Increase Driving Situation Awareness and In-vehicle Gesture-based Menu Navigation Accuracy with Heads-Up Display

Yusheng Cao

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Computer Science and Application

Myounghoon Jeon, Chair
Scott Mccrickard
Sang Won Lee

April 17, 2023
Blacksburg, Virginia

Keywords: Gesture interaction, In vehicle, Heads up display, Auditory Display

Increase Driving Situation Awareness and In-vehicle Gesture-based Menu Navigation Accuracy with Heads-Up Display

Yusheng Cao

(ABSTRACT)

More and more novel functions are being integrated into the vehicle infotainment system to allow individuals to perform secondary tasks with high accuracy and low accident risks. Mid-air gesture interactions are one of them. This thesis designed and tested a novel interface to solve a specific issue caused by this method of interaction: visual distraction within the car. In this study, a Heads-Up Display (HUD) was integrated with a gesture-based menu navigation system to allow drivers to see menu selections without looking away from the road. An experiment was conducted to investigate the potential of this system in improving drivers' driving performance, situation awareness, and gesture interactions. The thesis recruited 24 participants to test the system. Participants provided subjective feedback about using the system and objective performance data. This thesis found that HUD significantly outperformed the Heads-Down Display (HDD) in participants' preference, perceived workload, level 1 situation awareness, and secondary-task performance. However, to achieve this, the participants compensated by having poor driving performance and relatively longer visual distraction. This thesis will provide directions for future research and improve the overall user experience while the driver interacts with the in-vehicle gesture interaction system.

Increase Driving Situation Awareness and In-vehicle Gesture-based Menu Navigation Accuracy with Heads-Up Display

Yusheng Cao

(GENERAL AUDIENCE ABSTRACT)

Driving is becoming one of the essential daily activities. Unless a fully autonomous vehicle is made, driving will remain as the primary task when operating the vehicle. However, to improve the overall experience during traveling, drivers are also required to perform secondary tasks such as changing the AC, switching the music, navigating the map, and other functions. Nevertheless, car accidents may happen when drivers are performing secondary tasks because those tasks are considered a distraction from the primary task, which is driving safely. Many novel interaction methods have been implemented in a modern car, such as touch screen interaction, voice interaction, etc. This thesis introduces a new gesture interaction system that allows the user to use mid-air gestures to navigate through the secondary task menus. To further avoid visual distraction caused by the system, the gesture interaction system integrated a head-up display (HUD) to allow the user to see visual feedback on their front windshield. The HUD will let the driver use the system without looking in the other directions and keep peripheral vision on the road. The experiment recruited 24 participants to test the system. Each participant provided subjective feedback about their workload, experience, and preference. In the experiment, driving simulator was used to collect their driving performance. The eye tracker glasses were used to collect eye gaze data, and the gesture menu system was used to collect gesture system performance. This thesis expects four key factors to affect the user experience: HUD vs. Heads-Down Display (visual feedback types), with sound feedback vs. without sound feedback. Results showed that HUD helped the driver perform the secondary task faster, understand the current situation better, and reduce workload. Most of the participants preferred using the HUD over using HDD. However, there are some compensations that drivers needed to make if they use HUD: focusing on the HUD for more time while performing secondary tasks and having poor driving performance. By analyzing result data, this thesis provides a direction for conducting HUD or in-vehicle gesture interaction research and improving the users' performance and overall experience.

Contents

List of Figures

List of Tables

List of Abbreviations

HDD	Heads-Down Display
HUD	Head-up Display
SA	Situation Awareness
SAGAT	Situation Awareness Global Assessment Technique
SUS	System Usability Scale

Heads-Down Display for in-vehicle use, referring to those monitor/displays which need the driver to move their visual direction away from the forward road way and look down to see it. All modern vehicle dashboard and infotainment system are considered to be the implementations which used Heads-Down Display.

Head-up Display for in-vehicle use refers to those devices that project the information onto the front windshield. The project area remains partially transparent, allowing the user to see the information and, in the meantime, see through it to pay attention on the road simultaneously.

System Usability Scale is a reliable tool for measuring the system usability. It will help with analysing and evaluating the product and service including hardware and software.

The Situation Awareness Global Assessment Technique is a query technique that was developed by Endsley. SAGAT is based on information-processing theory. SAGAT is one of the best publicized and most widely known measure of SA.[12]

Chapter 1

Introduction

1.1 Research Background

Because car accidents are one of the leading causes of death, much research has been produced to provide solutions to reduce potential accidents worldwide. The European Commission and National Highway Traffic Safety Administration have suggested that in-vehicle interfaces, including navigation displays and media players, may be one reason drivers look away from the forward roadway [9]. In addition to the off-road glancing, movements that interfere with driving, such as reaching out to a function button or touch screen, also have the potential to cause accidents.

In the past decade, gesture-based interactions have been introduced to the in-vehicle infotainment system to reduce unnecessary body movement while the driver performs secondary tasks (e.g., control AC temperature, perform navigation, control radio and music). Ample amount of research has suggested that gesture-based systems have lower levels of distraction and unnecessary movement compared to traditional touch-based interactions. Additionally, research has shown that gesture-based systems also decrease drivers' workload and the number of glances toward the interface [15][42].

However, the multiple resources theory has indicated that when people are required to utilize the same resource to carry out dual-tasks, the performance will be significantly degraded [51]. While carrying out both driving and using the gesture-based interactions, a driver will need to rely upon a manual resource and a visual resource, which means the driving performance may be degraded depending on the difficulty of the gesture selection task.

Moreover, research has demonstrated that auditory feedback has been effective for reducing driver distraction [38]. Using non-visual modalities can mitigate the risk of accidents by limiting the competition for visual attention for dual tasks [30]. Sterkenburg et al. [42] showed that the gesture-based interface combined with well-designed auditory feedback can improve drivers' performance for both driving and menu selection tasks. However, Sterkenburg et al. [42] also stated that auditory feedback alone is not enough to avoid all unnecessary glances away from the road. The drivers cannot stop themselves from seeing the menu on the monitor because they might not fully trust the auditory-only display. Thus, they instinctively seek a visual aid in such circumstances [42]. Furthermore, Shakeri et al. [38] suggested that visual scanning is one of the sources of dual task interference. The division of the visual attention to the road and the visual feedback for gesture interactions will be a factor, which will make qualitative distinctions because they are not placed within the useful field of view (UFOV) (4 - 30 degree) but 'head field' (>30 degree) instead [38]. With this type of setup, the user will not need to move their head to search for the interest point (visual aid for the gesture interactions). Instead, they only need to switch their eye focus point between the HUD and the road.

Based on Sterkenburg et al. [42] and Shakeri et al. [38] suggestions, the present study combined auditory feedback with a novel visual approach to improve the overall accuracy of the system and user experience in addition to eliminating unnecessary glances. The Heads-Up Display (HUD) has been introduced to vehicles in recent decades and have shown great potential providing information efficiently while allowing drivers to keep their gaze on the road [32]. Compared to the traditional infotainment display position, HUD will keep the driver's attention on the road and provide enough information and visual assistance to operate the gesture system [32]. In this thesis, a HUD will be implemented to provide a visual aid to assist the driver in carrying out gesture interactions. When the gesture system is activated, the HUD will activate and shows gesture menu-related information. The emergency warning sign will be able to interrupt the gesture selecting process when there is a hazard ahead. The HUD is placed in foveal vision (<4 °visual angle) with the road position, which will create a significant distinction from the head down display (HDD) (> 30°) [38].

In this thesis experiment, all the participants used this gesture system with or without HUD and with or with out auditory feedback during driving on the simulator. Using the gesture system, I collected participants' data on the overall accuracy and time consumed navigating the menu system. I also collected participants' data on driving performance and eye glance from the driving simulator. To investigate whether the HUD can eliminate most of the unnecessary visual distractions, on-road hazardous events were set up during the experiment to evaluate participants' situation awareness.

Chapter 2

Related Work

2.1 In-vehicle Gesture Interaction

Gesture interaction has become one of the new input methods for in-vehicle interactions. Researchers have put a great effort into investigating gesture interaction in the following three research areas. First, related to the current research, researchers investigated the best hand gesture for the driver to interact with the in-vehicle functions, both primary [52] [24] and secondary tasks [26] [42].

There are three ways of designing in-vehicle gesture interaction systems. The first research direction is 'direct mapping with functions,' meaning each gesture represents a specific vehicle function [26]. For example, drawing a circle in mid-air using the index finger could signal an increase or decrease in the music volume. Most in-production BMW vehicles are equipped with a gesture interaction system based on this design [36].

The second approach used the traditional graphical user interface design guidelines, employing the 'WIMP' (window, icon, menu, and pointer) paradigm. Based on Jacob et al.

[19], they could make the window icon and menu remain the same (they can be visible on HDD or HUD), but instead of using the touch screen or mouse as a pointer, they chose to use a mid-air hand position with respect to the hand detector, to work as pointer input to interact with the user interface [42]. This gesture design used a set of reprogrammed gestures to let the user navigate through the interface. However, this type of gesture system design has one major draw back. Sterkenburg et al. [42] showed that compared to other conditions (auditory feedback only, no audio, visual only), a combination of visual feedback and auditory feedback showed higher workload and less eyes-on-road time. Despite those major drawbacks, it showed a faster secondary task completion times, and was preferred by the participants. This implies that participants prefer more visual feedback when it is available [42].

Researchers have investigated the best way to improve the overall performance of the in-vehicle gesture interactions by implementing additional feedback. So far, the studies have focused on finding the best feedback method: auditory feedback [42] or haptic feedback [37]. However, little research has attempted to use visual feedback to improve the overall performance and user experience.

Overall, based on the current literature review, hand-gesture interactions with properly designed feedback can reduce unnecessary glances away from the road, which are considered to be removing a hazard for driving, compared to the traditional touch screen interface [26]. Furthermore, although Sterkenburg et al. [42] suggested that using the visual/auditory combination feedback has the highest workload for in-vehicle gesture interaction, this type of feedback has the shortest selection time, which means less distraction from driving. It is worthwhile to conduct further investigation into visual/auditory feedback set up because previous work in this area used a non-efficient visual feedback design (the visual feedback placed at the right bottom (HDD) and not aligned to the road direction (HUD)).

2.2 Heads-Up Display (HUD)

The HUD has been developed over decades as a useful in-vehicle information system. Compared to traditional Heads-Down Displays, it shows a significant performance improvement, causing less visual distraction and enabling smooth transitions between the incoming traffic and the information provided [14, 22, 22, 31, 39]. Research shows that the HUD helps drivers to respond more quickly to unanticipated road events under both low and high driving loads [21].

According to Beck et al., eleven high-priority information items are displayed in the HUD [2]. These can be classified into hazard warnings, traffic notifications, and information about secondary tasks. There has been research about vehicle information [29][11], traffic warnings such as lead vehicle warnings, [35][1][6][34] and pedestrian warnings [10][34] for the hazard warning. There is also much research about situation awareness, such as traffic sign warnings [10][11] and lane departure warnings [6][34]. Furthermore, there is some research about driving instructions, such as navigation instructions [6][34][7] and drivers' state alert [3]. As for the position of the HUD, Tippey et al. presented a wearable HUD [44], but the other studies' design chose to place the HUD on the front windshield. In literature, drivers have only passively reacted to the information presented on the HUD, but they have not actively interacted with it.

Shakeri et al. [38] stated that visual scanning is one of the sources of dual-task interference. Taking Sterkenburg et al. [42]'s research as an example, there are two visual channels in their research. One is on the road for driving while the other is the visual feedback of the gesture system. Those two visual channels are placed separately so that they are larger than the driver's useful field of view (UFOV) (4 - 30°). This means that the driver needs to rotate their head to switch between those two visual channels. Without a doubt, this experiment design impacts the overall user experience because the experiment participants cannot put their peripheral vision on the road while seeing the visual feedback simultaneously.

Based on the existing work in this area, this thesis will use the HUD as a more advanced source of visual feedback for the in-vehicle gesture interaction system. This is because drivers have been found to usually seek visual feedback for the gesture interactions [40] than auditory displays during the gesture interaction process. Due to driver's preference, using HUD will provide information more efficiently by reducing the attention transfer time from the road to the source of information. Moreover, because the HUD can provide more complex information (e.g., navigation), this thesis postulates that the HUD will be possible to provide necessary visual feedback for the gesture interaction system. According to Shakeri et al. [38], HUD can put two visual sources within the drivers' useful field of view so that drivers will only need to switch their visual focus point while performing driving and min-air gestures at the same time. Ideally, using HUD is expected to make the whole process easier and safer.

2.3 Auditory Display

In addition to HUD, the second feedback source used in this experiment was the auditory display. The researchers have focused on finding the best type of auditory displays to increase the overall system performance and driving safety. A previous experiment compared three auditory display types in combination with the gesture-based menu navigation system for in-vehicle use. The three types of auditory display tested were Auditory Icon, Earcon, and Spearcon.

Auditory displays can have similar effects to that provided by the HUD, in terms of reducing eyes-off road time and encouraging the driver to focus more on the road [40]. Jason et al. came to the conclusion that the auditory-only menu system required drivers more time to navigate [40]. The time-sensitivity nature of driving made drivers more eager to seek other types of information as well as audio that would help them carry out a given task.

Considering that no similarly designed product existed and was widely used by the public, this thesis assumes that all the participants recruited had no experience using the gesture-base system. Moreover, according to Lucas [23], both auditory icons and earcons have some learning effect, which means that the user shows a different performance when they fully understand and memorize what each auditory cue means. Therefore, this thesis will consider text-to-speech and spearcon as potential auditory display types based on the learnability factor. However, since driving is a time-sensitivity work, text-to-speech is not considered because it takes a relatively longer time than spearcon. Using spearcon allowed the user to get more condensed information from the system without the risk of misunderstanding and save time. Finally, during the pilot study, the participants' feedback suggested that spearcon is the most intuitive to use and provides the most accurate auditory feedback. Compared to auditory icons and earcons, spearcon led to significantly better navigation efficiency, accuracy and learning rate [48]. Thus, based on the previous literature review, this thesis used spearcons as the choice of the auditory feedback.

2.4 In-vehicle situation awareness

Situation awareness (SA) is one of the fundamental concepts for maintaining safety in a dynamic task environment. Endsley stated that SA is how human operators develop and maintain a complete understanding of their surrounding environment. [12]. As for driving, drivers are given more specific tasks regarding their SA: The drivers need to know the vehicle's current position, distance to potential hazards, relative behavior to other vehicles or potential risks, and precise predictions of the possible changes to all the variables mentioned before [17]. To be more specific, Matthews et al. categorized SA as spatial identity, temporal, goal, and system awareness [25]. However, regardless of how complicated the definition for the in-vehicle SA is, it can be interpreted as the driver will keep locating the potential hazard during driving and respond to it properly in the given environment. Various research projects have shown different methods to shorten the time for the driver to locate the hazard and respond.

The most recent research in this domain presents its own way of improving the drivers' SA. Gang et al. suggested that using spatialized warning sound will increase drivers' SA [13]. Andreas and Alistair presented their unique approach using HUD to reflect the surrounding environment, which helped the drivers' SA. Calhoun et al. suggested a HUD-like system to increase UAV operators' (which can also be considered as the drivers) SA by highlighting the interest area with a high density information (highlighting color and text labels) to help them assess the environment correctly and quickly [5]. The findings of McDonald [27] research showed that the HUD has great potential in improving users' SA by making the warning signs more salient to drivers. Moreover, Wang et al. [49] compared the HUD with the dashboard and the result shows that HUD has significantly higher SAGAT result.

Moreover Sterkenburg et al. [41] have attempted to show that although using the different auditory displays improves the gesture interaction performance, it still has not shown improvement in accuracy and caused visual distraction when visual feedback was added. [41]. Yang et al. presented a good way of improving the drivers' SA [53] by placing led light under the windshield, using the led to highlight the hazard direction. But it may not help gesture interactions since it needs more complicated and high information density visual feedback. Despite this, this research shows that the specially designed visual feedback improved the driver's SA. In conclusion, HUD may have a great potential of solving problems mentioned before, while also beholding the possibility to increase drivers' SA. Based on this background, this experiment will implement a gesture interaction system, adopting HUD as a source of virtual feedback to improve drivers' SA.

Chapter 3

Experiment

3.1 Research Questions

This thesis will attempt to answer the following questions: If an appropriate visual aid is provided, can the driver maintain a safer drive while carrying out the secondary tasks? If an appropriate visual aid is provided, can the driver react better to incoming hazards while carrying out the gesture-based secondary tasks? Finally, can auditory feedback improve the gesture interaction performance with the HUD? In this regard, we present the following research questions:

RQ1. Between Heads Up Display and Heads Down Display, which source of visual feedback makes the participant show better performance in driving and in-vehicle gesture interactions for a secondary menu navigation task?

RQ2. Does the HUD increase the driver's situation awareness while operating a gesture-based interaction system to help avoid potential accidents?

RQ3. Between with auditory display and without auditory display, which condition makes the participant show better performance in driving and in-vehicle gesture interactions for a secondary menu navigation task?

RQ4. Does the visual and auditory feedback have any interaction effects on driving performance and gesture interaction performance?

3.2 Experiment Setup

3.2.1 System Design

Gesture Interaction Design

The gesture system in thesis was designed based on Sterkenburg et al. [42]'s design. The system contained 4 different gestures to allow users to use to navigate through the system. Activate gesture: the participant needs to grab a fist to activate the system. Selecting gesture: The participant needs to hover their hand horizontally or vertically to select menu item. Without activating the system, hovering around the menu will not have any visual or auditory feedback. Switch menu gesture: the participant can perform a swiping gesture to switch the menu groups. The menu contained three layers, four items on each layer, and twelve menu options in total. Selecting gesture: the participant can perform a tapping gesture to select the menu item.

Visual Feedback Design

The gesture-based in-vehicle menu navigation system provided three user interfaces (UI) on both the HDD and HUD: stand-by, interact, and warning. The stand-by UI had the standard grid of the menu but gray background (Note that this UI is not the thesis' research variable). The Stand-by UI will appear on the HUD when the gesture interaction system is not activated. The gray background on the HUD will make the UI even more transparent so that the participants will see the road easier. The menu grid will remain white, which will help the user have an impression of where the menu is located.

When the participant performs the activate gesture (making a fist), the interact UI will replace the stand-by UI to provide a visual aid in response to the gesture interaction. The gesture interaction system will be activated and the interact UI will show a two-by-two grid where each grid element represents a selection. As the participant hand moves, the UI will highlight the participant's current selection by changing the UI grid menu background into red to indicate where the participant's hand is with respect to the virtual physical position of the UI element. In the meantime, a sound representing the menu item will identify which UI element is highlighted. After the swipe gesture is successfully performed, the UI will be replaced with the new menu option set, and a swipe sound will play. Finally, the participant will perform the select gesture, and a confirmation sound will be played.

The warning UI will monitor the driving scenarios during the participant interacts with the UI or drives. If there is a potential hazard (a pedestrian crossing or an emergency vehicle appearing), the warning UI will replace the current UI to warn the driver to react. If the warning UI replaces the interact UI, the selection procedure will be stopped.

A pre-programmed audio prompt will present the command to let the participant select one random specific item within the menu, and the system will record the participant's performance using the gesture system (accuracy and speed).

Auditory Feedback Design

This thesis used the spearcon generator software [48] to transfer the menu item to spearcon audio files with compression ratio of 70% (software default setting). We kept the selection confirmation sound and swipe sound in both auditory conditions. We also used an auditory icon for the selection confirmation and swipe to differentiate those auditory displays from the spearcons which were designed to help the user locate which menu item they are selecting. An earcon was used for the warning sound for hazard.

3.3 Equipment

A medium-fidelity driving simulator (NADS-MiniSim) was used to simulate the driving scenarios [8]. The simulator was built by three 50 inch Sony displays covering a 120 degree horizontal field of view, a steering wheel, an adjustable seat, pedals for gas and brake, and a stereo speaker. A custom HUD device placed between the display and the steering wheel made by acrylic board with partial reflection coating. An Apple iPad pro 12 inch was placed under the acrylic board as the source of the visual display (Figure 3.1). The participant was able to see the iPad image and at the same time see through it to view the simulated environment. Natural reader [45] online text-to-speech web application was used to generate all the voice commands. A Tobi Pro Glasses 2 [16] - an eye tracker device was used to capture participants' gaze fixation during the study. A Leap-motion hand tracking camera [46] was placed on the side of the simulator for tracking the participant's hand movements. The Heads-Up-Display was handcrafted in the lab. A transparent acrylic board was mounted between the driving station and the monitor in certain angle. A one-way reflective film was placed on the acrylic board to create the HUD-like reflection (drivers can both see through it and the reflection in other direction).

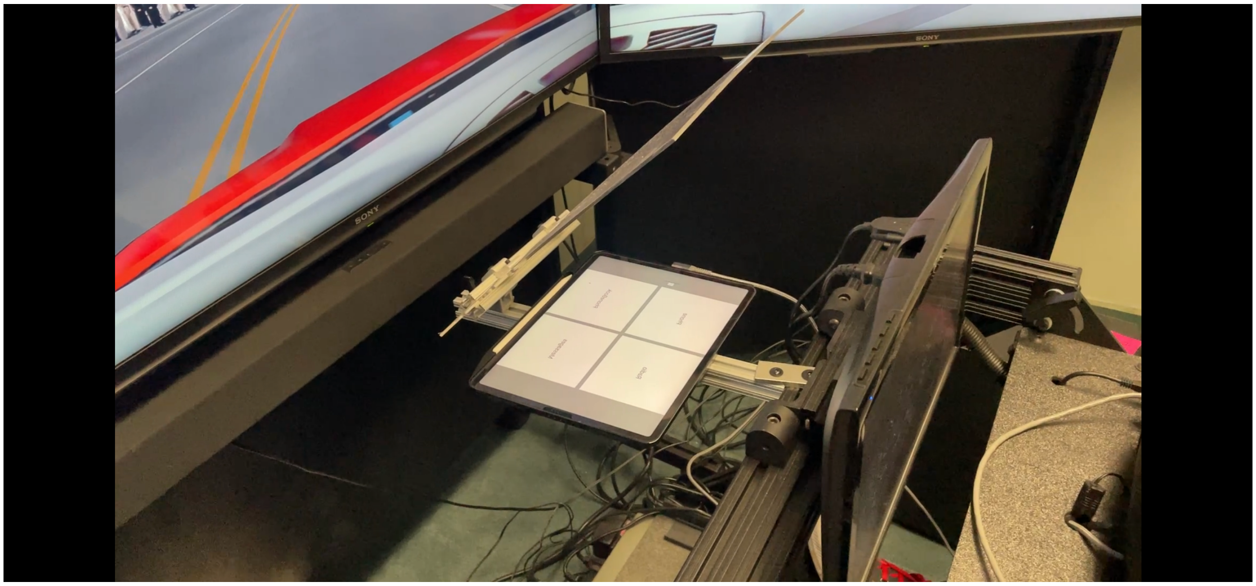


Figure 3.1: The Heads Up Display setup

3.4 Participants

For this experiment, 24 participants were recruited and compensated 10 for their participation. This experiment adopted 2 x 2 repeated measures ANOVA test [43]. This thesis ran a power analysis with the parameters of moderate to average effect size of 0.06, alpha value of 0.05, power value of 0.8, group number of 1, measurements number of 4 and received the participant number(sample size) 23. To ensure a more accessible experimental design, this thesis set the number be 24. After the Institutional Review Board (IRB) approved the experimental protocol, The recruitment form was sent out to Virginia Tech community and nearby neighborhood. Moreover, before each experiment, the participant was asked to give a written consent by signing the consent document. The participants had to fulfill the following pre-requisites to be eligible for the study:

- Participants were required to have a valid driver's license and must be an active driver.

- Participants were required to be an active driver for at least one year

- Participants needed to be between 18 - 45 years old.

- Participants do not have severe motion sickness during driving.

Based on those criteria, a total of 24 participants were recruited for this study. The participants included 7 females and 17 males, were 26 years old on average, held a driver's license for 7 years, drove 84130 miles in total and drove 8.8 times per week.

3.5 Experimental Design

This study adopted 2 (HDD or HUD) (Figure3.2) by 2 (no-sound or with spearcon) within-factors design, which created four parallel experimental conditions. The participants were asked to drive through four different driving scenarios under the following four conditions: with HUD or with HDD (Figure3.2), with or without spearcon sound feedback. Each auditory display representing each menu item was played when the item was ready to be selected. Spearcon was made by compressing the speech of the menu items; in the no-sound condition, the system played no sound to indicate which menu item is currently selected (the select and swipe sounds remained).

Figure 3.2: HUD vs. HDD

During pilot study, after try out for multiple times the test participants shows gradually better performance. The order of the conditions were counterbalanced to minimize the learning effect. Thus every participants experienced the conditions in different orders. Table 3.1 depicts the experimental design.

	scenerio1	scenerio2	scenerio3	scenerio4
p1	HUD with sound	HUD without sound	HDD with sound	HDD without sound
p2	HDD without sound	HUD with sound	HUD without sound	HDD with sound
p3	HDD with sound	HDD without sound	HUD with sound	HUD without sound
p4	HUD without sound	HDD with sound	HDD without sound	HUD with sound
...	...	...	...	...
p24	HDD without sound	HDD with sound	HUD without sound	HUD with sound

Table 3.1: Scenerio Design

While the percipient drives in the simulation under each condition, 15 voice commands asked the participant to select a specific menu item (3 with hazard interruptions, 12 without hazard interruptions). To make the data consistency facilitate analysis process, the experiment used A pre-programmed audio prompt to giving out commands. The pre-programmed order limited the choice of the menu layer. This process ensured that the participant experience the same difficulty in each experimental phase. However, the system randomly chose which specific menu item on the current layer for which to send out the command. Each driving scenario lasted around 6 minutes. Moreover, hazard warning (emergency vehicle ahead/pedestrian crossing) interrupted the participant's driving and selecting process during the driving scenarios (Figure3.3). To ensure the driving data is worth analysis, the participants didn't respond to the hazard in driving (move to other lanes, break/accurate). They were asked to find the location and type of the hazard and reflect them in the SAGAT questionnaire.

Figure 3.3: Caution! pedestrian ahead.

3.5.1 Task

The participants asked to drive safely by following a simulated leading vehicle at a consistent distance in the driving scenarios. An automated command were given out at certain intervals, ask the participant to use the gesture-based menu navigation system to select a specific menu item within the given time. The new command were give out, regardless of participants' selection speed and accuracy when the time run out (30s). In the meantime, the participant have to remain aware of the traffic conditions and react to the hazards.

3.5.2 Procedure

At outside the study, the participant was fill the IRB consent form first. Then, the participant was asked to fill out questionnaires to collect demographic information (age, driving information). Next, the participant completed the first part of the motion sickness measurement questionnaire to create a baseline for subjective measurement of discomfort. After completing a 2 minute simulator sickness scenario with eye-tracker glasses on, the participant was asked to fill out the motion sickness measurement post-questionnaire to determine if the participant has motion sickness. Participant was allowed to proceed with the experiment only if they have no motion sickness. When each participant is qualified for the experiment, they will watch a short tutorial video about using the gesture-based menu navigation system. Then, the participant will have a 5-10 minutes to become familiar with the system. During the experiment, the participant will start with an introduction of the auditory display/visual feedback condition they will experience. The participant will have some time to get familiar with the system equipped with the auditory/visual feedback condition. Once the participant is ready, they will proceed with the driving scenario.

During the driving scenario, they will select menu item based on the command given by the system. The participant will also be asked to maintain the same distance to the leading vehicle in the driving scenario and also react to any potential hazards that appear in the scenario. The driving performance data was collected by the driving simulator, and the menu system will collect the gesture selection performance data. The driving scenario was forced to stop after the last hazard warning appear for 5 second, the individual participants was asked to fill the The Situation Awareness Global Assessment Technique (SAGAT) questions used to measure all three situation awareness levels (perception, comprehension and projection) of the driver. NASA-TLX[18] was also used to record the perceived workload after each scenario. And finally the System Usability Scale (SUS) Questionnaire [4] to measure the usability of the certain condition. After all four conditions were completed, the participant was asked about their preferred auditory/visual feedback conditions and to provide some further comments regarding the overall experience. Finally, the participant was compensated with fifteen dollars. The whole experiment lasted around one and half hours.

3.6 Measurement

The data measurements in this experiment were categorized into two different types: subjective measures and objective measures. Subjective measures were collected by the followings: questionnaires after the participant finished each scenario: System Usability Scale (SUS) Questionnaire [4]. Participants' workload level will be collected through NASA-TLX [18]. A situation awareness questionnaire asked about the hazard encountered during the experiment. The SAGAT questionnaire[12] was used to get the participants' situation awareness about the emergency. The SAGAT asked participants about their element of interest on screen to measure perception. To measure comprehension, the SAGAT asked ”What do these elements tell you about the current situation (are you safe or in danger)?” and ”What is currently happening in the scenario? ” and finally the SAGAT will asked user ”What do you think will happen next?” and ”What do you think you need to do next?” to measure projection. As objective measures, the participant’s eye gaze data was collected by the eye-tracking glasses. The driving simulator will collect the participants’ driving performance data. Detail will be described in the next section.

Chapter 4

Results

A repeated-measures ANOVA [43] was conducted to determine if there were any differences among the four conditions. The gestrue system analyzed the task performance for gesture interaction data by comparing the participants' accuracy and time. This thesis will examine the time spend difference in different focus areas using the eye-tracking data collected (on-road, off-road, HUD, HDD). This thesiswill use the NASA-TLX[18] result to determine the workload for each condition. Data collected by System Usability Scale (SUS) Questionnaire will show the user preference of the system. Data collected by SAGAT questionnaire [12] will determine the level of situation awareness during driving in different conditions.

4.1 Situation Awareness

For level 1 SA, data were extracted from questions 1, 3-5 (appendicesA.1) to make sure participants located the source of the hazard. For level 2 SA, data were extracted from questions 2 and 6 to make sure participants understood the current situation fully. For level 3 SA, data were extracted for the last 2 questions to make sure participants have right predictions of the future. All the correct answers were marked with score 1 and incorrect answers were marked with score 0. Results were analyzed with a 2 (Visual types) x 2 (Sound types) nominal logistic model fit. There was a statistical difference for level 1 SA (Figure 4.1).

Average of Situation Awareness

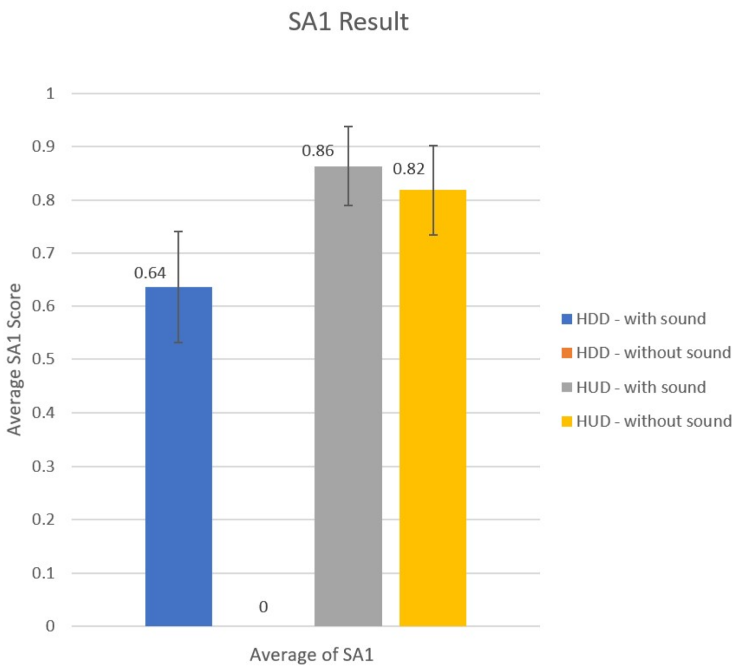


Figure 4.1: Situation Awareness Result

The ANOVA result shows a statistically significant difference between visual feedback types ($F(1, 21) = 56.3909$, $p < .001$, $\eta_p^2 = 0.3543$). HUD ($Mean = 0.84$, $SD = 0.369$) had significantly higher SA1 score compared to HDD ($Mean = 0.318$, $SD = 0.471$). Results also show a statistically significant difference between sound feedback types ($F(1, 21) = 31.7114$, $p < .001$ $\eta_p^2 = 0.1892$). with sound ($Mean = 0.75$, $SD = 0.438$) had significantly higher SA1 score compared to without sound ($Mean = 0.40$, $SD = 0.497$). The interaction between visual feedback and sound was also significant ($F(1, 21) = 17.3122$, $p = 0.0004$, η_p^2

= 0.1492). In addiction to the repeated measures ANOVA, LSD t-test also shows the sound(t-test matrix result = 0.14229) and visual (t-test matrix result = 0.34319) has differences between pairs of means within the overall ANOVA model and they are significantly different. HUD with sound (*Mean* = 0.86, *SD* = 0.3512) had the best performance HUD without sound (*Mean* = 0.8181, *SD* = 0.394) were second best, HDD with sound goes third (*Mean* = 0.63, *SD* = 0.492) and HDD without sound was worst (*Mean* = 0, *SD* = 0, all failed the test).

For the level 2 SA and level 3 SA, the ANOVA result shows that there were no significant differences among the conditions.

4.2 Workload

The NASA-TLX data were analyzed with a 2 (Visual types) x 2 (Sound types) repeated measures ANOVA, which showed the following results.

4.2.1 Mental Demand

A repeated measures ANOVA shows that for the mental demand, visual feedback types had a statistically significant difference ($F(1, 21) = 11.1332$, $p < 0.0031$, $\eta_p^2 = 0.8676$). The result shows that HUD (*Mean* = 43.97, *SD* = 22.608) had lower mental demand than HDD (*Mean* = 59.31, *SD* = 17.27). There were no main effect of sound types ($F(1, 21) = 2.3857$, $p = .2523$) and no interaction effect between visual types and sound types ($F(1, 21) = 1.4734$, $p < 0.2383$).

4.2.2 Physical Demand

By applying repeated measures ANOVA for the physical demand data, in the parameter estimate for Physical data model fit, visual feedback shows ($F(1, 21) = 8.7833$, $p < 0.0064$, $\eta_p^2 = 0.9105$) that it have significant difference. For sound and sound/visual interaction, it has result of $F(1, 21) = 0.2347$, $p = .6331$ and $F(1, 21) = 1.6766$, $p = 0.2094$, which means they have no significant difference. The result shows HUD($mean = 39.65$, $SD = 21.65$) has lower physical demand then HDD($mean = 52.84$, $SD = 21.03$).

4.2.3 Temporal Demand

By applying repeated measure ANOVA for the temporal demand data, in the parameter estimate for temporal demand model fit, visual feedback shows ($F(1, 21) = 8.4723$, $p = 0.0084$) means it have significant difference. for sound and sound/visual interaction, it have ($F(1, 21) = 1.0897$, $p = 0.3088$) and ($F(1, 21) = 0.4641$, $p = 0.5032$) means they have no significant difference. The result shows HUD($mean = 40.34$, $SD = 23.58$) have lower temporal demand then HDD($mean = 55.45$, $SD = 21.45$)

4.2.4 Performance

By applying repeated measure ANOVA for the Performance data, in the parameter estimate for Performance data model fit, visual feedback shows ($F(1, 21) = 36.3359$, $p < 0.0001$, $\eta_p^2 = 0.8961$) means it have significant difference. for sound and sound/visual interaction, it have ($F(1, 21) = 0.5983$, $p = .4479$) and ($F(1, 21) = 0.1794$, $p = 0.6762$) means they have no significant difference. The result shows HUD($mean = 39.659$, $SD = 24.385$) have higher performance demand then HDD($mean = 57.045$, $SD = 23.13$)

4.2.5 Effort

By applying repeated measure ANOVA for the effort data, in the parameter estimate for effort model fit, visual feedback shows ($F(1, 21) = 8.7833$, $p < 0.0064$, $\eta_p^2 = 0.9374$) means it have significant difference. for sound and sound/visual interaction, it have ($F(1, 21) = 0.2347$, $p = .6331$) and ($F(1, 21) = 1.6766$, $p = 0.2094$) means they have no significant difference. The result shows HUD($mean = 51.59$, $SD = 22.97$) have lower Effort than HDD($mean = 62.95$, $SD = 22.13$)

4.2.6 Frustration

By applying repeated measure ANOVA for the frustration data, in the parameter estimate for frustration model fit, visual feedback shows ($F(1, 21) = 5.2083$, $p < 0.0330$, $\eta_p^2 = 0.8790$) means it have significant difference. for sound and sound/visual interaction, it have ($F(1, 21) = 2.2530$, $p = .1482$) and ($F(1, 21) = 3.8344$, $p = 0.0636$) means they have no significant difference. The result shows HUD($mean = 37.04$, $SD = 25.204$) have lower frustration then HDD($mean = 47.95$, $SD = 25.319$)

4.2.7 Overall

By applying repeated measure ANOVA for the overall workload data, in the parameter estimate for moverall workload model fit, visual feedback shows ($F(1, 21) = 14.0468$, $p = 0.0012$, $\eta_p^2 = 0.8418$) means it have significant difference. for sound and sound/visual interaction, it have ($F(1, 21) = 0.9328$, $p = .3451$) and ($F(1, 21) = 2.3314$, $p = 0.1417$) means they have no significant difference.

The result shows HUD($mean = 45.85$, $SD = 19.5$) have lower overall workload than

HDD($mean$ = 60.77, SD = 15.22)

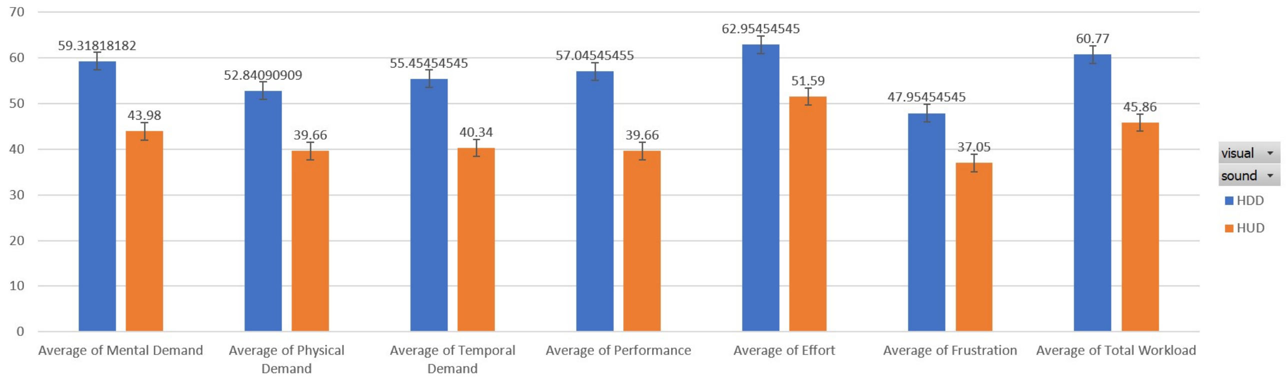


Figure 4.2: Overall Workload Result

In conclusion, there is no significant result shows in sound and sound/visual interaction. For the visual feedback conditions(Figure 4.2), HUD out-perform HDD in every factors of the workload.

4.3 System Usability Scale (SUS)

The participants were ask to fill out the SUS questionnaire for each condition (Appendix A.2). The results were analyzed using a repeated measures ANOVA. In the parameter estimate for SUS model fit, visual types had a statistically significant result $F(1{,}22.65) = 20.0278$, $p = 0.0002$, $\eta_p^2 = 0.8692$. There were no main effect of sound types $F(1{,}20.77) = 1.5473$, $p = 0.2274$ and no interaction effect between visual types and sound types $F(1{,}16.54) = 0.0068$, $p = 0.9351$. The result shows HUD ($mean = 68.85$, $SD = 19.44$) have higher system usability than HDD ($mean = 53.8$, $SD = 19.76$).

4.4 Menu Navigation Performance

Participants' menu navigation performance was collected by the gesture interaction system. Time for each individual selection was recorded. Time between the end of the command and the participant's successful selection of the right menu item was measured. If the participant did not select the right item in 25 seconds, the system considered it ”selection failure” and the record was left blank. Accuracy for each individual scenario was calculated afterwards. The menu navigation time and accuracy data were analyzed using repeated measures ANOVA.

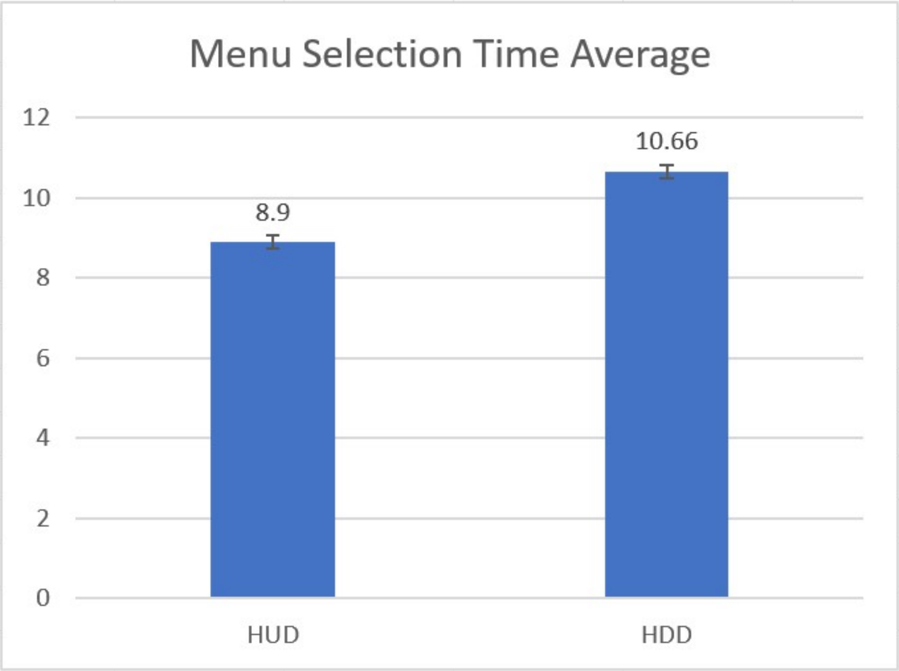


Figure 4.3: Menu Navigation Performance Result

Results showed a statistically significant difference between visual types $F(1{,}23.05) = 8.2027$, $p = 0.0088$, $\eta_p^2 = 0.9652$. The result (Figure 4.3) shows that participants took less average time using HUD ($Mean = 8.9$, $SD = 4$) than using HDD ($Mean = 10.66$, $SD = 5.2$). There were no main effect of sound types $F(1{,}21.14) = 3.728$, $p = 0.067$ and no interaction effect between visual types and sound types $F(1{,}21.8) = 2.3386$, $p = 0.1406$.

4.5 Eye Tracking Data

Tobii eye tracker [16] provided three types of data based on the area of interest (AOI) we defined: Menu and road. Tobii eye tracker provided eye movement types (fixation, saccade, and eyes not found), gaze event duration (ms), and AOI hit point. Based on those data we can categorize AOI hit into: selecting (eye focus on HDD/HUD for selecting, considered as distraction), driving (eye focus on road), or in-between (eyes moved rapidly between the menu and road, considered as distraction). The data were analyzed using repeated measures ANOVA. There were statistically significant differences for selecting and in-between driving and selecting.

4.5.1 Selecting Fixation

By filtering out all the selecting record which eye movement type is fixation and put in repeated measure ANOVA model fit. A repeated measures ANOVA showed a statistically significant difference between visual types $F(1{,}15.91) = 32.0716$, $p < 0.001$, $\eta_p^2 = 0.0603$. However, there were no main effect of sound types ($F(1{,}10.19) = 0.0317$, $p = 0.8622$) and interaction effect between visual types and sound types ($F(1{,}15.7) = 0.1988$, $p = 0.6618$). To get a better comparison, further calculation was made: first the average time for each participants was calculated, and also count the number of record occurred which is both selecting and fixation.(Table B.2) Then use the number of occurred as weight, the weight average time was calculated Based on the result (Figure 4.4), HUD had more fixation time($mean = 229.35$, $SD = 591.94$) compare to HDD($mean = 91.89$, $SD = 200.85$)during selecting menu items. It means that on average, participants stared at the HUD longer than HDD during selection.

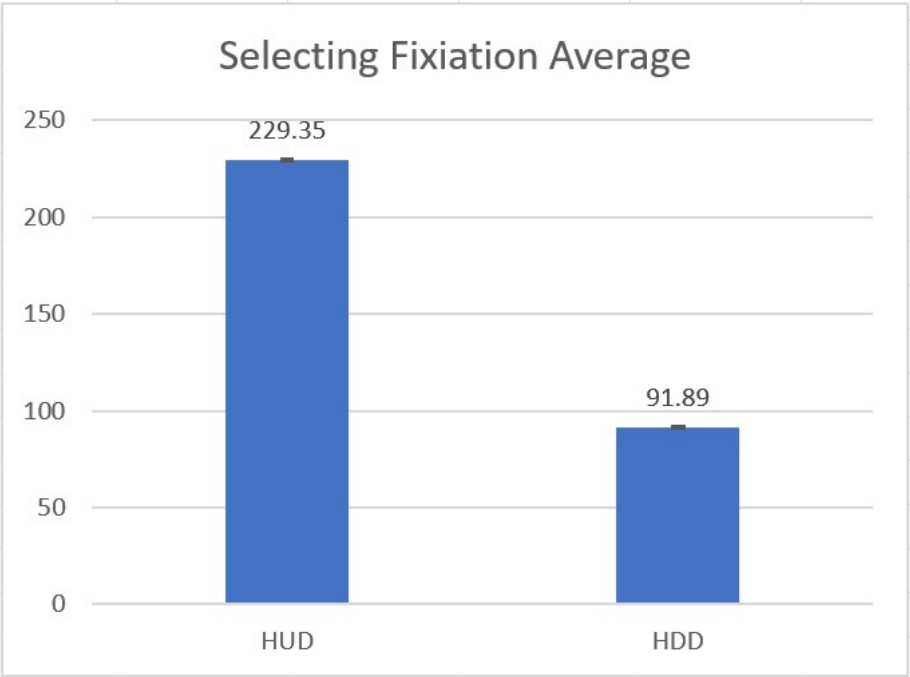


Figure 4.4: Selecting Fixation time Result

4.5.2 In-Between Saccade

By filtering out all the in between record which eye movement type is saccade and put in repeated measure ANOVA model fit, A repeated measures ANOVA showed a statistically significant difference between visual types $F(1{,}16.77) = 53.3471$, $p < 0.001$, $\eta_p^2 = 0.0234$. However, there were no main effect of sound types ($F(1{,}17.48) = 3.5139$, $p = 0.0777$) and interaction effect between visual types and sound types ($F(1{,}15.84) = 4.0022$, $p = 0.0629$).

To get a better comparison, further calculation was make like in the previous section: first the average time for each participants was calculated, and also count the number of record occurred which is both selecting and fixation. Then use the number of occurred as weight, the weight average time was calculated(Table B.3) Based on the result (Figure 4.5), HUD had($mean = 34.14$, $SD = 1071.382$) less saccade time compare to HDD($mean = 41.067$, $SD = 1198.728$) in between driving and selecting. It means that on average, participants' short glance time at the HDD longer than HUD in-between driving and selecting menu items.

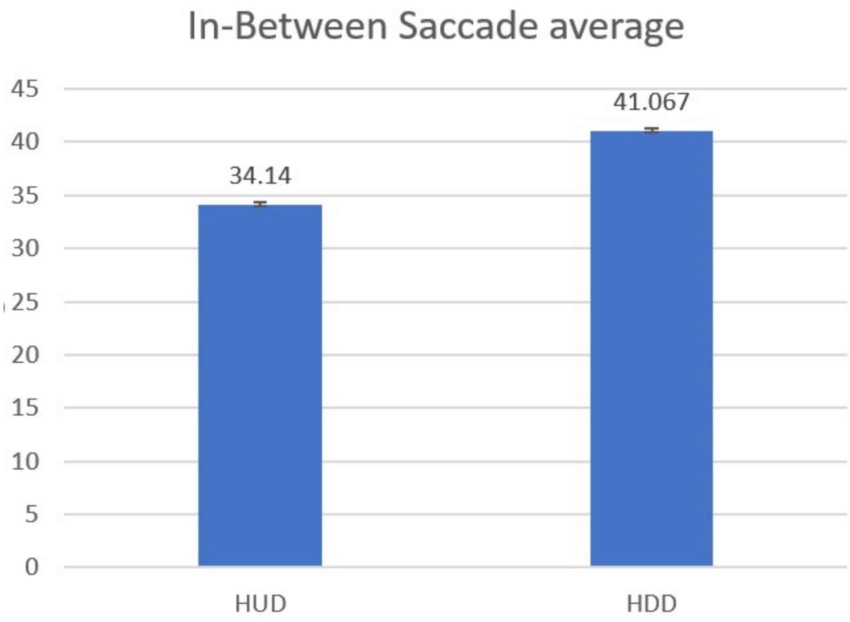


Figure 4.5: In-Between Saccade time Result

4.6 Driving data

NADS-MiniSim [8] provided data for each scenario: brake pedal force (lb), steering wheel angle, steering wheel torque (lb), collisions, bumper to bumper distance (ft), time to collision (s), and lane center offset (ft). After fit in repeated measures ANOVA model: steering wheel angle and lane center offset showed statistically significant differences. Considering that all participants experienced the same sets of driving scenarios and all conditions were counterbalanced, it is logical to assume that there are no other factors affecting the driving performance other than the experiment condition: HUD vs HDD, or sound vs no sound. All driving scenarios contain same amount of straight lane, curve lane and number of merge to other lane.

4.6.1 Steering Wheel Angle

Steering wheel angle measured how much participants turned the steering wheel. A repeated measures ANOVA showed a statistically significant difference between visual types $F(1,22.57) = 5.5419$, $p = 0.0277$, $\eta_p^2 = 0.0286$. There were no main effect of sound types ($f(1,23.68) = 0.0001$, $p = 0.9910$) and interaction effect of visual types and sound types ($F(1,23.22) = 0.2222$, $p = 0.6417$). The average for each participant was calculated afterward, and the following data table was calculated based on that (Figure: 4.6): HUD($mean = 1.651$, $SD = 4.79$) had lower average steering wheel angle compared to HDD($mean = 1.7582$, $SD = 3.149$). This suggests when using the HDD, participants became less stable when driving the car.

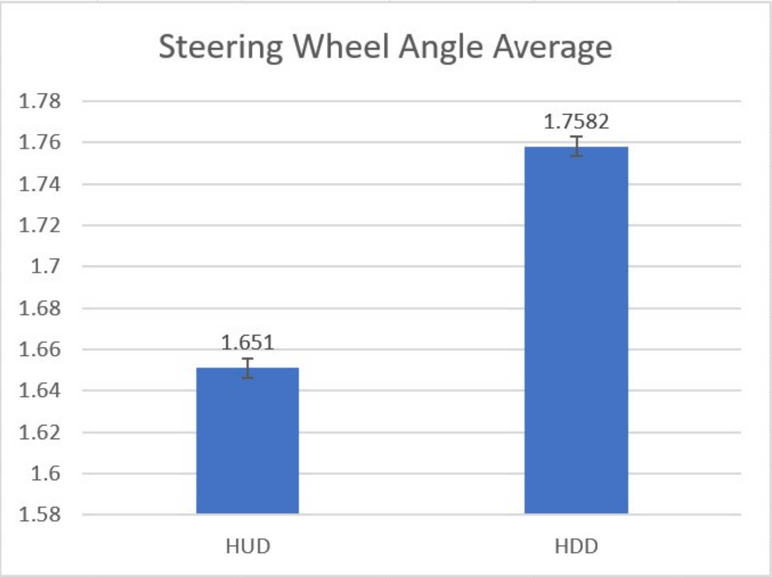


Figure 4.6: Steering Wheel Angle Result

4.6.2 Lane Deviation

lane center offset measured how much participant driving away form the center of the road. By putting all data to repeated measure ANOVA model, the parameter estimates suggest, visual feedback has a statistically significant difference from each other with $F(1{,}23.24) = 9.7091$, $p = 0.0048$, $\eta_p^2 = 0.0284$. Sound($F(1{,}23.33) = 0.3204$, $p = 0.5768$) and sound/visual combination($F(1{,}23.3) = 0.6904$, $p = 0.4144$) have no statistically significant difference, so there is no further data analysis for those conditions. The average for each participant was calculated afterward, and the following data table was calculated based on that (Figure: 4.7): HUD($mean = 0.667$, $SD = 0.951$) had higher lane deviation compared to HDD($mean = 0.5627$, $SD = 0.972$). Combine with the steering wheel angle result, It is reasonable to deduce that: even though driver maintain good control on the steering wheel when using HUD, when the land changed (vehicle entering curve or need to merge) the driver shows poor performance adapting the updating driving environment.

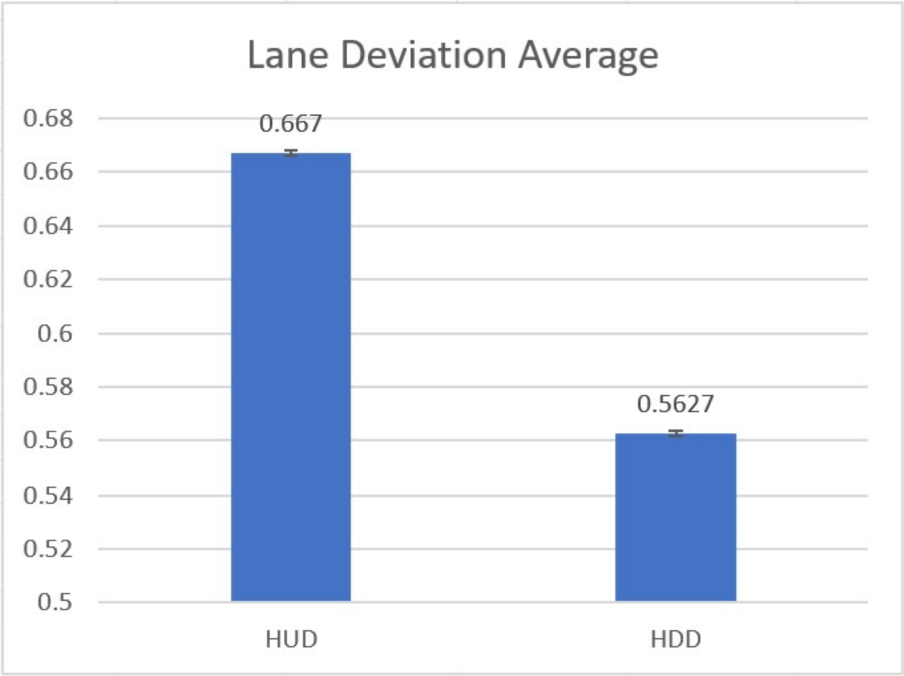


Figure 4.7: Lane Deviation time Result

Chapter 5

Discussion

This thesis aimed to increase the overall in-vehicle secondary task performance and user experience as well as driving safety by implementing a gesture interaction menu navigation system which integrated a specific HUD. To achieve this, this thesis analyzed driver situation awareness, workload, system usability, eye glance, diving performance and menu navigation performance under four different conditions (HUD with sound, HUD without sound, HDD with sound and HDD without sound).

5.1 Revisiting the results

5.1.1 Situation Awareness

There was no difference in the results of situation awareness levels 2 and 3 among the conditions. However, HUD with sound condition resulted in significantly higher level 1 situation awareness compared to other conditions (HUD without sound, HDD with sound, HDD without sound). HDD without sound showed worse SA level 1.

This result is in agreement with Gang et al. [13]'s work which reveals that spatialized warning sound increases drivers' level 1 SA. However, in their research, both level 2 and level 3 SA were also increased. Therefore, it is reasonable to suspect that the difference is caused by the application of the spatialized warning sounds. Olsson and Burns [33] suggested that the larger the visual distraction is, the longer the reaction time will be, which matches with our results. Also, the HUD was placed in a similar location as in Yang et al. [53], which provided a similar feature andincreased driver SA.

The SA results suggest that even though lower level SA is a necessary condition for the higher level SA, it is not a sufficient condition. More advanced approaches (e.g., spatial sounds about hazard) need to be implemented to increase all levels of SA.

5.1.2 Workload

Among all the workload factors (mental demand, physical demand, temporal demand, performance, effort frustration, and overall workload), HUD showed statistically significant differences compared to HDD. Furthermore, there were no significant differences between sound and no sound, which shows contradictory results with Sterkenburg et al. [41]. This system shared a similar gesture system design with their research. Their visual and audio combinations (same as this thesis' HDD with sound condition) showed significantly lower workload than the visual-only condition (same as this thesis' HDD without sound condition). However, in this thesis, there was no significant difference between with and without sound conditions. Villalobos-Zúñiga et al. [47] suggest that using HUD will significantly reduce the workload. Taken together, we can cautiously infer that when a sufficient visual display is given, drivers' workload can be reduced regardless of the presence of auditory cues.

5.1.3 System Usability Scale (SUS)

For the system usability scale, there was no statistically significant difference between sound and no sound. Moreover, there was no interaction effect between visual and sound conditions. However, there was a significant difference between visual conditions. HDD had a mean of 53.8 and HUD had a mean of 68.85. Based on the SUS score category, HDD belongs to Grade D (Poor) and HDD belongs to Grade C (Okay), which may mean that the system design can be improved further.

5.1.4 Menu Navigation Performance

For the gesture performance, there is no statistically significant difference in selecting accuracy among conditions. The results are in agreement with Sterkenburg et al. [41], which also shows that there is no difference in menu selection accuracy between conditions. Thus, to more accurately compare the gesture performance among conditions, this thesis further compared the menu navigation time. By using the repeated measures ANOVA model, the result shows that there was a significant difference between visual conditions but no difference between sound conditions. There was no interaction effect between visual and sound conditions. The result shows that HUD had a shorter time than HDD, and thus, better performance than HDD. Sterkenburg et al. [41] results demonstrated that with visual feedback, gesture performance was better than non-visual feedback. This result is also in agreement with Villalobos-Zúñiga et al. [47]'s results, which showed that the HUD condition had better performance compared to the HDD condition.

5.1.5 Eye tracking Data

For the eye tracking data, there were significant differences among the conditions. During the selection, there was a significant difference with fixation eye movement type between HUD and HDD. Also, between driving and selecting, saccade eye movement had a significant difference between HUD and HDD. There was no significant difference between sound and sound/visual combinations. When the participant was using HUD, they had more fixation time on the menu compared to using the HDD. In the mean time, HUD had less saccade between driving and selecting compared to HUD. This result matched Metz et al. [28] discussion point: drivers take the demands of the driving task into account and try to distribute available resources in a way that presumably preserves driving safety. Based on the equipment setup, HUD was placed within the user's field of view during driving, and HDD was not. Therefore, when the driver had a peripheral vision on the road, they might think that they could shift their eye focus quickly and thus, they spent more time on the menu for better selection performance (i.e., shorter selection time mentioned above). For HDD, the driver might need to move the head to see the menu. Also, because they could not see the road while viewing the HDD, they chose to switch visual focus points between the road and HDD more rapidly. Based on these two reasons, HDD showed a longer saccade weight average time between driving and selecting.

5.1.6 Driving Performance

In driving data, there was no significant difference in most factors other than steering wheel angle and lane deviation offset. There was no significant difference between sound and sound/visual combinations. HUD demonstrated better performance in steering wheel angle compared to using HDD. However, in lane center deviation, using HUD had a higher average center offset which showed that when the lane situation changes: vehicle entering a curve, the driver maintained going straight and adjusted slower than using the HDD. These results are in agreement with Sterkenburg et al. [41] who found the presence of sound had no significant difference for lane deviations. However, Villalobos-Zúñiga et al. [47] suggested otherwise, finding HUD, when compared to HDD, significantly improved the Lane deviations. Nevertheless, in their research, they had no detailed description of their driving scenario design. If their driving scenario has similar difficulty as this thesis's design, the difference in lane deviation may exist.

5.2 Situation Awareness

The situation awareness result suggests that both visual and audio conditions affect level 1 SA, which means HUD with sound helps the driver understand the current situation better than using an HDD without sound. It shows that HUD works better than HDD, with sound working better than without sound. The result is very reasonable since the time for the user to locate the hazard is extremely short. Using HUD and sound will significantly shorten the reaction time (hence causing this result). However, on level 2 and level 3, there was no significant difference among the conditions, which indicates that even though the driver has a good situation understanding, visual and audio feedback in this design may not help them properly assess the current situation and do corresponding actions based on that. While in other similar research, their HUD had improved all three levels of the driver's overall situation awareness. Considering this, it is reasonable to question if directly adopting the HDD UI design to HUD is the right choice. Since all three levels are equally important, the result didn't sufficiently show that poorly designed HUD or sound increase the driver's overall situation awareness.

5.3 Visual Feedback and Sound Feedback

All the repeated measures ANOVA model analyses indicated that there was no significant difference among visual/sound feedback combinations other than level1 SA. This suggested that compared to sound feedback, visual feedback in this thesis has some dominant effect on the driver to perform gesture-based secondary tasks. However, considering there are many contradicted results stated otherwise in similar research, this statement is only limited to that UI adopting the same design as this thesis did. Some better visual and auditory display designs show better performance in some areas.

5.4 Driving Safety

Driving safety is based on driving performance and distractions. Based on the driving performance result, when using the HUD, the driver maintains good control of the steering wheel but fails to react to lane changes, such as a vehicle entering a curve.. This can be explained by looking at eye tracking results. When using HUD, users looked at the menu for a longer time and took fewer short glances compared to HDD. Both results suggest that during selection, when a driver uses HUD, the driver keeps peripheral vision on the road so that they have worse performance on keeping the vehicle in the center of the lane. HUD can be considered a distraction from driving during selecting. Furthermore, a worse lane deviation amount indicated that peripheral vision is not safe for driving. Usually the driver who use HUD noticed the lane changing slower than HUD . While using the HUD, the driver tended to use quick glace more than stationary fix eye movement. They unusually chose to see the HDD when they think they are in a safe driving scenario. However, based on the SA result, when they misjudged the current situation and the system interrupted the secondary task and warned them about the hazard, they usually found difficult to locate the source of the hazard in a concise amount of time. Overall, using HUD or using HDD both have a positive and negative effect on driving safety. Both HUD and HDD showed no optimum driving performance. With the HUD, the driver could locate potential emergency hazards quickly (High SA1 score), but when the situation changed gradually (vehicle entering a large curve), the driver failed to notice the environment change. Meanwhile, when using HDD, if drivers misjudged the current situation, they would find it difficult to avoid the danger and fix their actions.

5.5 Secondary Task Performance

Without a doubt, using HUD had significantly better performance than using HDD. Using HUD took less average time per task compared to using HDD. However, this supports, perhaps conclusively, that the good performance here is made by compensation for driving. If the primary task (driving) has bad performance, would it be meaningful that the driver showed good performance in the secondary task?

5.6 Subjective Perception

The SUS result indicates that the drivers prefer using the HUD over HDD. The NASA-TLX result indicated that using HUD had less mental demand, physical demand, temporal demand, effort, frustration, higher performance, and overall performance. This result can conclude that using HUD had less workload than using HDD. However, the SUS result indicated that even the best performance among all conditions is still rated C, which means it is not an adequate system to use. Therefore, further improvements still need to be made to the current system.

5.7 Theoretical Implications

5.7.1 Visual Feedback and Situation Awareness

According to the SA result, HUD had a significant improvement on driver's level 1 SA compared to HDD. This result is partially in agreement with the McDonald [27] and Wang et al. [49] results which show that HUD improved the driver's overall SA from level 1 - to level 3. The experiment setup caused this result. To make the comparison consistent, the HUD and HDD provide the same warning sign (just warning, not indicating what and where the source of warning is). But in the literature, past attempts used a visual approach to improve driver SA with the capability to show where they want the drivers to pay attention. Yang et al. [53]'s design used LED to point to direction. Calhoun et al. [5]'s design highlighted the interest area. McDonald [27]'s design indicated where the landmark is. With the ability to designate the direction of interest point/area, it can shorten the driver's response time. With some specific way of highlighting (text note, change color), the system can put the driver in action even faster.

5.7.2 Auditory Display and Situation Awareness

According to the result section, the presence of sound only had significant influence on level 1 SA but not on other levels. This means that without an auditory display, the driver tends to have worse performance in analyzing the current situation. The result matches similar research efforts that discovered the HUD warning is not sufficient to improve the driver's overall SA. Since visual feedback is also an important aspect which can improve situation awareness, the visual-auditory in-vehicle warning system is more useful[20]. However, in other related research that investigated the advisory 3D sound cue to improve Driver's SA, Wang et al. [50]'s result demonstrate that in a similar pedestrian crossing scenario, the driver's level 2 SA was also improved. Moreover, their overall SA was also improved[50]. It is also similar in Gang et al. [13] results. These findings do not agree with this thesis's outcome, but may offer a potential direction to improve the system since the auditory cue used herein did not show the direction of the hazard.

5.7.3 Driving and Secondary Task Performance

For the gesture-based menu navigation, using HUD and sound both have advantages and disadvantages. While using HUD, the driver will have better secondary task performance, less workload, and more system usability. Drivers will keep peripheral vision on the road so that they can remain to assess the current situation. But this builds a false sense of security for the driver, making them believe they can correct the error in time, while in actuality, they still have poor driving performance. Since their focal vision is fully focused on the HUD (longer fixation time), it is possible that they can use peripheral vision to find incoming hazards like jaywalking pedestrians, but not notice some gradually appearing hazards (vehicle off center of the lane, vehicle entering curve). While using the HDD, the driver tends to drive more safely, but lack control of the vehicle since they know what they are going to do is dangerous. This is because HDD causes too much body movement and distraction. They use different strategies while operating the gesture menu system such as increased short glances and rapidly switching between the menu and the road to ensure driving safety. This leads to poor secondary task performance. This strategy has only worked when the driver had the right assessment of the current situation, but the result suggested drivers have low level 1 SA, so HDD can't help driver perform the secondary task safely.

5.8 Practical Implications

5.8.1 Guidelines for future research

Even though the HUD result in this thesis is not optimum, it still shows great potentials in increasing the driver's overall safety while performing the gesture-based secondary tasks by improving the performance of the gesture-based selection and partial situation awareness. It is worth carrying out research addressing the following points of using HUD as established in this thesis. Instead of showing improvements in level 2 and level 3 SA, it created distractions in the driver's attention and caused poor driving performance, specifically in maintaining driving in the lane. The current HUD UI design is for HDD use but not HUD use. Considering the position of the HUD place, the HUD can display specifically tailored information to increase drivers' overall SA (pointing the direction of the hazard, for instance). But it is also worth doing research about different elements conveying different density information, using color to represent the current level of emergency. This would include using icons to represent complex and priority information or using text to represent complex but non-priority information. Also, the size of the HUD is another direction worth investigating. A larger HUD will provide more information but, in the meantime, create a more visual distraction. How to ensure that a large-size HUD does not affect driving safety is yet another topic to investigate. The sound, spearcon, used in this thesis, is decided based on the ease of learning. It is not necessarily the most efficient sound due to the time last on average. Also, different sounds have different information densities. The different sounds may show different performance in different tasks (warning, helping gesture interaction), alone or in combination with HUD. In this thesis, all scenarios did not have turns larger than 45, and the leading vehicle helped the driver assess the environment in some way. In further research, it will be of interest to include more complex driving scenarios and real-life cases (intersection, cut in the lane, overtake[50], environment hazard) There is still room for improving the gesture interaction system by changing the interaction direction and interaction logic and making it work more seamlessly.

Chapter 6

Limitations

In this thesis, the leap-motion motion capture camera, released in July 2013, was used for hand tracking and inconsistently performed during the experiment. Different hand shapes, hand sizes, and hand colors differently performed using the same program. However, if a more accurate motion camera was in use, the gesture-related result would be better. It also slightly affected the workload measurement and system usability measurement. Moreover, the driving simulator used in this thesis does not allow pausing during the ongoing scenario, so the only possible window for collecting the SAGAT data was at the very end of the driving scenario. This led to a very small sample size of the SAGAT data compared to other similar research. To make the comparison consistent, the size of the HUD was the same as the size of the HDD. But in a real-life situation, the average HUD size equipped in a production car is significantly smaller than the HUD tested in this thesis. If the same size HUD was used in this thesis, the result may be affected.

Chapter 7

Conclusion

Using the gesture interaction system is one of the future ways of improving the overall driving experience. However, to make sure drivers safely conduct gesture interactions during driving, the use of both visual and sound feedback is effective in different ways. In this thesis, a HUD was introduced to improve the driver's overall performance, including driving, gesture menu selection, and situation awareness. It shows that using HUD will improve level 1 situation awareness and improve gesture menu selection performance. However, as a trade-off, users sacrificed driving performance in poor control of the steering wheel to maintain in lane and more time not focusing on the road despite keeping peripheral vision on the road. In the meantime, sound had no significant influence on driving and menu selection, but rather, on level 1 situation awareness, which indicates the sound choice: spearcon is not necessarily the optimum sound type choice for in-vehicle gesture interaction with HUD. Implementing gesture interactions with HUD as visual feedback is a niche domain in both the HUD and gesture interaction research area. This thesis may not present the most optimal system yet, but attempts to illustrate the great potential that a well-designed HUD can help drivers perform in-vehicle gesture integration in a more appropriate way.

Bibliography

[1] Patricia R. J. A. Alves, Joel Goncalves, Rosaldo J. F. Rossetti, Eugenio C. Oliveira, and Cristina Olaverri-Monreal. Forward collision warning systems using heads-up displays: Testing usability of two new metaphors. *2013 IEEE Intelligent Vehicles Symposium (IV)*, 2013. doi: 10.1109/ivs.2013.6629438.

[2] Donghyun Beck, Jaemoon Jung, Juhee Park, and Woojin Park. A study on user experience of automotive hud systems: Contexts of information use and user-perceived design improvement points. *International Journal of Human–Computer Interaction*, 35(20):1936–1946, 2019. doi: 10.1080/10447318.2019.1587857.

[3] Gilbert Beyer, Gian Mario Bertolotti, Andrea Cristiani, and Shadi Al Dehni. An adaptive driver alert system making use of implicit sensing and notification techniques. In *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, pages 417–424. Springer, 2010.

[4] John Brooke. Sus: a “quick and dirty’usability. *Usability evaluation in industry*, 189 (3), 1996.

[5] Gloria L Calhoun, Mark H Draper, Michael F Abernathy, Michael Patzek, and Francisco Delgado. Synthetic vision system for improving unmanned aerial vehicle

operator situation awareness. In *Enhanced and Synthetic Vision 2005*, volume 5802, pages 219–230. SPIE, 2005.

[6] Vassilis Charissis and Stylianos Papanastasiou. Human–machine collaboration through vehicle head up display interface. *Cognition, Technology Work*, 12(1):41–50, 2008. doi: 10.1007/s10111-008-0117-0.

[7] Kar-Hai Chu and Sam Joseph. Using second life to demonstrate a concept automobile heads up display (a-hud). In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*, pages 497–498, 2008.

[8] IOWA college of engineering. minisim, may 2022. URL `https://nads.uiowa.edu/minisim`.

[9] European Commission. Commission recommendation of 26 may 2008 on safe and efficient in-vehicle information and communication systems: update of the european statement of principles on human-machine interface (notified under document number c(2008) 1742). URL `https://eur-lex.europa.eu/eli/reco/2008/653/oj`.

[10] Rebecca Currano, So Yeon Park, Dylan James Moore, Kent Lyons, and David Sirkin. Little road driving hud: Heads-up display complexity influences drivers' perceptions of automated vehicles. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021. doi: 10.1145/3411764.3445575.

[11] A. Doshi, Shinko Yuanhsien Cheng, and M.m. Trivedi. A novel active heads-up display for driver assistance. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1):85–93, 2009. doi: 10.1109/tsmcb.2008.923527.

[12] Mica R Endsley. Situation awareness global assessment technique (sagat). In

Proceedings of the IEEE 1988 national aerospace and electronics conference, pages 789–795. IEEE, 1988.

[13] Nick Gang, Srinath Sibi, Romain Michon, Brian Mok, Chris Chafe, and Wendy Ju. Don't be alarmed: Sonifying autonomous vehicle perception to increase situation awareness. In *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*, pages 237–246, 2018.

[14] Kenneth W Gish and Loren Staplin. Human factors aspects of using head up displays in automobiles: A review of the literature. 1995.

[15] Lisa Graichen, Matthias Graichen, and Josef F. Krems. Evaluation of gesture-based in-vehicle interaction: User experience and the potential to reduce driver distraction. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 61(5): 774–792, 2019. doi: 10.1177/0018720818824253.

[16] Tobii Group. Tobii pro glasses 2 - discontinued, may 2022. URL `https://www.tobiipro.com/product-listing/tobii-pro-glasses-2/`.

[17] Leo Gugerty and William Tirre. A pc-based driving simulator for assessing situation awareness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 41, pages 895–895. SAGE Publications Sage CA: Los Angeles, CA, 1997.

[18] Sandra G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909. URL `https://doi.org/10.1177/154193120605000909`.

[19] Robert J.K. Jacob, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit

Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. Reality-based interaction: A framework for post-wimp interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, page 201–210, New York, NY, USA, 2008. Association for Computing Machinery. ISBN 9781605580111. doi: 10.1145/1357054.1357089. URL https://doi.org/10.1145/1357054.1357089.

[20] Uijong Ju, Lewis L. Chuang, and Christian Wallraven. Acoustic cues increase situational awareness in accident situations: A vr car-driving study. *IEEE Transactions on Intelligent Transportation Systems*, 23(4):3281–3291, 2022. doi: 10.1109/TITS.2020.3035374.

[21] Yung-Ching Liu. Effects of using head-up display in automobile context on attention demand and driving performance. *Displays*, 24(4-5):157–165, 2003.

[22] Yung-Ching Liu and Ming-Hui Wen. Comparison of head-up display (hud) vs. head-down display (hdd): driving performance of commercial vehicle operators in taiwan. *International Journal of Human-Computer Studies*, 61(5):679–697, 2004.

[23] Paul A Lucas. An evaluation of the communicative ability of auditory icons and earcons. Georgia Institute of Technology, 1994.

[24] Udara E. Manawadu, Mitsuhiro Kamezaki, Masaaki Ishikawa, Takahiro Kawano, and Shigeki Sugano. A hand gesture based driver-vehicle interface to control lateral and longitudinal motions of an autonomous vehicle. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 001785–001790, 2016. doi: 10.1109/SMC.2016.7844497.

[25] Michael L Matthews, David J Bryant, Robert DG Webb, and Joanne L Harbluk. Model for situation awareness and driving: Application to analysis and research for intelligent transportation systems. *Transportation research record*, 1779(1):26–32, 2001.

[26] Keenan R May, Thomas M Gable, and Bruce N Walker. A multimodal air gesture interface for in vehicle menu navigation. In *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '14, page 1–6, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450307253. doi: 10.1145/2667239.2667280. URL https://doi.org/10.1145/2667239.2667280.

[27] Zachary McDonald. *Augmented Reality HUDs: Warning Signs and Drivers' Situation Awareness*. Rochester Institute of Technology, 2016.

[28] Barbara Metz, Nadja Schömig, and Hans-Peter Krüger. Attention during visual secondary tasks in driving: Adaptation to the demands of the driving task. *Transportation research part F: traffic psychology and behaviour*, 14(5):369–380, 2011.

[29] Muhammad Syafiq Syed Mohamed and Chi-Ngong Pow. Kansei engineering approach towards automotive heads up display (hud) interface design. 2017.

[30] Engineering National Academies of Sciences, Medicine, et al. Analysis of naturalistic driving study data: Safer glances, driver inattention, and crash risk. 2015.

[31] Hyung Jun Oh, Sang Min Ko, and Yong Gu Ji. Effects of superimposition of a head-up display on driving performance and glance behavior in the elderly. *International Journal of Human-Computer Interaction*, 32(2):143–154, 2016.

[32] Shigeru Okabayashi, Masao Sakata, Jun'ichi Fukano, Shigetoshi Daidoji, Chikara Hashimoto, and Tomonari Ishikawa. Development of practical heads-up display for production vehicle application. *SAE Transactions*, 98:638–647, 1989. ISSN 0096736X. URL http://www.jstor.org/stable/44472311.

[33] S Olsson and Peter C Burns. Measuring driver visual distraction with a peripheral detection task. *Obtained from August*, 2000.

[34] Hyesun Park and Kyong-ho Kim. Efficient information representation method for driver-centered ar-hud system. In *International Conference of Design, User Experience, and Usability*, pages 393–400. Springer, 2013.

[35] Marina Plavić, Heiner Bubb, and Markus Duschl. Ergonomic design and evaluation of augmented reality based cautionary warnings for driving assistance in urban environments. 2009.

[36] Bianka Rommel. Bmw gesture control - the next level of idrive interaction, Oct 2020. URL https://www.bimmer-tech.net/blog/item/124-bmw-gesture-control.

[37] Gözel Shakeri, John H. Williamson, and Stephen Brewster. May the force be with you: Ultrasound haptic feedback for mid-air gesture interaction in cars. In *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '18, page 1–10, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450359467. doi: 10.1145/3239060.3239081. URL https://doi.org/10.1145/3239060.3239081.

[38] Gözel Shakeri, John H. Williamson, and Stephen Brewster. Novel multimodal feedback techniques for in-car mid-air gesture interaction. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2017. doi: 10.1145/3122986.3123011.

[39] Russell J Sojourner and Jonathan F Antin. Journal of the human.

[40] Jason Sterkenburg, Steven Landry, and Myounghoon Jeon. Eyes-free in-vehicle gesture controls: Auditory-only displays reduced visual distraction and workload. In

Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct, AutomotiveUI '17, page 195–200, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450351515. doi: 10.1145/3131726.3131747. URL https://doi.org/10.1145/3131726.3131747.

[41] Jason Sterkenburg, Steven Landry, and Myounghoon Jeon. Eyes-free in-vehicle gesture controls. *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct*, 2017. doi: 10.1145/3131726.3131747.

[42] Jason Sterkenburg, Steven Landry, and Myounghoon Jeon. Design and evaluation of auditory-supported air gesture controls in vehicles. *Journal on Multimodal User Interfaces*, 13(2):55–70, 2019. doi: 10.1007/s12193-019-00298-8.

[43] Lars St»hle and Svante Wold. Analysis of variance (anova). *Chemometrics and Intelligent Laboratory Systems*, 6(4):259–272, 1989. ISSN 0169-7439. doi: https://doi.org/10.1016/0169-7439(89)80095-4. URL https://www.sciencedirect.com/science/article/pii/0169743989800954.

[44] Kathryn G. Tippey, Elayaraj Sivaraj, Wil-Johneen Ardoin, Trey Roady, and Thomas K. Ferris. Texting while driving using google glass. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58(1):2023–2027, 2014. doi: 10.1177/1541931214581422.

[45] Text to speech online. Free text to speech: Online, app, software, commercial license with natural sounding voices., may 2022. URL https://www.naturalreaders.com/online/.

[46] ultraleap. Leap motion controller, may 2022. URL https://www.ultraleap.com/product/leap-motion-controller/.

[47] Gabriela Villalobos-Zúñiga, Tuomo Kujala, and Antti Oulasvirta. T9+hud: Physical keypad and hud can improve driving performance while typing and driving. In *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Automotive'UI 16, page 177–184, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450345330. doi: 10.1145/3003715.3005453. URL `https://doi.org/10.1145/3003715.3005453`.

[48] Bruce N Walker, Jeffrey Lindsay, Amanda Nance, Yoko Nakano, Dianne K Palladino, Tilman Dingler, and Myounghoon Jeon. Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. *Human Factors*, 55(1):157–182, 2013.

[49] Jianmin Wang, Wenjuan Wang, Preben Hansen, Yang Li, and Fang You. The situation awareness and usability research of different hud hmi design in driving while using adaptive cruise control. In *International Conference on Human-Computer Interaction*, pages 236–248. Springer, 2020.

[50] MinJuan Wang, Sus Lundgren Lyckvi, Chenhui Chen, Palle Dahlstedt, and Fang Chen. Using advisory 3d sound cues to improve drivers' performance and situation awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 2814–2825, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346559. doi: 10.1145/3025453.3025634. URL `https://doi.org/10.1145/3025453.3025634`.

[51] Christopher D Wickens. Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3(2):159–177, 2002.

[52] Deyou Xu. A neural network approach for hand gesture recognition in virtual reality

driving training system of spg. In *18th International Conference on Pattern Recognition (ICPR'06)*, volume 3, pages 519–522, 2006. doi: 10.1109/ICPR.2006.109.

[53] Yucheng Yang, Burak Karakaya, Giancarlo Caccia Dominioni, Kyosuke Kawabe, and Klaus Bengler. An hmi concept to improve driver's visual behavior and situation awareness in automated vehicle. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 650–655, 2018. doi: 10.1109/ITSC.2018.8569986.

Appendices

Appendix A

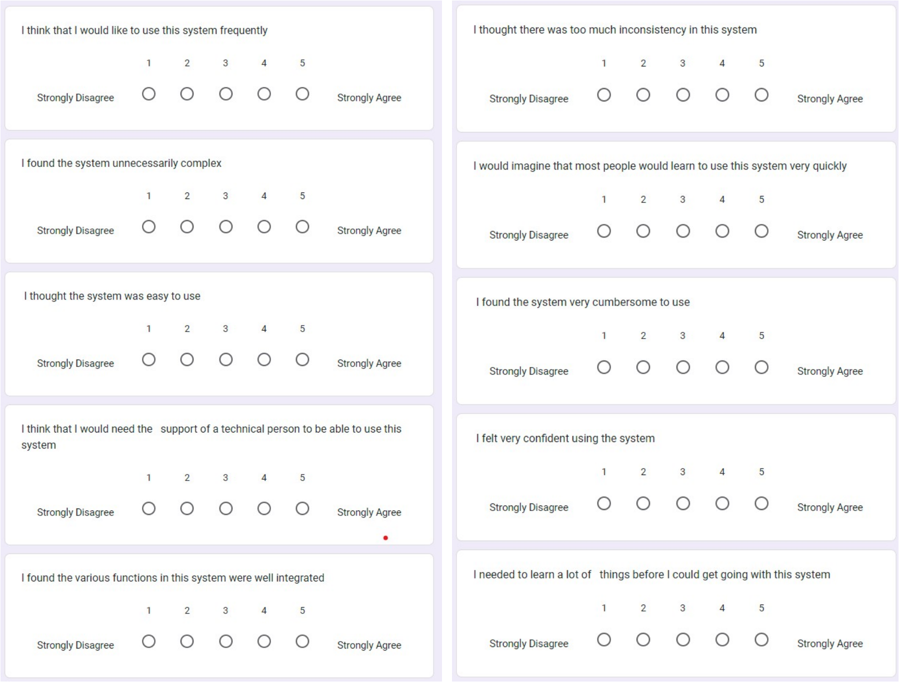
supported Questionnaire

A.1 SAGAT

SAGAT Situation Awareness Questionnaire

1. What elements of interest do you see on the screen?
2. What do these elements tell you about the current situation(are you safe or in danger)?
3. What vehicles did you notice around you?
4. What buildings did you notice around you?
5. How many pedestrians do you see on the screen?
6. What is currently happening in the scenario?
7. What do you think will happen next?
8. What do you think you need to do next?

A.2 SUS Questionnaire



I think that I would like to use this system frequently

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the system unnecessarily complex

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought the system was easy to use

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I think that I would need the support of a technical person to be able to use this system

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the various functions in this system were well integrated

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I thought there was too much inconsistency in this system

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I would imagine that most people would learn to use this system very quickly

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I found the system very cumbersome to use

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I felt very confident using the system

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

I needed to learn a lot of things before I could get going with this system

	1	2	3	4	5	
Strongly Disagree	○	○	○	○	○	Strongly Agree

Figure A.1: SUS Questions

Appendix B

data

B.1 Eye Tracking

Table B.1: Average Time for Selecting(fixation)

	HDD		HUD	
ID	count	average	count	average
1047	71	203.0986	510	690.9902
1288	72	340.4722	50	769.7
1379	11	316.3636	184	886.9022
1385	33	172.7273	199	1440.739
1528	16	141.25	102	914.8235
1985	128	280.4297	143	513.5455
2121	131	336.9008	340	627.6324
2233	192	225.2969	155	430.3806
2309	31	189	172	664.4709
2462	85	219.2824	221	999.9548
2895	1	140	258	728.4419
2937	170	227.3824	105	930.6667

Table B.2: Average Time for Selecting(fixation)(continue)

	HDD		HUD	
ID	count	average	count	average
6688	219	259.6073	393	921.2316
7121	201	486.408	519	849.8189
7122	27	177.037	135	578.1333
7258	154	288.3961	337	531.0386
7625	87	213.3333	260	545.0615
8031	14	307.1429	227	543.7048
8898	115	218.4	140	652.4
9005	1	100	291	1286.928

Table B.3: Average time for in-between(saccade)

	HDD		HUD	
ID	average	count	average	count
1047	39.70032	1268	31.846	1013
1288	45.31879	1192	35.88235	612
1379	36.80723	664	32.05769	1040
1385	48.08989	445	31.27983	461
1528	43.48294	1114	31.96721	732
1985	41.21569	2040	38.30735	898
2121	37.34139	993	28.49817	1365
2233	41.62188	2084	37.31088	2049
2309	38.52761	489	33.89281	1418
2462	44.93292	969	36.57551	841
2895	38.16781	1168	32.88889	1080
2937	43.52148	1187	37.58958	614
3668	39.75104	964	32.29917	1083
3927	42.63014	365	36.26016	492
4343	45.8954	1262	33.24355	891
4455	34.72308	1300	33.40881	1590
6688	40.0854	1171	32.96329	1117
7121	42.96678	1746	38.40871	1194
7122	37.49086	1642	33.57739	1557
7258	45.02398	1668	38.39219	1331
7625	41.68096	1166	30.65693	822
8031	43.12245	980	32.52252	888
8898	36.12903	1023	35.62558	1079
9005	35.91398	744	31.3799	587