Feasibility of AR-VR use in autonomous cars for user engagements and its effects on posture and vigilance during transit

JOSEPH MUGURO*a,b,* , PRINGGO WIDYO LAKSONOc, YUTA SASATAKEd, MUHAMMAD ILHAMDI RUSYDIe, KOJIRO MATSUSHITAd, MINORU SASAKIa,b,*

a. Intelligent Production Technology Research & Development Center for Aerospace (IPTeCA), Tokai National Higher Education and Research System, Japan
b. School of Engineering, Dedan Kimanthi University of Technology, 657-10100, Nyeri, Kenya
c. Industrial Engineering, Universitas Sebelas Maret, Surakarta, Indonesia
d. Department of Mechanical Engineering, Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu, 501-1193, Japan
e. Department of Electrical Engineering, Engineering Faculty of Universitas Andalas, Padang 25163, Indonesia

ABSTRACT: Autonomous driving system (ADS) is anticipated to revolutionize travel by reclaiming lost time and improve safety on the roads. With automation, user-engagements that enhances road monitoring should be considered to maintain vigilance and safety. From the literature, virtual reality (VR) usage in cars offer productivity and increased privacy. This paper explores the efficacy of passenger use of VR headsets to enhance user engagement during transit. User-engagement was quantified using physiological measures (pupillary response and electrodermal activity) during an in-car VR game/activity experiment. Further, the impacts of engaging with secondary tasks was evaluated using reaction time of pop-up objects. We designed a driving simulation with inbuilt entertaining activities, no-task, game-task, video-task, and mixed-task, played in a real car with a FOVE VR headset on the perimeter track of the Gifu University campus with 15 subjects (average 25.6 years, SD = 6.4).

From reaction time, significant difference between tasks was found using one-way ANOVA ($F(3,231) = 2.75, p = .0437$). A post-hoc test revealed that game and mixed task reaction times were significantly different ($p = .0126$ and $p = .016$, respectively) suggesting that task design should consider hazard recognition in a real car. From physiological measures, an increased/sustained effect of user engagement was noted compared with baseline (no-task) suggesting effectiveness in maintaining vigilance. The results also reported a 10-fold improvement in sitting posture with baseline (no-task) suggesting effectiveness in maintaining vigilance study in an ADS.

KEYWORDS: Autonomous Driving Systems; Driver/passenger engagement; Driving Related Tasks, 3D-VR/AR, In-Car VR

1. INTRODUCTION

One of the anticipated technology advancements is the autonomous driving system (ADS), which seeks to replace the monotonous drive with a more productive transit time. Following automation levels defined by the society of automotive engineers (SAE) standard, in level 3 and above, the focus is on what activities the driver engages in as he/she shifts from controlling the car (Standard, 2018). ADS, though in active development, is being explored through different conceptual designs by automobile manufacturers.

A consensus with researchers and automakers is on the foreseeable transition to car-of-the-future, wherewith a remodeling or fine-tuning of the current vehicles is necessary (Anderson et al., 2014). Conceptual designs have been explored with the suggested reorientation of the interiors to have forward and backward seating, as noted by (Salter et al., 2019). Adjustment of window size has also been proposed to reduce production costs (Wada, 2017), (Deahl, 2018). Manufacturers are working towards providing an environment for the driver/user to work, communicate, be entertained, or even nap during transit.

Such an environment can be attained in a moving vehicle through augmented/virtual reality (AR/VR). AR/VR is a promising venture that has been applied in education, automotive industry, robot operations, surgeries, to name a few. Developers have proposed immersive experiences to reinroduce artificial experiences to the car user (Audi, 2019; Deahl, 2018; Mark B. Rober, Sawyer I. Cohen, Daniel Kurz, Tobias Holl, Benjamin B. Lyon, Peter George Meier, Jeffrey M. Riepling, 2018). These experiences would cover games, holographic meetings, remote working environments, avatar conferencing, amongst others. In this application form, the user is supplied with visual and other haptic contents that are either synchronized or disconnected to the present physical reality. As a use-case, a user playing an immersive in-car game with a head-mounted display (HMD) will be interacting with game elements visually and experience the dynamics of a moving car. As such, there are two distinct realities (physical and cyber) that will either be competing or working together. The disconnect between physical and virtual reality has been shown to increase discomfort (motion sickness/cyber sickness), particularly in HMD. Reading a book and other passive activities similarly increases discomfort, manifesting as motion sickness in a moving car.

With the introduction of VR in the car, special care should be accorded as per user experience.

The proposal to introduce VR in cars has mixed views and perceptions in research communities. On the one hand, drawing from the discomfort formed in the usage of VR systems, some see it as a less feasible solution to the problem at hand...
This paper focuses on using VR in a car environment to offer an environment that can be manipulated to be conducive for productivity and entertainment. VR usage within the car environment is the so-called in-car VR, which touches on the use of HMD in an actual moving car whether as a passenger or a driver. In the proposed setup, the investigation is on the use of VR by a passenger.

Research targeting in-car VR use has been done in the past. McGill et al. explored the first on-road immersive VR with varying visual presentation of the real-world motion (McGill et al., 2017a). The authors investigated optimal visual presentations of motion in VR in a bid to minimize sensory conflict. From the results, the paper found no particular best system that balances immersion and sickness. As a primer to the study, experimentation of in-car VR use was recommended.

A paper (Paredes et al., 2018) explored in-car VR to create a calm and mindful experience for ADS users. The authors utilized dynamic and static scenes to investigate an ideal experience for a moving and a parked car system. In the study, congruency was explored in either a static physical environment (parked car) with a static virtual environment (congruent condition) compared with a static physical environment and a dynamic virtual environment (incongruent condition). The users experienced a fully immersed, underwater exploration movement synchronized (loosely) with the car motion. The users were virtually translocated in a calming beach scene with no car motion in both car movement and car-parked case on the static scene. The authors reported that diving in the ocean in a moving car had lower autonomic arousal levels than static VR in a stationary car condition. Also, the authors noted the incongruence between car movement and VR content.

Another research utilizing in-car VR is reported in (Hock et al., 2017) targeting VR entertainment. The participants engaged in a rail-shooter game in a static (parked car) and dynamic (moving car) environment in the setup. The setup features synchronization of physical space to cyberspace in the sense that kinesthetic congruence between visual (virtual world) and vestibular information (physical car movements) is maintained. This was achieved by relegating car motion (from the onboard diagnostic board) as VR scene motion commands. The authors concluded that perceived kinesthetic forces caused by in-car VR potentially increase enjoyment and immersion while reducing simulator sickness compared to a static environment.

Other research on in-car VR has focused on challenges of passenger experience, cooperative game-play, VR/AR for driving, and posture alignments (Broy et al., 2011; McGill et al., 2017b; Zuckerman et al., 2014).

From the above, when car motion cues and visual information from HMD are mismatched, there is a surge in nausea and general user discomfort (Wada, 2017), (McGill et al., 2019). As such, synchronicity should be considered adequately for the in-car VR experience. Synchronized tasks consider physical car attributes like acceleration, braking, turns, and geo-location and integrate this in the VR environment. This would feature as scene in VR that moves or turns with every turn of the vehicle. A use case is an infotainment system that gives contextual information to users.

A paper by Kuiper et al. suggested that knowing what is coming is helpful to prepare and anticipate motion, thereby reducing motion sickness (Kuiper et al., 2020). Visual feedback is the natural response (or the lack of) in anticipatory reactions and responses. According to Wada et al., the posture of the car occupants can be fine-tuned to mitigate sickness (Wada, 2017). Superseding visual information has been pointed out as a strategy to reduce discomfort in an ADS [19]. At present, we investigated what tasks would encourage postural adjustments.

An equally important inquiry has to do with what task engages users of ADS. Inquiry of in-car VR/AR scenes that is task-specific and exploring the challenges presented in the areas is vital for the overall acceptability of the ADS system. With this in mind, this paper seeks to investigate in-car VR user experiences with varying tasks in an actual car.

Role changes anticipated in ADS, where the driver/occupant is not mandated with visual supervision of the road, will increase the loss of situational awareness and vigilance (Korber et al., 2015; X. Li et al., 2020). Situational awareness is significant for level 3 and 4 of SAE standards that requires human intervention. The design of in-car VR or tasks can supplement this by offering contextually relevant information alongside the engagement modality. Researchers seek what precipitates a good resumption of control. Several authors have investigated non-driving-related tasks, the activities performed by the driver during an active driving session, to measure its impacts (Jeong & Liu, 2019; X. Li et al., 2020). Since ADS will eliminate the need for active driving inputs and constant monitoring of the road, activities performed by the driver will not be categorized as a distraction; driving will be the new distraction as roles get reversed (Miller et al., 2015) (Dingus et al., 2006). In this paradigm shift, distraction is desirable in a car environment, i.e., driving-related tasks (DRT). In this case, a strong appeal is to keep the users vigilant by activities/engagement that helps in indirect road monitoring as a safety measure. From (Standard, 2018), a fallback-ready user should be receptive to requests or eminent vehicle system failure whether a take-over request is issued or not. Waymo’s® road safety performance data reported 47 collision and minor contacts for 2019/2020 operations (Schwall et al., 2020). Besides this, news about the fatal accidents involving self-driving cars still looms with the usual human fault in the fallback-ready user, as is the case in [28]. From the above, the limitations of the ADS will continue being bottlenecked towards safety as long as ADS share roadways with human drivers, way past the fully autonomous levels are arrived at (Schwall et al., 2020; Waymo, 2020). What needs to be addressed is to optimize safety by ensuring direct or indirect road monitoring of fallback users for readiness to take-over control.

In this paper, borrowing from previous research, we propose to explore further the usage of in-car VR for entertainment modality which enhances drivers’ vigilance. Compared to the previous research, this paper goes further in investigating in-car VR, utilizing game and video tasks, the two most engaged in pass-time activities, and explore the physiological aspects of engagement using eye data and electrodermal activity (EDA) levels for each engagement. Further, we explored the use of road monitoring for pop-up objects, theorized as monitoring for hazards along the drive path. The recognition/reaction time is recorded for analysis and comparison.

We investigated the feasibility of performing four distinct tasks; no task/baseline, game task, video clip task, and mixed video and game task, experienced in VR within a moving vehicle environment. To ascertain feasibility and the engagement levels, we measured body sways, eye gaze information, pupil size variations, and EDA levels of each of the task. These physiological measures will be used to determine engagement levels of different activities. In summary, the research sought to inquire on the following.
- Investigate user’s recognition of threatening driving scenarios while engaging in different virtual tasks.
- Evaluate user engagement with different in-car VR elements.
- Investigate the influence of content design on posture (head movement) in 3D space.

The rest of this document is divided as follows. Section 2 presents the materials and methods applied; section 3 describes the results derived from the experiment. Section 4 gives discussion and a corresponding conclusion in section 5.

2. MATERIALS AND METHODS

2.1 Driving simulator

We designed a custom car simulator scene with Unity 3D game engine for the proposed driver analysis. The VR content was designed to be experienced by passengers in a real-world moving car as opposed to previous design that targeted parked cars or reduced car dynamics (straight lines) (Hock et al., 2017; Muguro et al., 2021). The car in use was a Subaru Stella Matic, with a displacement of 600cc, with no special relaxation features in place. The experiment was conducted on the premises of Gifu University, Japan. The area covers an approximate distance of 5 Km. The area map was selected as a testing ground as it featured typical road conditions in Japan, albeit with a reduced drive speed. The campus has a speed limit of 20 Km/h, six speed bumps, three barrier gates for entrance and exit, and six crosswalks around the perimeter. The environment is relatively flat, with one-way asphalt roads.

The primary building block (VR setup) involved the following units: a gaming PC and VR HMD for rendering the virtual scene, a GPS tracking device(s) for localization in the virtual map, and a physiological recording unit, palmar EDA. FOVE® HMD was used for content rendering in the prototype. FOVE is a 6D VR that avails positional and rotational data during use. Also, the device has an inbuilt eye-tracking system. The simulation was run on a windows-10 gaming PC (Mouser G-Tune P5 ) with an Intel® Core i7 processor and GeForce GTX 1650 graphics card.

GPS was opted for to avoid the processes for signal acquisition from the OBD of the car. In this case, the position of the virtual vehicle is controlled by the position received by the actual car. We utilized two GPS sensor arrays (AE-GNSS-EXTANT (K-13850) from Akizushi Denshi Co.), with an update rate of 1 Hz. The GPS signal is fed directly to the laptop PC rendering the game using serial communication and parsed to extract latitude and longitude coordinates. The received coordinates are then converted to X,Y and Z coordinates with World Geodetic System 1984 (WGS84) reference system. From these, X and Z coordinates were set as target move position of the virtual vehicle in Unity 3D. Y-axis (height from ground) was maintained as a constant in the scene, giving the virtual vehicle a hovering-like motion. The rotation towards the target position was calculated to orient the virtual vehicle in the y-axis only. Instead of a closed car environment, we opted for a convertible type of vehicle that does not occlude the 3D scene view and gaze.

Physical car rotation has been reported to interfere with VR rotation in previous studies whereby vehicle turnings are perceived by the HMD inertial measurement unit as head movements with a corresponding disorientation of view. As a remedy, different authors opted to either preconfigure rotations or have third parties orient the view (Hock et al., 2017; Paredes et al., 2018). In this paper, the virtual vehicle’s rotation is calculated by the game engine. Thus, the HMD rotation has to be reoriented to a forward-facing view in reference to the virtual vehicle. To achieve this, we availed a user-controlled rotation correction using joystick(DualShock 4 Wireless Controller) buttons. The experimental setup is as shown in figure 1.

During preliminary experiment design, we noted that scenes with elaborate details like houses and road markings, though desirable for context, were incongruent when the VR car collided or deviated from the markings. As a workaround, we reduced scene details and utilized a checkered ground to maintain a metric for movement.

2.2 Game mechanism design

The target of the paper is to evaluate passenger’s engagement with in-car VR elements. The paper proposes gaming in a car as a pass-time activity that will improve engagement and offer entertainment. A 3D driving course was designed in Unity 3D, featuring four tasks: no activity, game task, video task and mixed video-game task. To investigate the impacts of engaging with each activity, we included pop-up objects randomly spawned in the scene. The user is required to press a specific joystick button to acknowledge recognition. The sample scene for the different tasks is shown in figure 2.

In the experiment, the order of game and video task were selected random while as no-task/baseline and mixed task appeared at the beginning and the end, respectively.

2.2.1 No Task

As the name suggests, there was no extra task rendered in this phase. This was meant to get the user accustomed to the 3D scenes, capture the natural response and baseline of physiological responses. Besides the user reaction to pop-up objects, the user was also required to orient the VR to a forward-facing direction using a joystick.

2.2.2 Gaming task

The driver actively engages with elements on the road while monitoring for pop-up objects on the game setup. A controllable paddle object (player) is located a few meters from the cars’ position, clearly visible by the user. The paddle moves with the virtual vehicle and is controllable along the x-axis. As the car moves autonomously, collectible objects are spawned at an interval of 2 seconds. When the controller paddle (Unity 3D game object) collides with the mesh of the spawned object, a score is registered (intercepted) and the contrary for a missed object.

2.2.3 Video Task

We propose to set up 2D video on the system for an effortless transition from video to road monitoring. This was achieved by projecting the contents on the environment instead of having the video player in a cockpit. With this setup, the video is rendered in the upper part of the drive path, leaving
enough window for road monitoring. The setup is meant to reduce the time taken to have eyes on the road. The setup is as shown in figure 2(c).

In this case, the video is rendered on a Unity 3D texture material located 200 meters from the camera position. Since the camera in use is a VR, users head rotations shifted the positioning of the video in the virtual scene. We utilized an information video that required users attention to grasp the content. In this task, no other content was issued apart from pop-up objects.

2.2.4 Mixed (Game and Video) Task
In the mixed task, an entertainment video (music video) was played on the rendering texture described in the Video task. The video was meant to be casual in its utility, allowing the users to engage with a music video as they play a game. In this scene, the users played the game, engaged with video, and scanned the environment for pop-up objects, simultaneously. The setup is as shown in figure 2(d).

2.3 Evaluation parameters
In this experiment, we propose to utilize objective data measures to quantify or describe engagement models. To this end, subjective methods like questionnaires were sparingly utilized, albeit in the design and feedback of experience. As such, no questionnaire findings are reported. We rely on physiological measures, pupil size, and EDA response for inferring engagement.

2.3.1 Physiological measure of engagement
Physiological signals have been applied as objective measures in varying fields and topics like emotion recognition, affective computing, decision-making processes, among others (Čegovnik et al., 2018; Le et al., 2020; Shukla et al., 2019; Zimasa et al., 2019). The pupil radius and gaze information were recorded from FOVE HMD, while palmar EDA was collected from the less dominant hand of the user. EDA is derived from measured skin conductance (SC), converted from the potential difference in the electrodes positioned in the subjects’ skin. A report by (Braithwaite et al., 2015) recommends the location of the signal and sampling rate. From this, the current paper utilized palmar EDA with a sampling rate of 1 Hz. The current experiments were conducted on weekends between 3-6 pm as recommended for EDA to be within the same circadian rhythm (Wang et al., 2019).

From literature, eye pupil size change based on the illumination (direct response) or focus point (accommodative response). Other causes of change include drug use, health issues, among others. Amongst these, direct and accommodative responses are essential to this research. In the setup,
we utilized a global illumination setting to ensure uniformity in lighting. Concerning accommodative response, game and video elements were positioned in the exact location for consistency. Research is replete with the parasympathetic and sympathetic innervation of the eye, where it has been established that pupillary dilation is similarly caused by mental processing and engagement (Henderson et al., 2018; Steinhauser et al., 2004). Our target is to identify differences in pupil size, which are explainable by the engagement task.

2.3.2 Posture and head movement

Body sways have been utilized in rehabilitation research as indicators of instability as applied in studies in virtual spaces (Fransson et al., 2019; Gandolfi et al., 2017). In public/shared spaces however, posture is investigated primarily from a safety-oriented viewpoint as cases on collision with real world are heightened (Mai & Khamis, 2018; Mathis & Khamis, n.d.). Posture and improper body movement have practical safety and comfort concerns in any transportation system when the real world is obstructed as pointed out by various authors (Chang et al., 2021; Ekchian et al., 2016; J. Li et al., 2021; Vibert et al., 2001). In the current setup, we investigated which content encourages posture adjustments and head movements following recommendations from authors in (McGill et al., 2019), (Wada et al., 2010). 3D position of head movements is recorded from HMD and utilized for analysis. The indices of interest are the lateral and transverse directions in reference to car movement. The transverse direction is the HMD y-axis which corresponds to physical cars’ forward direction. In 3D VR, this was registered when participants leaned forward or backward. Lateral movement corresponds to the HMD x-axis and results from head rotations.

2.4 Experiment protocol and participants

The experiment was conducted in a moving car environment, as shown in figure 1 with users experiencing VR contents. GPS readings were parsed to the game engine using serial communication as soon as the location was received. Also, the user had to keep orienting the scene view (using joystick) to a forward-facing direction.

Students comprised the participants in the study and were recruited following approval from the relevant ethics committee. Fifteen subjects (9 male and 6 female) took part in the experiment (average age = 25.6, SD = 6.4). Real-life driving or gaming experience was not considered in the current study. A preparatory scene was presented prior to the recording of data. In the preparatory test, the subjects were introduced to the controls buttons and the general objective of the experiment. No incentives were offered to the subjects. The volunteering subjects were instructed to stop the experiment in case of any motion sickness effect or any ensuing complications.

Each subject’s gaze information, button presses, scores, and interaction with game elements are logged in an excel file for further processing. Each session lasted between 10-15 minutes for all subjects. Data analysis was performed using Matlab® software.

3. RESULTS

3.1 Reaction time

The following section describes the results obtained from the experiment. The reaction time associated with individual tasks is shown in figure 3, with * indicating groups with statistically significant differences (p <= .05). Table 1 gives descriptive statistics and the results of the tests performed. From the table, no-task had the best reaction time (mean = 1.06 sec, SD = 0.86) followed by video-task (mean = 1.39 sec, SD = 1), mixed-task (mean = 1.48 sec., SD = 0.94) and game-task (mean = 1.49 sec., SD = 0.88), respectively. The interquartile range (IQR) of the no-task was considerably compact as the users did not have any additional tasks. The remaining tasks had a comparatively similar recognition time as indicated by the box plot and quartiles. Of the three tasks, the video task had a lower reaction rate owing to the hands not performing any immediate operation compared to the game task.

Figure 3. Popup object reaction time of all subjects

<table>
<thead>
<tr>
<th>Task</th>
<th>Mean (s)</th>
<th>Median (s)</th>
<th>Std. Deviation</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1.06</td>
<td>0.80</td>
<td>0.86</td>
<td>0.31</td>
</tr>
<tr>
<td>Game</td>
<td>1.49</td>
<td>1.19</td>
<td>0.88</td>
<td>0.64</td>
</tr>
<tr>
<td>Video</td>
<td>1.39</td>
<td>0.93</td>
<td>1.00</td>
<td>0.99</td>
</tr>
<tr>
<td>Mixed</td>
<td>1.48</td>
<td>1.02</td>
<td>0.94</td>
<td>1.11</td>
</tr>
</tbody>
</table>

One-way ANOVA

<table>
<thead>
<tr>
<th>Source</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean squares</th>
<th>F</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columns</td>
<td>6,988</td>
<td>3</td>
<td>2.32918</td>
<td>2.75</td>
<td>0.0437</td>
</tr>
<tr>
<td>Error</td>
<td>193.244</td>
<td>228</td>
<td>0.84756</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>200.223</td>
<td>231</td>
<td>0.00000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fisher’s LSD post-hoc test

<table>
<thead>
<tr>
<th>Group comparison</th>
<th>Mean Estimation</th>
<th>95% Confidence Interval</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>low</td>
<td>high</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>-0.43</td>
<td>-0.77</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>-0.33</td>
<td>-0.66</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-0.41</td>
<td>-0.75</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>0.10</td>
<td>-0.23</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>0.01</td>
<td>-0.32</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>-0.09</td>
<td>-0.43</td>
</tr>
</tbody>
</table>

NB: Group 1 – No task, 2 – Game, 3 – Video, 4 – Mixed

Table 1. Descriptive statistics and test results performed between group means.

One way ANOVA was performed on the data as shown in Table 1. There was a statistically significant difference...
between groups as shown with \( F(3,231) = 2.75, \ p = .0437 \). A Fishers’ (least significant difference) post-hoc test revealed that the means of game and mixed task reaction time were statistically significantly different (\( p = .0126 \) and \( p = .016 \), respectively) compared to no-task, as highlighted in figure 3. As shown in Table 1, the video task had no significant difference from the no-task (\( p = .06 \)).

3.2 Engagement model from physiological signals

The experiment relied on physiological signals to infer engagement, i.e., pupil size variation and EDA. Figure 4 shows sample data of pupil radius and palmar EDA readings recorded in the experiment. In the figure, colored patches represent each of the tasks under evaluation. In pupil radius, a moving average filter has been applied for visualizing the resultant signal. Each of these signals is analyzed in the next section for all the participants.

3.3 Posture (Head movement)

In VR usage, where the user cannot correctly tell the vehicle’s orientation from visual cues, it may cause disorientation. Figure 6 shows the histogram of head movement in different tasks. An overlaid boxplot gives additional information as per the distribution of occurrences. From the results, when subjects were not engaged in no-task, there were widespread lateral deviations compared to when there was an engagement task. From the histograms, the use of the video task encouraged postural adjustment as witnessed by a more compact interquartile ranges (IQR) and peaks around the center, as shown in Figure 6(a). Lateral IQR were 0.069, 0.014, 0.012, and 0.007, for no-task, game, video, and mixed tasks, respectively. Transverse IQR were 0.005, 0.008, 0.0007, and 0.0003, for no-task, game, video, and mixed tasks, respectively. Also, there was reduced backward-forward head movement, represented by a left skew in video and mixed task of Figure 6(b).

4. DISCUSSION

The paper has explored different engagement modalities that are applicable in a car VR experience. We designed four scenes to be experienced in an actual driving session and utilized pop-up events to investigate the impacts of recognition with each of the tasks. Specifically, this study was concerned with the content design, description of engagement (reaction time of pop-up events and physiological measures), and posture and head movements. Each of these focus areas is discussed below.

Several recommendations have been made towards in-car VR. Scenes with minimal or no incongruences are highly recommended. In the game design, we found that highly detailed scenes with obstacles (buildings, trees) colliding with the users had an unsettling effect similar to those reported in (Paredes et al., 2018). In the experiment design, the research employed GPS sensors for localization, i.e., to map and navigate the virtual car. The GPS sensors had a positioning error of ±2 M from the actual vehicle position. The actual environment in Gifu university featured narrow roads and many physical structures, thereby increasing chances that there will be localization problems. At the preliminary testing stage, incongruences emanating from deviation from the main road or disorientation were verbally pointed out as unsettling by participants. The rotation and position errors around road markings introduced further incongruences.
To mitigate this, we reduced terrain details (building and road markings) and used checkered ground. In further and or subsequent experiment, a real-time kinematic (RTK) GPS is recommended to improve on positioning accuracy.

From practice, motion sickness in real cars is thought to be reduced when participants focus on an outside environment, in which case, the visual and vestibular disconnect is reduced (Koch et al., 2018; Leung & Hon, 2019). In the design of the scene, we capitalized on an environment-centric design to emulate real-world practices. This was achieved by taking the engagement tasks outside the car, ensuring that the user is aware of external movements, though virtual, as he/she engages with entertainment task, simultaneously.

Disengagement has been identified as a course of the present and future challenges to be facing autonomous drivers (Lin et al., 2018; Steinberger et al., 2017). The research explored in-car VR performing different tasks as a proof-of-concept, applicable as an alternative albeit richer infotainment system in the much-anticipated advent of ADS. From figure 3, the users engaging with the no-task scene had a high recognition rate, which is desirable in a road monitoring situation. However, sustained distraction-free vigilance is hard to maintain and is bound to manifest fatigue faster in this state as compared to other conditions (Bench & Lench, 2013; Thiffault & Bergeron, 2003). Thus, a more feasible mode of engagement is needed to offer a tradeoff in vigilance, as suggested in the current research.

In the design of experiment, four tasks are considered ranging from no task to mixed task. The choice was decided with consideration of what is most feasible in a moving car environment. In this case, tasks requiring excessive bodily movement was discounted as unnatural while maintaining eyes on the road for monitoring. In the current evaluation scheme, learning effects (improvement in performance as users get used to similar tasks) that can potentially appear in later scenes was not a major point of concern since there were no task-wise performance indices like game score. VR fatigue, which is a form of scene adaptation, has been reported to manifest in continued usage (Wang et al., 2019). Primarily, fatigue would feature as a decline in stimulation as the scenes progress, which was not observed as shown in physiological measures.

The average reaction time performance of subjects evaluated in tasks were as follows; no-task had the quickest reaction (mean = 1.06 sec., SD = 0.86), video-task (mean = 1.38 sec, SD = 1), mixed-task (mean = 1.48 sec., SD = 0.94) and game-task (mean = 1.49 sec., SD = 0.88). One-way ANOVA found significant difference in the means of the groups ($F(3,233) = 2.75, p = .0437$). Further, Fishers’ (LSD) post-hoc testing indicated that game and mixed tasks were significant ($p = .0126$ and $p = .016$, respectively) compared to no-task. This was expected, as the hands actively control a game, and therefore reaction time is slightly delayed. From table 1, no-task had an average mean difference of 0.43 and 0.41 s. for the game and mixed tasks, respectively. Video task had similar performance with no task.

Overall, the average results of the experiment showed that the difference in reaction time was less than 1 second for all tasks. From this, we concluded that the inclusion of additional tasks did not negatively affect hazard recognition. This is in agreement with the same findings we identified in a related study reported in (Muguro et al., 2021). Research on take-over time, which hazard recognition points to, has indicated that cognitive- and visual-loaded secondary tasks do not have major impacts on take-over performance (Happee et al., 2017; Zeeb et al., 2016). Time to first reaction, time to hands on steering, time to eyes on the road, amongst others is what was identified as determinants of the quality of take-over by researchers (Doubek et al., 2020; Lotz et al., 2019; Wu et al., 2021). Following from this, the design of hand-interactivity between user and the infotainment system should be carefully considered. In addition, a system that allows for eyes-on-the-road, i.e., an environment-centric infotainment system, is much preferred and feasible as discussed herein.

Figure 5 shows average pupil and EDA activity; no-task reported the least engagement while as game and video tasks exhibited highest engagement. The experiment was conducted within 10-15 minutes of use. Therefore, the engagement model may vary in continued usage; however, if the initial parameters are to be maintained, a method of capturing the user’s attention can be re-introduced depending on the current engagement levels. Thus, the setup avails an indirect measure of engagement during road monitoring session. This way, vigilance can be supported for extended periods. More details on this have been discussed in (Muguro et al., 2021). In this regard, the results support the idea that ADS-human handover would be enhanced through the identified driver engagement levels, where the system only performs handover when the user is demonstrably vigilant.

The research also sought to understand postural and head movement since the HMD completely occluded the physical environment. This was done using lateral and transverse head movements. No-task engagement mode had higher lateral
and transverse head movements, as shown in Figure 6(a) & (b). This may present as a challenge in VR usage as excessive head movements and sways have been linked with car/simulation sickness (Wada, 2017; Wada et al., 2010). Figure 6(b) shows that no-task had almost the same transverse head movements as a game-task. The game is designed to spawn objects a few meters (20 m) ahead, and the user was to scan the environment to guide a paddle to collect the objects. This way, deviations on the x-axis are expected to be high in a game task (as is the case) but not in no-task engagement. The other possible explanation for this would be on the user’s undirected gazes wandering into the scene and terrain, associated with boredom and loss of engagement (Bench & Lench, 2013; Bixler & D’Mello, 2016).

A striking difference is noted in a game-compared to a mixed-task, which primarily had the same game setup. In the transverse direction, users moved the head far lesser distance in a mixed environment than in a game setup in attaining the same goal. There was more than 10X increment in interquartile range between no-task and game/mixed task. In a mixed task, users exhibited far less lateral and transverse movements due to the video scene being overcast in the scene. This was thought to have given the user a contextualization of their posture since the video positioning (y-axis position and x-axis rotation) was affected by the head turns and rotations. The effect is visible in Figure 6(b), where the users had a symmetrical distribution around zero in mixed and video tasks compared to the game task. From this, the video texture that followed head rotation and positions was thought to avail contextual feedback of users’ positioning in 3D space, thereby improving posture significantly. A similar effect can be obtained using floating objects or avatars. In-car VR posture and contextual information should be investigated further to enhance simulator sickness and its effects on road monitoring.

The scene design strategies, gathered from previous research and practice, were employed in the current experiment. This includes terrain design to reduce vertigo, eyes-on-road setup, amongst others. The effectiveness of the motion sickness reduction strategies applied in the present study cannot be fully confirmed because we did not investigate motion sickness issues which was beyond the scope of our study. Further investigation and analysis are required. From the completed subjects’ verbal assessment, there was no, or reduced discomfort induced by the content apart from one subject (who was reported to be highly susceptible to car motion sickness) who could not complete the experiment. Further experimentation is needed to verify the utility of the current setup in comparison with alternatives. The contents in use should also be expanded to get a comprehensive view of engagement and its role in hampering motion sickness effects.

In summary, the study sought to investigate user’s hazard recognition time, assessment of user engagement levels, and the influence of content design on posture in virtual space. The proposed environment-centric infotainment design would enhance road-hazard monitoring. The current game design can be replaced with a see-through VR headset, where the popup objects are the actual potential hazards (vehicles and other obstacles) that the ADS is uncertain about or would want the user to be aware of, in a bid to increase driving environment situational awareness. The proposed setup was demonstrated to sustain vigilance as well as avail a mechanism for quantification of user engagement, all of which find direct export in an ADS-human handover process.

The current investigation had several limitations, chief of which was sample test subjects and a limited test scenario. The present paper investigated four tasks, game, video, and a combination of these, and compared that to baseline, no-task. The activities are not exhaustive with a bias towards entertainment. Further tests and investigations are still needed to fully understand the dynamics of experiences targeting entertainment and other office work. At present, when users fail to hit a target owing to physical car movement and rotations, the failure was attributed to game-hardship setting, and therefore an acceptable loss. Other serious tasks would require higher accuracy, minimal scene movements, redesign of user interfaces, to name a few. Besides this, a sample population featuring students is similarly limiting. Further investigation will be performed to capture a greater audience with different demographics.

As a future endeavor, vehicle-centric infotainment design, where the engagement content is displayed in a 2D screen as described in (Elliott et al., 2019), will be compared with the current environment-centric design. Other investigation points identified include but not limited to the use of RTK GPS to increase localization, VR contents and the emergent motion sickness effects, take-over performance while using the described secondary tasks, among others. Last but not least, we note that the proposed scheme can seamlessly be incorporated in trains and airplane infotainment during transit, achieving the discussed objectives and more.

5. CONCLUSIONS

The paper investigated the use of in-car VR as an entertainment modality. This is in anticipation of what the ADS users will be engaged in as an autonomous drive takes effect. The paper focused on an in-car VR content design and the effects this has on engagement levels as inferred from physiological measures as well as hazard recognition time. The final design employed a variation of VR scenes (no-task, game-task, video-task, and mixed (video & game) -task), experienced in a moving vehicle. Analysis was done using reaction time and physiological measures (eye pupil size and electrodermal activity) to infer engagement levels.

From the results, no-task had the highest reaction time, followed by video task, game task, and mixed task. Overall, the experiment confirmed that the difference in reaction time was less than 1 second for all tasks, suggesting that the usage of the proposed secondary tasks did not negatively affect the recognition of hazardous driving events on the road. As seen from pupil size variation and EDA activity measurements, extra tasks were desirable and increased engagement levels linearly. In addition, virtual content affected bodily posture with emphasis on head motion. This was thought to emanate from a contextualization of the users orientation in the virtual space. Similar postural effects can be attained by use of floating objects or avatar systems.

The paper discussed potential merits that would accrue using the proposed an environment-centric infotainment design. These include enhanced road-hazard monitoring, sustained vigilance during transit as well as a mechanism for quantification of user engagement. Optimization of these operations would potentially increase performance of ADS-to-human handover process. Towards the realizability of the proposal, the current design can be replaced with a see-through VR headset, and the popup objects are replaced with augmented driving scenes (vehicles and other obstacles) to derive situational awareness increments discussed herein.

The activities evaluated herein, with a bias towards entertainment, are not exhaustive and further experiments are needed to investigate different designs and the role in hampering or otherwise, the effects of motion sickness, amongst others. Further tests and investigations are still needed to understand the dynamics of VR use in a vehicle. The investigation should further be expanded to multiple forms of transit including trains and airplanes.


