Methodology

3.1. User study

The study was conducted in a simulated driving environment consisting of a motion-based driving simulator with real car parts (seat, steering wheel, and pedals) and a physical dashboard. The visuals were displayed on three 49-inch curved TV screens that provided a 145° field of view of the driving environment (**Figure 1**). The driving scenario was developed for the purpose of the study in SCANeR Studio [30]. It had a length of 13 km and simulated a route from a suburb to a city center. In the study, we used a conditionally automated vehicle (SAE L3) [31]. During the driving scenario, several intersections with crosswalks and other road users formed the driving environment to create an object-rich test environment.

Figure 1: Driving simulator set-up used in the study

In each trial, there were four prompts to turn on the automated driving system (hereafter referred to as handover request) and four prompts to take over control of the vehicle (hereafter referred to as takeover request). The takeover requests occurred due to both critical (e.g., a busy crosswalk or a complicated intersection) and non-critical events (this was to simulate the vehicle simply losing communication with the infrastructure or the vehicle's sensor system failing).

3.2. Participants

28 (14 male and 14 female) participants took part in the study. The drivers ranged in age from 21 to 57 years (M = 30.17 years, SD = 10.60 years) and had held a valid driving license for an average of 11.77 years (SD = 10.12 years). About half of them (53.33%) reported driving daily, 36.66% several times a week, 6.66% several times a month, and 3.33% several times a year. 20% had no experience with vehicles with automated features (any advanced driving assist systems (ADAS)), while 6.66% had driven a vehicle with multiple ADAS systems once, 13.3% a few times, and 60% several times. Data from 9 of the 28 participants had to be excluded due to motion sickness or because the data sets were recorded partially and were hence unusable due to technical problems for the ICA-DSA method development. The data available for all 28 participants were used for the evaluation of the HUD. As a thank you for their participation in the study, the participants received a gift voucher for €10.

3.3. Experiment design and procedure

The goal of the study was two-fold. First, it aimed to collect data necessary for the development of the ICA-DSA method. Second, it aimed to present the application of the method for assessment a novel human-computer interaction (HCI) solution.

For the later purpose, the study had a within subject design – all participants performed two trials:

- a baseline trial,
- a trial with the addition of a head-up display (HUD), shown on **Figure 1**.

The HUD displayed information about the vehicle speed, the speed limit, indicated speeding, highlighted traffic signs, indicated too short distances to the vehicle ahead, and highlighted (with bounding boxes) important road participants during take-over, which could affect the course of driving. The HUD was intended to help drivers with the perception of the environment (SA level 1), but also to contribute to safer and smoother longitudinal control of the vehicle. Half of the participants started the study with the baseline trial, and the other half with trial with the HUD.

The study began with the experimenter explaining the purpose of the study and informing participants that their only task was to operate the vehicle safely. Participants were informed that they can stop their participants at any time if they felt any discomfort or motion sickness. Upon receiving all information concerning the study, the participants were provided with an informant consent, which after it was signed, was followed by a demographic questionnaire. In this study, we did not collect any personal information. After signing the consent forms, all of the data was recorded under unique randomly assigned IDs.

The participants were then subjected to a practice drive in which they were shown how to use the driving simulator, and how to turn on and turn off the automated driving system (ADS). They received the handover request in the form of a pre-recorded voice message saying “Turn on automated driving system”. The ADS could be turned on by pressing a specifically dedicated ADS button on the bottom left lever of the steering wheel. They received the takeover request as a visual and auditory notification 5 seconds prior to the ADS turning off on itself. The visual takeover notification was a text message “Takeover”, which was accompanied with a countdown from 5 to 0, showing the time remaining for takeover. The auditory notification was a 4000 Hz pure tone [32] which was played at 65 dB from the start of the takeover notification until the driver took over control of the vehicle. Participants were able to take over control of the vehicle by pressing on the brake or gas pedal for at least 40 N, steering the wheel for at least 6° or by pressing the same ADS button on the bottom left lever of the steering wheel used for turning on the automated system.

The participants then proceeded to completing the two trials. In between there was a 2-minute break. After completing the trials, they were provided with the gift voucher.

3.4. Development of ICA-DSA

3.4.1. Eye tracking

Video recordings from the driver's point of view recorded with Tobii Pro Glasses 2 eye tracker [33] were used as the main data source for assessment of SA level 1 (perception of elements in the environment [1]). Tobii Pro Glasses 2 are a head-mounted eye tracker with sampling frequency of 50 Hz. The eye tracker provides information about the gaze coordinates within the coordinate system of the video recording, enabling visualization of the gaze position as a small green circle in the video scene (see **Figure 2** annotation view_marker).

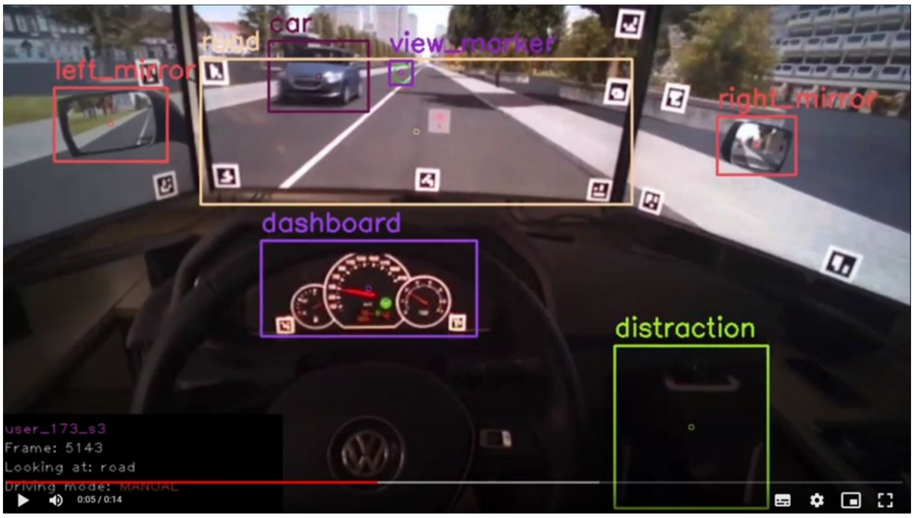

Figure 2: Objects of interest observed during driving

For analyzing the eye tracking videos, we used You Only Look Once (YOLO) object identifier for object detection and recognition. YOLO can recognize and locate multiple objects within an image or video. Its libraries typically provide several pre-trained models for object detection, classification and segmentation in driving scenes, mainly based on real-life driving recordings. For our model, we used driving-based video recordings from a driving simulator. Consequently, we could not use existing YOLO models, but had to perform the training process from scratch with the simulator-based recordings and data. To train the model, we decided to observe only the driver's SA during the handover and takeover requests, which resulted in 8 video extractions. The video extractions started with the handover/takeover requested and lasted 15 seconds after the driver's accepted the requested, which (in average) resulted in 20 seconds long videos for the handovers, and 30 seconds long videos for the takeovers.

With the goal of capturing driver's SA level 1, objects of interest observed were the rear mirrors (left, right and center), other road participants (vehicles, pedestrians, cyclists), the physical dashboard in the simulator, the projected HUD (in the trials with the HUD) in the simulation, the physical head-down display (HDD) for display of entertainment content, and the gaze position of the test participant. All of the points of interest are presented in **Figure 2**.

The processing was performed in. The first step of the process was to select 300 static images in the video recordings to be used as a training set and mark all objects of interest visible in the individual image. This process was performed by YOLO Label tool [34] which outputs a configuration file with bounding boxes of all objects. The set of images is split into subsets for training (70%), testing (15%) and validation (15%). The model itself was trained using the YOLO Custom Training model.

The final step was then to analyze all of the video recordings using our pre-trained model by implementing it in PyTorch [35]. A position of the gaze in each frame was compared against the other detected objects and the potential overlap was recorded as "seen object". **Figure 2** shows a screenshot of the analyzed video with a set of predefined objects of interest and user's gaze position (view marker). In the right bottom corner of the video, the experimenter can see at all times the data being recorded, allowing for monitoring of the reliability of the detection model. Attached to this submission is a short video extraction showing the process in practice.

3.4.2. Driving performance

ICA-DSA further foresees use of driving performance data as indicators for assessment of SA level 2 (comprehension on the meaning of elements in the environment) and SA level 3 (projection of near future events), mainly focusing on lateral and longitudinal control of the vehicle [1].

For observing SA level 2, ICA-DSA looks at:

- % longitudinal accelerations < 1.23 m/s^2 < 2.12 m/s^2,
- % longitudinal decelerations < 1.13 m/s^2 < 2.02 m/s^2,
- % lateral accelerations < 1.64 m/s^2 < 1.87 m/s^2, and
- % speeding 10% above the speed limit (as defined by national regulation).

For observing SA level 3, ICA-DSA looks at:

- % longitudinal accelerations > 2.12 m/s^2,
- % longitudinal decelerations > 2.02 m/s^2,
- % lateral accelerations > 1.87 m/s^2, and
- % number of accidents (collisions with other road participants or road infrastructure).

The acceleration and deceleration ranges were derived from de Winkel et al. study results [36] in which they suggested standards for lateral and longitudinal acceleration rates in automated vehicles, which were somewhat in line also with standards for accelerations in (manually operated) public transportation [37]. They defined longitudinal accelerations, longitudinal decelerations and lateral accelerations below 1.23 ms^{-2}, 1.13 ms^{-2}, and 1.64 ms^{-2} respectively as "good", whereas above 2.12 ms^{-2}, 2.02 ms^{-2} and 1.87 ms^{-2} respectively as terrible. Based on their results, we defined accelerations above the "good" rating as indicators of poor SA, which due to lack of comprehension of the meaning of status of elements in the environment requires sudden adjustments of the lateral and longitudinal control of the vehicle. Excessive accelerations on the other hand, usually result in accidents or near accidents [38], due to the driver's lack of projection of the status of the elements in near future.

For the purpose of this study, the driving performance data was captured with a motion-based driving simulator and SCANeR simulation software. The data was aggregated with using a data logger with sampling frequency of 100 Hz. The data logger records positions and velocities of all elements in the driving environment, including the ego-vehicle. To be able to align the driving performance data with the eye tracker data, which is logged at 50 Hz, the driving simulator data was down sampled from 100 HZ to 50 Hz.

3.5. Independent variables used in the study

For presentation of results obtained with ICA-DSA, we defined two independent variables in the study: absence and presence of the HUD. Additionally, we observed separately the data for the handover and takeover request. The data was collected from receiving the notification (for handover or takeover) until accepting it, and 15 seconds after acceptance to differentiate between manual and automated driving mode. During the automated drive (after handover and before takeover), we observed only the eye tracking data, as the driving performance data was not from the driver but from the ADS.

4. Results

In addition to applying the ICA-DSA model, the obtained data for each trial was compared with paired samples t-test to compare the values between the baseline trial and trial with HUD. The Shapiro-Wilk test of normality showed that the eye-tracking and driving performance data were not normally distributed ($p > 0.05$). As a result, the data was analyzed with Wilcoxon signed-rank non-parametric test.

4.1. Eye tracking

The eye-tracking model provides information about the driver's gaze in every recorded frame. Since the observed video recordings represent specific situations during takeover (handover and takeover), which have defined short time durations, we present the results as percentage of time the driver spent looking at a specific (type of) of object in interest during each situation.

4.1.1. Handover

Figure 3 presents the mean percentage of time drivers looked at a specific object during handover. The data is split in two time intervals: 1) from handover notification until the actual handover (**Figure 3**, left) when the driver operates the vehicle manually, and 2) 15 seconds after right after handover (**Figure 3**, right) when the vehicle is operated by the ADS. It should be noted that the reduced time spent looking on the road in the trials with the HUD is due to the fact that the figures do not include the percentage of time the driver's gaze was on the HUD during the trials with the HUD, which were $M = 9.07\%$ ($SD = 14.59\%$) during the manual, and $M = 6.32\%$ ($SD = 11.31\%$) during the automated driving intervals. The HUD was displayed on the lanes, therefore overlapping with the road bounding box. Because the HUD is semi-transparent, the drivers could still pay attention to road, so looking at the HUD should not be considered entirely as not looking on the road. The Wilcoxon signed-rank test revealed statistically significant increase in the percentage of time drivers spent looking at other vehicles during before the handover, $Z = 2.130$, $p = 0.033$, and pedestrians after the handover, $Z = -2.216$, $p = 0.027$.

4.1.2. Takeover

Figure 4 presents the mean percentage of time driver's looked at a specific object during takeover. The data is split in two time intervals: 1) from takeover notification until the actual takeover (**Figure 4**, left) when the vehicle is operated by the ADS, and 2) 15 seconds after right after takeover (**Figure 4**, right) the driver operates the vehicle manually. Again, the figures do not include the percentage of time the driver's gaze was on the HUD during the trials with the HUD, which were $M = 7.64\%$ ($SD = 11.05\%$) during the automated, and $M = 5.43\%$ ($SD = 11.06\%$) during the automated driving intervals. The Wilcoxon signed-rank test revealed only one statistically significant increase: in the percentage of time drivers spent looking at pedestrians, both before takeover ($Z = 2.547$, $p = 0.011$), and after takeover ($Z = -2.008$, $p = 0.045$).

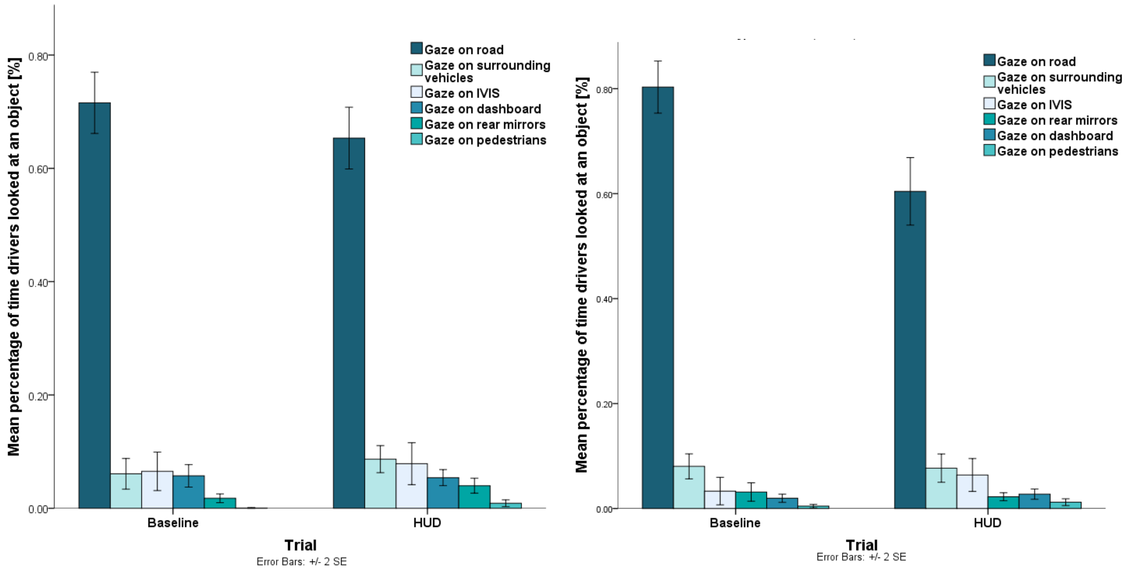

Figure 3: Mean percentage of time drivers spent looking at specific types of objects during handover. The left graph shows from handover request manual driving (left), whereas the right graph after handover automated driving (right)

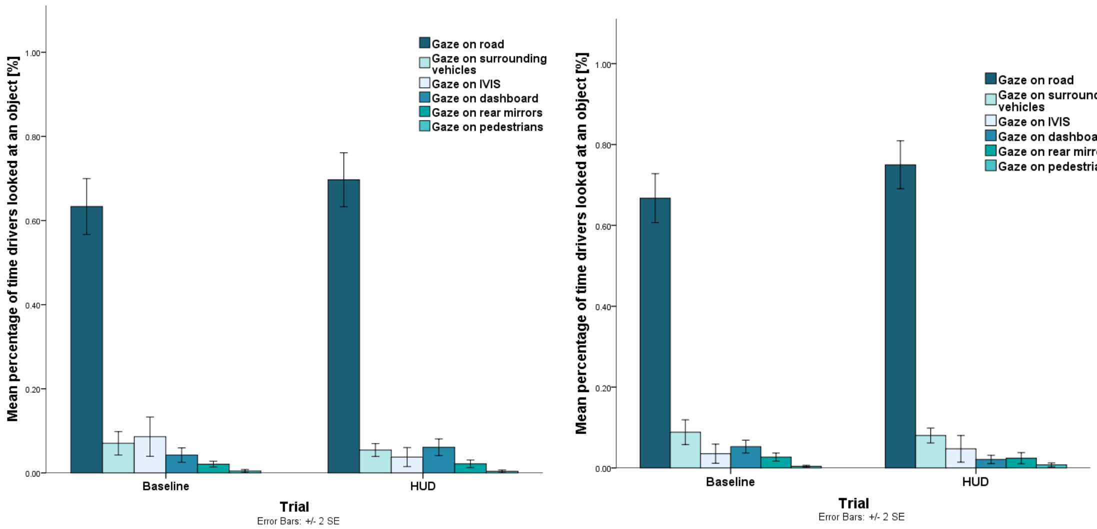

Figure 4: Mean percentage of time drivers spent looking at specific types of objects during takeover. The left graph shows from takeover request automated driving (left), whereas the right graph after takeover manual driving (right)

4.2. Driving performance

The eye-tracking model provides information about the driver's gaze in every recorded frame. Since the observed video recordings represent specific situations during takeover (handover and takeover), which have defined short time durations, we present the results as percentage of time the driver spent looking at a specific (type of) of object in interest during each situation.

4.2.1. Handover

Figure 5 presents the mean percentage of time drivers performed a specific driving performance indicator from the handover notification until the actual handover. The Wilcoxon signed-rank test revealed only one statistically significant increase - in the percentage of time the drivers' longitudinal deceleration was above 1.13 m/s2, Z = 2.450, p = 0.014 when operating the vehicle with a HUD compared to the baseline trial. There were no accidents in any of handover situations in the baseline and HUD trials.

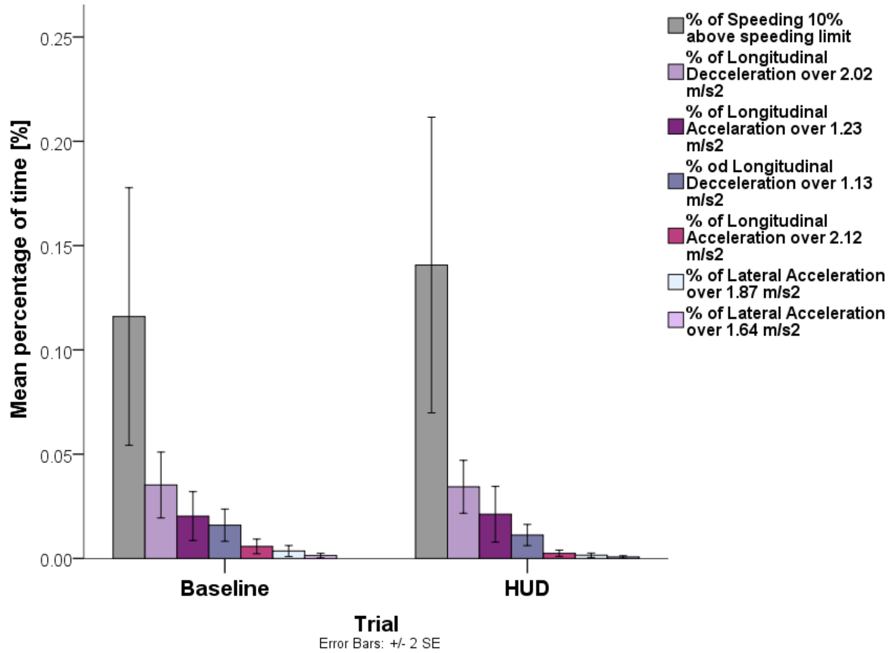

Figure 5: Mean percentage of time drivers performed a specific driving performance indicator before handover

4.2.2. Takeover

Figure 6 presents the mean percentage of time drivers performed a specific driving performance indicator 15 seconds after taking over the vehicle control. Also in this case the Wilcoxon signed-rank test revealed only a statistically significant decrease in the percentage of time the drivers' longitudinal deceleration was above 1.13 m/s2, $Z = -2.108$, $p = 0.035$ when operating the vehicle with a HUD compared to the baseline trial. There were no collisions in any of the takeover situations in the baseline and HUD trials.

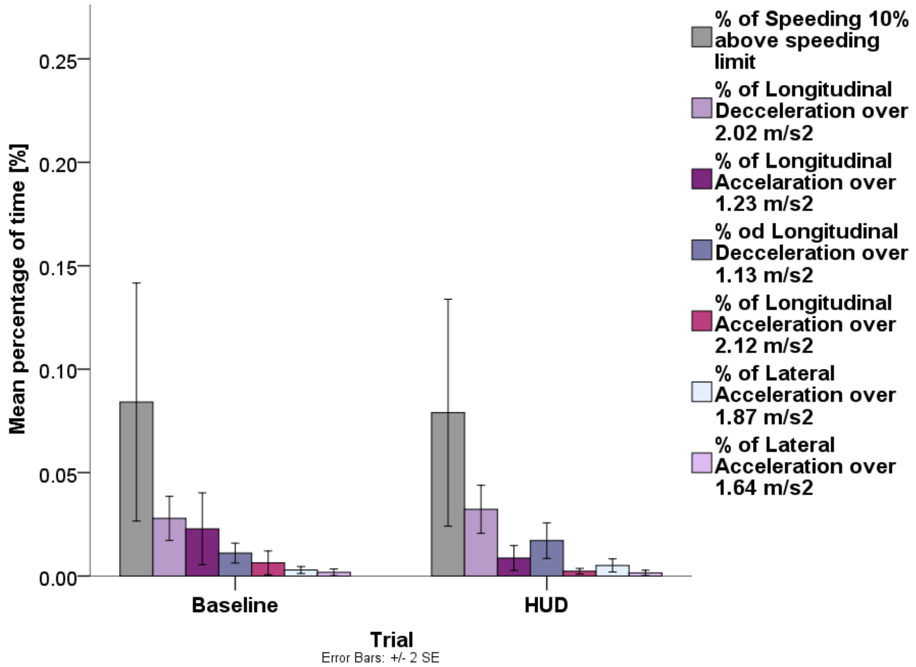

Figure 6: Mean percentage of time drivers performed a specific driving performance indicator after takeover

5. Discussion and conclusions

The goal of this study was to develop an inferential, non-intrusive, continuous, automated, and language-independent method for assessment of driver's SA at all three levels defined in SA theory. To do so, we conducted a user study to collect data that was used to train the models for SA level 1, which deals with the perception of elements in the environment. Based on the available literature, we used the data to further define driving performance indicators that can be used for assessment of SA level 2 and SA level 3, which can be retrieved in the same data format as the eye-tracking data.

The results of the user study demonstrate the application of the method to evaluate a novel HUD, designed to improve driver's SA during the transition of control between the driver and the vehicle in semi-automated vehicles. The results for SA level 1 revealed that the introduction of a HUD, which highlights other important road participants such as other vehicles and pedestrians, can increase the driver's attention to elements in the environment during the handover or takeover respectively. The results for SA level 2 showed that the addition of the HUD resulted in less excessive deceleration and thus a smoother ride. Sudden braking can occur to adjust longitudinal vehicular control to avoid speeding or collisions after (late) detection of elements in the environment. Because the HUD displayed information about speeding, too short distance to the vehicle in front as well as highlighted important road participants, drivers may have been able to better understand their significance and thus reduce the need for excessive braking. The observed driving performance indicators for SA level 3 did not show statistically significant differences, and some of them did not occur at all. The reason for this could be due to the relatively uneventful scenario that was mainly used to train the eye tracking model.

In addition to assessment of new interfaces such as the presented HUD, the ICA-DSA results could be used to gain an understanding about the driver in specific situations such as handover and takeover. For example, drivers seem to pay more attention to all road environment elements after operating the vehicle manually for some time (before handover). In contrast, they seem to spend more time looking straight on the road and fail to disperse their attention to other objects (including rear mirrors) during takeover after having spent some time being driven by the ADS. However, further studies dedicated for assessment of such specific situations are needed to further explore the existence of statistically significant differences and the what may be causing them.

Although it provides an automated method for capturing data for every object of interest and driving performance indicator in every frame, in its current form, the proposed ICA-DSA method does not provide a definite score, still requiring interpretation of the obtained results. Further research is needed to define a SA scale for every SA level, which could be then used to calculate an overall driver’s SA. As it is a continuous-based method, the scores could be calculated for a defined period of time; a duration that could be set around a time interval or situation-based. A potential step forward would be conducting an expert-based study, to define SA scale-based scores, which could also be calculated automatically, and could hence ease the interpretation process of the results. The study could also involve other existing query, self-rating and inferential SA methods (for example, SAGAT, SART and SABARS) to validate the sensitivity of ICA-DSA and check for existence of correlations among them.

Another limitation revealed with the study lays in the driving performance indicators used for assessment of SA level 3. A future step would be identification of more performance indicators that could indicate driver’s ability to predict the status of elements in near future, which do not require critical events such as (near) collisions.

Lastly, because the eye-tracking model was developed based on video recordings of driving simulator, this paper presents the application of ICA-DSA for studies conducted in a simulated driving environment. However, since YOLO models for object detection in videos from real vehicles are already available, ICA-DSA could be easily adapted and used also for assessment of driver’s SA based on real vehicle data.

At this point, we believe that in the presented form ICA-DSA provides a good starting point for overall assessment of driver's SA as it already minimizes the limitations of currently available solutions.

Acknowledgements

The work presented in this paper was financially supported by the Slovenian Research Agency within the project Modelling driver's situational awareness, grant no. Z2-3204 and program ICT4QL, grant no. P2-0246.

References

[1] Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. In Human Factors Journal, 37(1), 32-64.
[2] Bedny, G., & Meister, D. (1999). Theory of activity and situation awareness. *International Journal of cognitive ergonomics*, *3*(1), 63-72.
[3] Endsley, M. R. (1988, May). Situation awareness global assessment technique (SAGAT). In *Proceedings of the IEEE 1988 national aerospace and electronics conference* (pp. 789-795). IEEE.
[4] Fracker, M. L. (1991). Measures of situation awareness: Review and future directions.
[5] Smith, K., & Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human factors*, *37*(1), 137-148.
[6] Woods, D. D. (1988). Coping with complexity: the psychology of human behaviour in complex systems. In *Tasks, errors, and mental models* (pp. 128-148).
[7] Endsley, M. R. (2000). Direct measurement of situation awareness: Validity and use of SAGAT. In *Situational Awareness* (pp. 147-174).
[8] Bolstad, C. A. (2000, October). Age-related factors affecting the perception of essential information during risky driving situations. In *Human Performance Situation Awareness and Automation: User-Centered Design for the New Millennium Conference, Savannah, GA*.
[9] Johannsdottir, K. R., & Herdman, C. M. (2010). The role of working memory in supporting drivers' situation awareness for surrounding traffic. *Human factors*, *52*(6), 663-673.
[10] van den Beukel, A. P., & van der Voort, M. C. (2013). Retrieving human control after situations of automated driving: How to measure situation awareness. In *Advanced Microsystems for Automotive Applications 2013: Smart Systems for Safe and Green Vehicles* (pp. 43-53). Heidelberg: Springer International Publishing.
[11] Scholtz, J., Antonishek, B., & Young, J. (2004, January). Evaluation of a human-robot interface: Development of a situational awareness methodology. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the* (pp. 9-pp). IEEE.
[12] Sirkin, D., Martelaro, N., Johns, M., & Ju, W. (2017, May). Toward measurement of situation awareness in autonomous vehicles. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 405-415).
[13] Lu, Z., Coster, X., & De Winter, J. (2017). How much time do drivers need to obtain situation awareness? A laboratory-based study of automated driving. *Applied ergonomics*, *60*, 293-304.
[14] Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of experimental psychology: Applied*, *9*(1), 23.
[15] Stojmenova, K., Marinko, V., Komavec, M., & Sodnik, J. (2019). Effects of phoning during driving. In proceedings of the 9th International Conference on Information Society and Technology, ICIST 2019.
[16] Stojmenova, K. (2020). Assessing the attentional effects of cognitive load in driving environments. V: Wearable eye tracking: online user meeting: September 8 - September 9, 2020. Tobii Pro.

[17] Čegovnik, T., Stojmenova, K., Jakus, G., & Sodnik, J. (2018). An analysis of the suitability of a low-cost eye tracker for assessing the cognitive load of drivers. Applied ergonomics, 68, 1-11.
[18] Barnard, Y., & Lai, F. (2010). Spotting sheep in Yorkshire: Using eye-tracking for studying situation awareness in a driving simulator. In H*uman factors: a system view of human, technology and organisation.* Annual conference of the europe chapter of the human factors and ergonomics society 2009.
[19] Samuel, S., Borowsky, A., Zilberstein, S., & Fisher, D. L. (2016). Minimum time to situation awareness in scenarios involving transfer of control from an automated driving suite. Transportation research record, 2602(1), 115-120.
[20] Gold, C., Damböck, D., Lorenz, L., & Bengler, K. (2013, September). "Take over!" How long does it take to get the driver back into the loop? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 57, No. 1, pp. 1938-1942). Sage CA: Los Angeles, CA: Sage Publications.
[21] Schömig, N., & Metz, B. (2013). Three levels of situation awareness in driving with secondary tasks. *Safety science*, *56*, 44-51.
[22] Sonoda, K., & Wada, T. (2017). Displaying system situation awareness increases driver trust in automated driving. *IEEE Transactions on Intelligent Vehicles*, *2*(3), 185-193.
[23] Petersen, L., Robert, L., Yang, J., & Tilbury, D. (2019). Situational awareness, driver's trust in automated driving systems and secondary task performance. *SAE International Journal of Connected and Autonomous Vehicles, Forthcoming*.
[24] Jeon, M., Walker, B. N., & Gable, T. M. (2015). The effects of social interactions with in-vehicle agents on a driver's anger level, driving performance, situation awareness, and perceived workload. *Applied ergonomics*, *50*, 185-199.
[25] Komavec, M., Kaluža, B., Stojmenova, K., & Sodnik, J. (2019). *Risk Assessment Score Based on Simulated Driving Session*, in *2019 Driving Simulation Conference Europe*. 2019. p. 67-74.
[26] Merat, N., & Jamson, A. H. (2009). Is Drivers' Situation Awareness Influenced by a Fully Automated Driving Scenario? In Human factors, security and safety. Shaker Publishing.
[27] Ma, R., & Kaber, D. B. (2007). Situation awareness and driving performance in a simulated navigation task. *Ergonomics*, *50*(8), 1351-1364.
[28] Matthews, M. L., Bryant, D. J., Webb, R. D., & Harbluk, J. L. (2001). Model for situation awareness and driving: Application to analysis and research for intelligent transportation systems. *Transportation research record*, *1779*(1), 26-32.
[29] Nowakowski, C., Vizzini, D., Gupta, S. D., & Sengupta, R. (2012). Evaluation of real-time freeway end-of-queue alerting system to promote driver situational awareness. *Transportation research record*, *2324*(1), 37-43.
[30] AV Simulation. SCANeR Studio. Avaliable at: https://www.avsimulation.fr/solutions/.
[31] SAE, T. (2016). Definitions for terms related to driving automation systems for on-road motor vehicles. *SAE Standard J*, *3016*, 2016.
[32] Stojmenova, K., Policardi, F., & Sodnik, J. (2018). On the selection of stimulus for the Auditory Variant of the Detection Response Task Method for driving experiments. *Traffic injury prevention*, *19*(1), 23-27.
[33] Tobii Llc. Stockholm, Sweden. *Tobii Pro Glasses 2 wearable eye tracker*. Available online: https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/.
[34] YOLO Label tool. Available online: https://github.com/developer0hye/Yolo_Label.
[35] The PyTorch Foundation. PyTorch. Available online: https://pytorch.org/.
[36] de Winkel, K. N., Irmak, T., Happee, R., & Shyrokau, B. (2023). Standards for passenger comfort in automated vehicles: Acceleration and jerk. Applied Ergonomics, 106, 103881.
[37] Hoberock, L. L. (1976). A survey of longitudinal acceleration comfort studies in ground transportation vehicles. Council for Advanced Transportation Studies.
[38] Guillen, M., Nielsen, J. P., Pérez-Marín, A. M., & Elpidorou, V. (2020). Can automobile insurance telematics predict the risk of near-miss events? *North American Actuarial Journal*, *24*(1), 141-152.