IN-VEHICLE PERFORMANCE AND DISTRACTION FOR MIDAIR AND TOUCH DIRECTIONAL GESTURES

by

ARMAN HAFIZI

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirements for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2023

AUTHOR'S DECLARATION

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

STATEMENT OF CONTRIBUTIONS

This thesis is adapted from a manuscript written for publication [11]. Arman Hafizi was the main author for all chapters which were written under the supervision of Prof. Daniel Vogel and with collaboration of Jay Henderson, Ali Neshati, and Wei Zhou. I also want to thank my late supervisor, Prof. Edward Lank who assisted me so much during the early steps. His memory will be with me always.

ABSTRACT

We compare the performance and level of distraction of expressive directional gesture input in the context of in-vehicle system commands. Center console touchscreen swipes and midair swipe-like movements are tested in 8-directions, with 8-button touchscreen tapping as a baseline. Participants use these input methods for intermittent target selections while performing the Lane Change Task in a virtual driving simulator. Input performance is measured with time and accuracy, cognitive load with deviation of lane position and speed, and distraction from frequency of off-screen glances. Results show midair gestures were less distracting and faster, but with lower accuracy. Touchscreen swipes and touchscreen tapping are comparable across measures. Our work provides empirical evidence for vehicle interface designers and manufacturers considering midair or touch directional gestures for centre console input.

ACKNOWLEDGMENTS

This work was made possible by Waterloo-Huawei Joint Innovation Lab, NSERC Discovery Grant 2018-05187, and Canada Foundation for Innovation Infrastructure Fund 33151 "Facility for Fully Interactive Physio-digital Spaces."

DEDICATION

To my incredible parents, whose unwavering love and support have been the driving force behind my progress. This thesis is a testament to your guidance and sacrifices.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

QUOTE

"The force is strong with this one."

– Anakin Skywalker a.k.a. Darth Vader

1 INTRODUCTION

The automotive industry is progressively moving to centralize the display and control of vehicle systems on a single console screen [32]. Originally for “infotainment” like music and navigation, these are now frequently used for features like cabin climate control and vehicle status, and increasingly, even for critical functions like windshield wiper speed [31]. Likely due to familiarity and the manufacturing economy, a touchscreen is commonly used for primary input, but this means the driver has to momentarily redirect their visual focus away from the roadway [25]. Speech control [2, 7] and steering wheel buttons [13, 29] are potential eyes-free alternatives, but they have limitations. For instance, speech can be slow and unreliable, especially in loud environments, and it can disturb passengers [7, 15]. Steering wheel buttons increase manufacturing complexity and cost, and they are typically reserved for dedicated vehicle functions like cruise control.

Gestures performed on the touchscreen or in midair can in theory be performed eyes-free, and midair gestures are already used in some high-end vehicles. For example, the G11 7-Series BMW uses midair drags and swipes for functions like phone functions, media player, and navigation [4], and the S-Class Mercedes-Benz uses a midair swipe to control the sunroof [24]. Both use an index-to-thumb pinch to delimit midair gestures from other hand movements, and both use some form of directional movement [34]. However, the performance of midair gestures compared to conventional touchscreen input has only been evaluated in limited forms and higher-level tasks. For example, Parada-Loira, González-Agulla, and Alba-Castro [28] tested static and dynamic hand gestures moving through menus and selecting options, Wu et al. [38] tested very simple left-right directional gestures with a more open-ended interface navigation task, and Graichen, Graichen, and Krems [10] tested a set of completely different hand gestures to perform navigation related tasks. In particular, more expressive directional gestures have not been directly compared for midair and touch using a short-duration atomic type of selection task.

We compare expressive 8-direction gestures when performed as touchscreen swipes or as dynamic midair movements while forming a pinch posture. An equivalent 8-button condition using touchscreen tapping is included as a baseline. Participants use these conditions to complete intermittent short-duration single target selections while simultaneously performing the standard Lane Change Task [20] in a driving simulator. Input performance metrics include response time, selection time, and accuracy, and level of distraction is measured using Standard Deviation of Lateral Position (SDLP) [35], deviation from target vehicle speed, and frequency of off-screen glances. Results show that midair gestures are fast, more error prone, but least distracting. Touchscreen swipes are accurate and fast to respond with medium distraction. Touchscreen tapping is slower to respond and accurate, but most distracting. Since the participants behavioural space for midair gestures were wider than touch-screen gestures and midair gestures were clustered into more groups than touch-screen ones with the same parameters, we looked for ways to address this issue. Simulated experiments with data shows that if midair directions are reduced to 4, accuracy is comparable to both touchscreen conditions. Data and analysis scripts are available[1].

We contribute the first highly controlled study directly comparing more expressive 8-direction midair and touch gestures for in-vehicle control. Our results provide empirical evidence for vehicle interface designers and manufactures considering midair or touch gestures for centre console input.

1 https://github.com/exii-uw/invehicle-gestures

2 BACKGROUND AND RELATED WORK

After surveying elicitation studies to motivate the relevance of touch and midair directional gestures, we summarize previous work comparing different forms of gestures to other input methods under varying task conditions. We focus entirely on in-vehicle interaction given the specialized physical context and importance of considering input that is compatible with driving.

2.1 IN-VEHICLE GESTURE ELICITATION

Several works elicit what gestures people prefer to use when communicating with in-vehicle systems.

Wu et al. [37] used three studies to design user-defined in-vehicle midair gestures for interaction with an in-vehicle infotainment system. Using requirement analysis and function definition, gesture elicitation, and gesture vocabulary evaluation, they found up and down swiping, fanning oneself, and tapping were the most desired midair gestures by the participants. Up and down swiping is a form of directional gesture, our work tests a much expanded variation with 8 directions.

Fariman et al. [9] conducted a formal user elicitation study to establish a midair input vocabulary for in-vehicle systems. The results discovered some frequently-observed gestures in the context of in-vehicle control were not well described by Wobbrock, Morris, and Wilson in their general input elicitation study [36]. This suggests that input methods for vehicles need to be considered distinctly from general input. Based on their set of elicited gestures, Fariman et al. proposed a novel midair gesture vocabulary for in-vehicle systems called GestDrive. This included static and directional dynamic gestures performed with full-hand, two-fingers, and one-finger.

May, Gable, and Walker [22] combined a gesture elicitation study and an online gesture comparison to design a midair gesture set for in-vehicle interaction. Based on their results, they argue for three qualities to reduce driver workload with midair in-vehicle gestures: the gesture set should fit together conceptually; all gestures should adhere to a cohesive interaction metaphor and consistent spatial mapping; and the gestures should be unlikely to lead to system recognition errors due to distinctive shapes and clear user formation rules. They propose a set of gestures that include making a fist, displaying the palm, pointing, index finger flicking, and swiping. Our focus on a set of directional midair gestures is similar to their flicking and swiping gestures. In theory, using one hand posture and distinguishing a set of gestures based on direction satisfies the three qualities suggested by May, Gable, and Walker.

Ohn-Bar and Trivedi [27] suggested midair gesture sets including directional swipe, pinch, scroll, etc. for functionalities such as answering a call and song selection. These gestures were selected to evaluate the vision-based approaches for hand gesture recognition in a car.

Angelini et al. [3] proposed different ways to gesture on a steering wheel with a user-elicited taxonomy. They considered predefined commands such as selecting songs and and free commands as the participants desired. The study generated a set of gestures consisting of tap, swipe, and pressure using thumb, index, and whole-hand. Their swipe gesture on the steering wheel is quite similar to the direction swipe gesture we study on the centre console.

A collective observation from these past elicitation works is the dominance of directional gestures, like swipes, performed on a touch surface or in midair. In addition, for midair gestures, several studies found a pinch drag as a preferred gesture. We use these findings to justify our selection for comparing a touch screen and midair directional gesture input methods, and using pinch as a delimiter for midair.

2.2 COMPARISON WITH MIDAIR GESTURES

Researchers have implemented and tested different variations of midair and touch gestures for in-vehicle systems.

Parada-Loira, González-Agulla, and Alba-Castro [28] developed a system using an IR camera to detect midair gestures that are static (e.g. showing one to five fingers, an OK sign, an open or closed hand) and dynamic (e.g. directional movement, shaking, wrist twist). The gestures are used for general input, such as increasing and decreasing a continuous control like volume, as well as triggering dedicated commands, such as playing or pausing music. Using a questionnaire, they compared these midair gestures with existing touchscreen gestures in a driving simulator and in a parked prototype vehicle. The driving task was uncontrolled and open-ended: the participant only had to maintain a speed limit of 120 km/h. Their results show touchscreen gestures were more reliable and easier to use, but midair hand gestures were less distracting and more useful.

May, Gable, and Walker [21] proposed and tested a set of multi-modal barehand midair gestures for in-vehicle systems. They asked participants to drive in a low fidelity simulator while performing secondary menu navigation tasks using midair gestures with feedback sound, midair gestures without feedback sound, and a baseline using direct touch tapping. They found both midair gesture conditions had higher levels of mean lateral deviation compared to baseline, but not significantly different from direct touch. In following work, May et al. [23] measured the impact of menu length and midair gesture type on driver workload. Driver workload was measured using brake response time, task completion time, and NASA TLX. They found that swipe gestures were higher performing for short menus while hold to scroll systems, where one holds their hand with the palm facing down and changes the extension degree of wrist up or down, may provide versatility for longer menus.

Häuslschmid, Menrad, and Butz [14] examined the combined effect of scale and surface with small one and two finger gestures performed on the steering wheel ("micro-hand") and midair barehand gestures. They measure the level of driver distraction using the Lane Change Task. They found barehand gestures impeded lateral control meaning lower success rate (rate of correctly performed lane changes) and slightly higher reaction time but had less deterioration meaning less lateral deviation. Micro-hand gestures had a higher success rate and lower reaction time.

Angelini et al. [2] compared a combined midair and touch gestures performed on the steering wheel, such as a midair swipe upwards and downwards, and tap on the wheel. Baseline conditions included touch taps on centre console and speech commands. Using a primary driving task (driving simulator in a city centre with traffic condition) and secondary infotainment system input task, they measured performance and subjective workload. They observed no difference between input conditions regarding driving performance, but found there were fewer user interactions (higher performance) with speech and lower task completion time with touch.

Harrington et al. [12] evaluated the effect of ultrasound haptic feedback on midair and touchscreen gestures for atomic and complex input tasks. In a desktop car simulator, participants followed a lead car while performing two input tasks separately: target selection (atomic) and manipulate a slide-bar (complex adjust and select). Target selection was selecting one of four targets, and slide-bar manipulation was increasing or decreasing a value by dragging a slide-bar in a certain direction and magnitude. In the touch condition, an on-screen tap selected the target and right/left on-screen swipes adjusted the slide-bar. In the midair condition, targets were selected with downward movement of an open-hand and the slide-bar bar adjusted with right/left movement of an open hand and then making a closed fist. Each input method was performed with and without ultrasound haptic feedback. Their results showed that midair had longer off-road glances for target selection, less accuracy for slide-bar, and longer selection time than touch.

Wu et al. [38] compared touchscreen tapping, touchscreen swiping, midair tapping, and midair swiping. All methods used similar directional motions: tapping methods had left or right targets and swiping methods used motions to the left or to the right. Midair taps were done by pointing the index finger to two virtual targets defined at a fixed position beside the driver and recognized using a Leap Motion device. No feedback on correctness was provided for either of the methods. Midair swipes were done by rotating the wrist with no specific hand pose and detected using speed, magnitude, and direction of the trajectory. The input task was a "song selection task" that required a long sequence of individual menu navigation input actions (the duration of each task was approximately 10 to 30 seconds). To quantify driver attention, the study used a desktop driving simulator with "central lane maintaining" and "lead car following" tasks and associated metrics, in addition to subjective workload using the NASA TLX. They found the midair gestures were slower and had a higher workload than touchscreen gestures. They also report that swipe gestures, across both the midair and surface conditions, had a lower workload than tap gestures.

Most recently, Graichen, Graichen, and Krems [10] compared different dynamic gestures performed midair and and on a touchscreen, for in-vehicle systems. There were five types of midair input: making a fist posture, making a pointing posture, flashing the palm, and performing a 2-directional swipe with either a thumb to 4-finger pinch or an open hand (Fig. 2.1). The touchscreen condition was taps on 8 different buttons which were labelled with vehicle functions like show traffic information, mute, raise volume, and start navigation. They quantified acceptance and measured workload using NASA TLX with an objective measure of off-screen glances. These input actions were performed while performing two open-ended driving tasks in a virtual simulator. One task was a simple approximation of city driving with right and left turns and a maximum speed of 50 km/h. The other approximated motorway driving where the participant drove in the right side lane with occasional overtaking manoeuvres, all while maintaining a speed of 130 km/h. The results showed midair gestures required fewer and shorter off-screen glances. Midair gestures also had a higher acceptance rate and lower workload.

Figure 2.1: Multiple variations of midair gesture input used by Graichen, Graichen, and Krems [10].

Table 2.1 summarizes the works just discussed that also compare midair input for in-vehicle systems. All use either a complex multi-step input task, such as navigating infotainment menus or triggering sequences of vehicle functions (Fig. 2.2), or they include multiple variations of an input method within the same task, such as an up-down swipe to adjust volume, an open hand to zoom a map, and pointing for repeating navigation instructions (2.1). Using a complex task with sequences of input methods variations is intended to approximate specific use cases for in-vehicle interaction. However, these are hard to experimentally control and directly compare different input

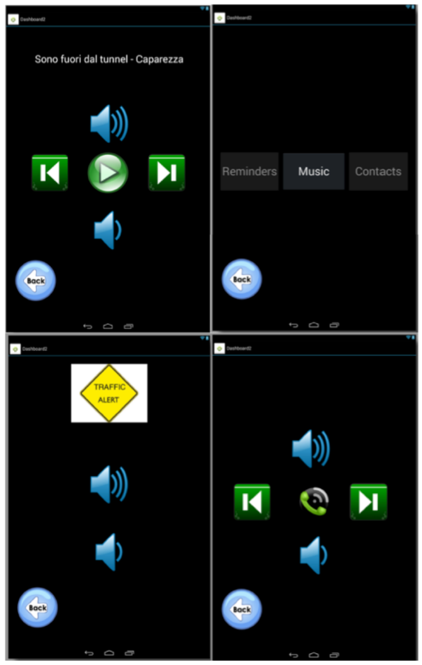

Figure 2.2: Multi-step input task used by Angelini et al. [2].

methods, and generalization is unclear for different interfaces and use cases. Our approach is to decompose complex tasks into "atomic" input actions, each a single, defined movement that can trigger an arbitrary in-vehicle command. Mapping an experiment task trials to single atomic actions lends more internal validity and can be easier generalize to a wider range of real car tasks. This approach is similar to how Fitts' law experiments test pointing using individual target selection trials since the goal is to isolate the fundamental motor action. The complex tasks used in previous work are valuable to test specific interfaces, similar to how usability tests are used with conventional computing systems. The atomic task approach also enables us to include more variations of each input method. While past work tested 1 or 2 movement directions, we evaluate 8. Finally, we adopt the lane change task with SDLP measure, a highly controlled driving task with high internal validity, which was also used by Wu et al. [38] and Häuslschmid, Menrad, and Butz [14].

Table 2.1: Summary and comparison of previous works that also compared a midair input method in terms of input methods included in the evaluation, the input task and measures, and the driving task. See text for more details.

		Input Task		**Driving Task**			
	Input Methods	**Type**	**Measures**	**Type**	**Apparatus**	**Metrics**	**Distraction**
Parada-Loira et al. [28]	midair *(1 directional)* midair other *(OK sign, open/close hand, shaking)* touch tap	complex media task (song, volume)	*questionnaire	maintaining speed limit	lab simulator, parked car	*questionnaire	*questionnaire
May et al. [21]	midair *(2-directional with open hand)* midair *(open/close hand)* touch tap	complex 3-level list, selection task	selection time, accuracy	following lead car	lab simulator	brake response time	off-screen glances
Häuslschmid et al. [14]	midair *(2-directional with open hand)* finger on-wheel *(2-directional)*	complex media task (song, volume)	*questionnaire	lane change task	lab simulator	SDLP; success rate	*not measured*
Angelini et al. [2]	midair *(2-directional)* touch tap speech	complex infotainment task (volume, call, song)	selection time, accuracy	driving with traffic in city centre	lab simulator	driving violations; speed; *NASA TLX	*not measured*
Harrington et al. [12]	midair *(2-directional with open hand then fist)* touch tap touch swipe	atomic and complex selection action	selection time, accuracy	following lead car	lab simulator	SDLP	off-screen glances
Wu et al. [38]	midair *(2-directional)* touch tap touch swipe	complex target selection (song menu)	selection time	maintaining central lane; following lead car	lab simulator	brake response time; SDLP; *NASA TLX	*not measured*
Graichen et al. [10]	midair *(fist, pointing, pinch drag)* touch tap	complex infotainment task (navigation, volume, radio)	*none*	driving with traffic in city and on motorway	lab simulator	*NASA TLX	off-screen glances
Our Work	midair *(8-directional pinch)* touch tap touch swipe	atomic selection action	selection time, accuracy	lane change task, maintaining speed limit	lab simulator	SDLP, speed, *NASA TLX	off-screen glances

3 INPUT METHODS

This section provides design and implementation details for the three input methods to be compared in the following experiment, as well as the driving simulation setup and input sensing equipment. All three methods support 8 different commands. Directional gesture or swiping methods differentiate based on movement direction and tapping differentiates based on position. In all methods, the target position or direction is referred to as up, up-right, right, down-right, down, down-left, etc.

Driving simulator. Our setup used OpenDS[1] software running on a Windows PC (CPU speed 3 GHz) with an ultra-wide 43" Samsung Odyssey G9 LED monitor. All driving input used a Logitech G29 Driving Force Racing steering wheel and pedals. The steering wheel was clamped to a 75 cm high desk with the monitor placed 90 cm from the steering wheel front. The driver is seated in a fixed leg, non-swivelling office chair. Figure 3.1 shows the setup.

Input sensing. A Galaxy Tab 9 tablet with 10.5" screen (22.6 $\times$ 14.2 cm) was used as a vehicle centre console. It was placed 20 cm to the right of the steering wheel centre on the desk clamped at a 20° tilt angle. The position and orientation were chosen to be easily reachable and approximate the typical position for a centre console in a vehicle.

We used a Vicon motion tracking system with Tracker software[2] for tracking the positions of fingers to enable midair gestures. Seven cameras were used in our setup (2 top front, 2 lower front, 2 top back, 1 top left). Two 13 mm markers were attached to the nails of the index finger and thumb on the participant's right hand using double-sided foam tape. The system's tracking frequency is 120 Hz with accuracy <1 mm for the markers we used.

In order to capture the off-screen glances of the drivers, a 4K webcam was placed 70 cm in front of the driver. Captured video was analyzed in Python using the Gaze Tracking library[3] to determine if the driver is looking at the driving simulator display or the centre console tablet.

Pilot test. The midair and swipe methods are implemented using parameterized rules to determine states and events like when a swipe motion starts, what direction is a swipe, and when a hand is pinching. To tune these parameters, we conducted a pilot with 4 participants in which they performed the same driving and input task used in the main experiment, but we only logged input movement data. We use mean values and 95% intervals from relevant data to set thresholds and direction tolerances for the different input

1 https://opends.dfki.de/
2 www.vicon.com/software/tracker
3 https://github.com/antoinelame/GazeTracking/

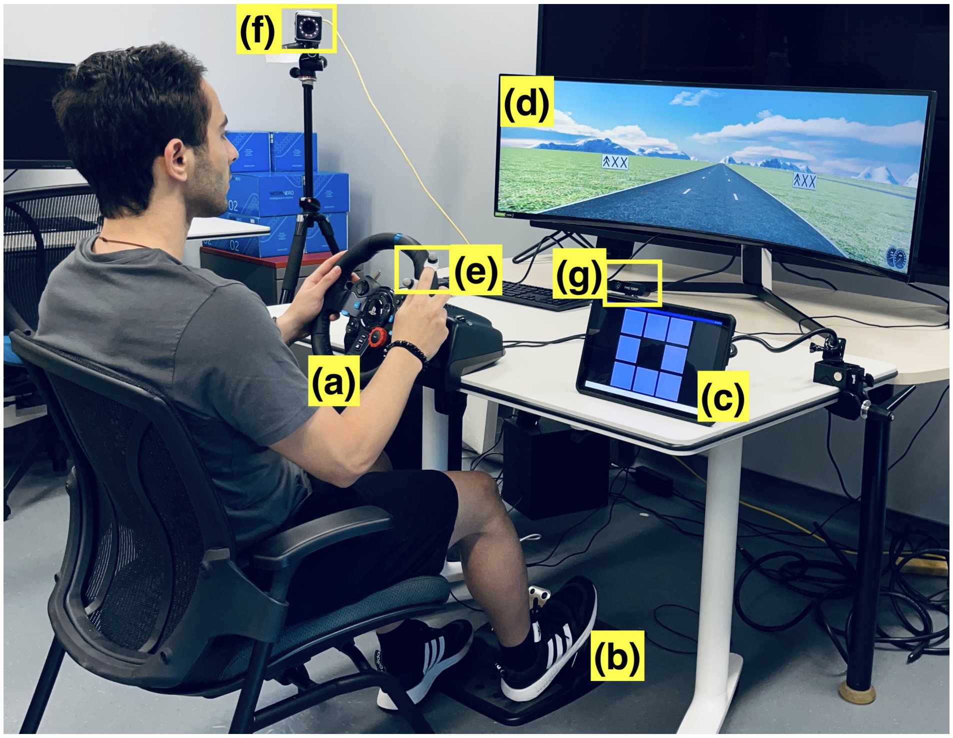

Figure 3.1: Setup used for driving simulation and input tracking: (a) steering wheel clamped to desk with fixed driver seat; (b) pedals below; (c) tablet used for infotainment screen; (d) OpenDS driver view displayed on 43" ultra-wide display; (e) markers attached to the index finger and thumb; (f) Vicon camera (the lower left front); (g) webcam used for eye tracking.

methods (each explained below). This pilot also functioned as a technical system test and it verified the driving task difficulty was appropriate.

3.1 TOUCH TAPPING

A touch tap is defined as a touch down event and then a touch up event without movement from the initial touch position. We used the standard Android threshold for distinguishing a touch tap. All 8 touch targets were shown in a 3 × 3 grid with an empty space in the centre cell coloured purple over a black background. Each target was 38 × 38 mm with 5 mm gaps between all targets to avoid overlapping (Figure 3.3a). No feedback for touch position or selected target was shown.

3.2 TOUCH SWIPING

A touch swipe is defined as a touch down, then an uninterrupted movement along the surface of the screen for at least 12 mm (100 pixels) with minimum speed of 12 mm/s (Figure 3.3b). Each event generated two time stamped

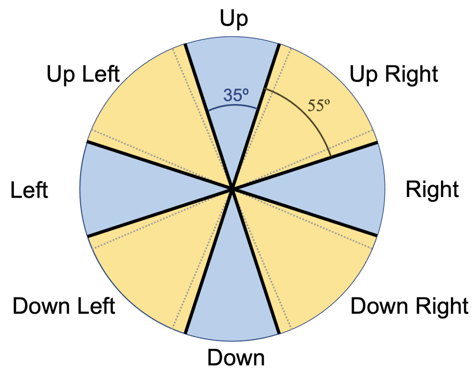

Figure 3.2: Non-uniform circular division for swipe and midair (grey dashed lines represent the original uniform sectors).

events. The screen was blank at all times, no touch feedback or target direction was provided.

Direction is determined by the slope of a linear regression result calculated from dragging trajectory data points. The pilot test showed it was easier to perform the four cardinal up-down and left-right directions, so we use a non-uniform division of the circular space. This non-uniform division is further supported by work finding different gesture movement directions do not share the same accuracy [39, 40]. Specifically, we assigned 35° circular sectors for cardinal directions and 55° circular sectors for diagonal directions (Figure 3.2). These creates a sector angle ratio between cardinal and diagonal directions of 7:11.

3.3 midair directional gestures

A midair gesture is initiated by bringing the thumb and finger together in a pinch. This was detected when the centre of motion tracking markers attached to each finger were less than 50 mm apart. Taking into account typical finger thickness, this is approximately 10 mm between fingertips. To avoid hysteresis between a pinch and open hand, the open hand threshold was 90 mm, approximately 50 mm between fingertips. These thresholds were calculated from mean values and 95% confidence intervals of marker position data from the pilot test. We found this simple pinch detection to be robust, there was no evidence of false positives or false negatives during the study and participants appeared to use it intuitively. Once a pinch was detected, subsequent movement away from the initial position determines the swipe action and direction (Figure 3.3c). The centroid of the two finger markers is projected to a virtual plane parallel to the front edge of the desk (i.e. parallel to the average plane of the monitor). Similar to the touch swipe method, an uninterrupted movement of the projected point along the virtual plane had to exceed a distance and speed threshold, in this case 50 mm with minimum speed of 50 mm/s. These values were computed from pilot test data. The difference in thresholds between midair and touch swipe is due to the relative

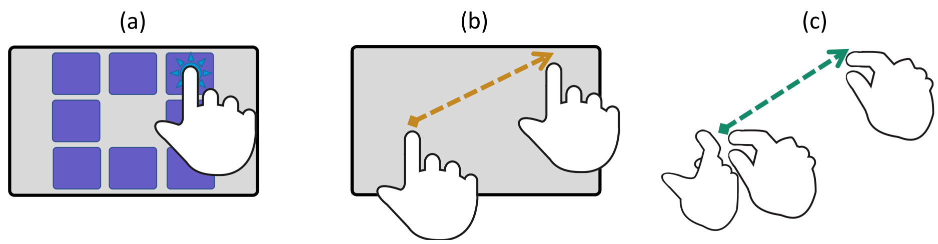

Figure 3.3: Input method conditions illustrating an equivalent *Up Right* direction: (a) TAP used standard tapping on one of 8 targets on a touchscreen centre console; (b) SWIPE used a drag movement along a blank touchscreen in the desired direction; (c) MIDAIR worked by first forming a thumb-to-index pinch, then moving in the desired direction.

sizes of the interaction areas, midair is approximately four times larger than touch swipe gestures (we examine interaction area in Section 5.5). To avoid pinch false positives when gripping the steering wheel, a 30×25 cm rectangle (slightly larger than the 26×23 cm steering wheel) was removed from the virtual plane. When the projected position of finger marker centroid was within this steering wheel area, it was ignored.

4 EXPERIMENT

The goal of this experiment is to directly compare the three input methods just described. For ecological validity and to measure attention required by each method, the Lane Change Task (LCT) is used with the driving simulator.

4.1 PARTICIPANTS

We recruited 18 participants, ages 19 to 41 (23.9 mean, 4.97 std), of which 12 identified as male and 6 as female. Two participants did not possess a valid driving license, but this had no observable effect on their driving performance (discussed in Section 5.4). Participants were recruited using recruitment e-mails throughout the university and word-of-mouth, and received $20 for successful completion of the study.

4.2 APPARATUS

The experiment used the software and hardware described in the previous section. Audio prompts are used for the input task stimulus, played through the Android tablet using the *TextToSpeech* class.

4.3 TASKS

The participants in this study were asked to perform two tasks simultaneously: a long duration driving task while responding to prompts to complete input tasks as quickly as possible.

DRIVING TASK. Following previous work [14, 19, 38], we use the standard Lane Change Task (LCT). The participant drives along a perfectly straight 3-lane roadway in a featureless landscape while changing lanes as instructed by two identical signs on both sides of the road which appear at controlled intervals. The signs are initially grey until within 20 m of the virtual vehicle, after which they indicate the required lane using two 'X's and an arrow (Figure 4.1b). Through-out the driving task, the participant is told to maintain 80 km/h. A speedometer is shown on the right bottom of the driving simulator display (Figure 4.1c) and the virtual vehicle uses an automatic transmission.

The experiment uses one unique "track" with pseudo-randomized lane change instructions (two subsequent signs do not indicate the same lane). The track has 18 signs, spaced approximately 150 m apart along the track. At 80 km/h, the track can be completed in around 2 minutes with pairs of signs appearing every 6 to 7 seconds, and each pair of signs visible for

Figure 4.1: Example of lane change driving task: (a) 3-lane roadway with drive in left lane; (b) pair of signs instructing driver to change to centre lane; (c) speedometer for driver to maintain 80 km/h.

approximately 1 second after which the participant is expected to complete the lane change.

INPUT TASK. As explained in Section 2.2, we use "atomic" input tasks since they have higher internal validity enabling the most direct comparison between input methods. Once the participant begins driving on the track, a spoken audio prompt indicating an input action direction is played every 7.5 seconds (e.g. "up-left"). The participant performs the input in the prompted direction and hears a beep to confirm an input action was recognized. No feedback about the correctness of direction is provided. During the 2-minute driving task, there are 16 input task prompts.

4.4 PROCEDURE

After consenting to participation and answering demographic questions, the markers were attached to the participant's finger and thumb and they were seated in the driving simulator setup. The experiment facilitator explained the lane change task, then the participant practised approximately 5 minutes to get comfortable with the simulator and the lane change task. Then, methods were explained and practised. Before starting blocks of measured tasks for an input method, the facilitator explained and demonstrated the method and the participant practised it without driving at the same time. This practice period lasted until the participants said they were comfortable performing the input method, typically 2 minutes. We reserved the simultaneous driving task and input task for the measured trials to capture unbiased performance.

Note that during this practice portion, the participant received feedback to indicate that the system detected an input action (a "beep" sound), and for midair and swipe, the recognized direction was displayed in text on the main display. This was necessary to learn how to perform the gesture and to gain familiarity with system characteristics, such as minimum movement thresholds and the location of the virtual plane. During the measured portions of experiment, feedback only indicated that a gesture was recognized with the same "beep" sound, but not the direction or whether it was correct. Using minimal feedback during the measured tasks avoided a confound: if the participant focused on feedback, they would be more distracted from driving.

Afterwards, the participant completed the main experiment using the three input methods with the driving and input tasks. They were encouraged to take breaks between input conditions. The main experiment required 15 to 20 minutes to complete. After all three conditions, they completed NASA TLX and answered the following questions:

- *Which gesture style did you like the most? Why?*
- *Do you have any additional comments about your experience?*
- *Do you own a valid driving license? If yes, from where?*

The total session took between 30 to 45 minutes.

4.5 design

This is a within-participants design with one primary independent variable: method with 3 types (tap for touch tapping, swipe for touch swiping, and midair for midair directional gestures). block and direction are secondary independent variables. There were 2 blocks per method, in each block all 8 directions were repeated twice. The order for method was counter-balanced using a balanced Latin square. To arrange the order of trials, in each block, we randomly assigned all 16 directions and used the same block for all participants. As a result, the directions were presented in the same order in each block for all participants. Each participant did 2 blocks for each of the 3 methods.

The measures computed from logs for the driving task are Standard Deviation of Lateral Position (*SDLP*) and *Speed*. *SDLP* is the standard deviation of the vehicle position from the centre of the correct lane (the intended lane 20 m after the signs appearing until the next sign) and *Speed* is the vehicle speed in km/h. Four dependent measures were computed from logs for the input task. *Selection Time* is defined as the duration from when the voice prompt finished until the participant finished performing the interaction (i.e. lifted their finger for tap or exceeded the movement thresholds for swipe and midair). In addition, we divide this overall task time into two component times: *Response Time* is the duration up until the participant started the interaction (i.e. contacted the touch screen for tap and swipe, or formed a pinch for midair); and *Completion Time* to capture the remaining duration when the interaction was performed. *Accuracy* is the percentage of directions that were performed correctly. A fifth measure was calculated using the 4K webcam and Python post-processing described above: *Glances* is defined as the number of off-screen glances that lasted for at least 500 ms during a block. In addition, NASA TLX questionnaire provides 6 subjective measures, each ranging from 0 (very low) to 20 (very high).

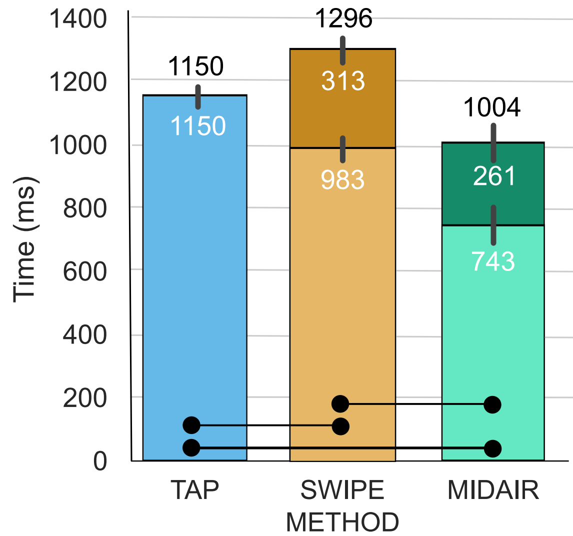

Figure 4.2: *Selection Time* by METHOD, light colours are *Response Time*, dark colours are *Completion Time*. Note *Completion Time* for tap is zero. White labels represent the values of *Response Time* and *Completion Time*, while the black labels show *Selection Time*. **In all graphs, significant differences between levels are shown as horizontal bars and vertical error bars are 95% confidence intervals.**

In summary: 3 METHODS × 2 BLOCKS × 8 DIRECTIONS × 2 repetitions = 96 data points per participant.

4.6 RESULTS

Trials with *Selection Time* greater than three standard deviations were removed from the dataset as outliers, which excluded 53 trials (3%). There were no outliers for other dependent measures. Trials were aggregated by BLOCK producing 108 entries in the final dataset (18 PARTICIPANTS × 3 METHODS × 2 BLOCKS). We use Shapiro-Wilk test to check data residual normality. If normally distributed, we use ANOVA with Tukey HSD post hoc tests, otherwise we use a Friedman test with post hoc pairwise Wilcoxon-signed rank tests.

4.6.1 *Selection Time*

MIDAIR was 22% faster than SWIPE and 13% faster than TAP (Figure 4.2). Between touch gestures, TAP was 11% faster than SWIPE. *Selection Time* was not normally distributed ($W = .96, p < .01$), so we use a Friedman test. METHOD had a significant effect ($\chi^2_F(2) = 25.7$, $p < .001$) with post hoc tests revealing MIDAIR (1004 ms) was faster than TAP (1150 ms) ($p < .05$) and SWIPE (1296 ms) ($p < .001$), and also TAP was faster than SWIPE ($p < .05$). We also examine the component response and completion times that make up the overall selection

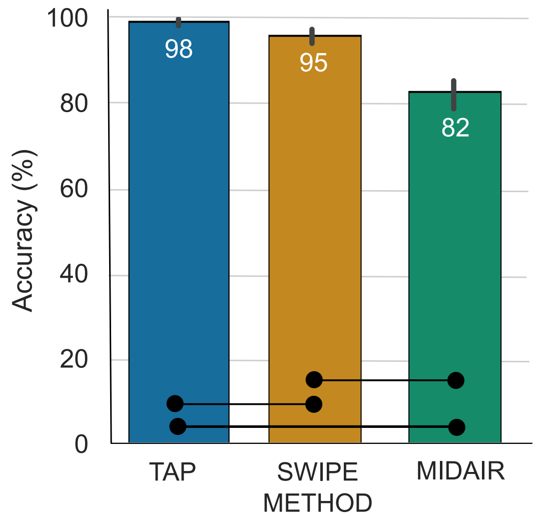

Figure 4.3: *Accuracy* by METHOD.

time using post hoc analysis (white numbers in Figure 4.2). For *Response Time*, MIDAIR was 35% lower than TAP and 24% lower than SWIPE (all $p < .001$). Tap has no *Completion Time* since it is a simple tap, but MIDAIR was 17% lower than SWIPE ($p < .05$).

4.6.2 *Accuracy*

Participants were less accurate with directional MIDAIR gestures, achieving 82% compared to 95% with SWIPE and 98% withTAP. Since data was not normally distributed ($W = .74, p < .001$), we use the Friedman test which revealed a main effect of METHOD on *Accuracy* ($\chi^2_F(2) = 47.2$, $p < .001$). Post hoc analysis showed that MIDAIR is approximately 15% less accurate than both TAP and SWIPE methods (all $p < .001$) and SWIPE 3% less accurate than TAP ($p < .01$).

4.6.3 *SDLP and Driving Speed*

Overall, there was no main effect of METHOD on *SDLP* and *Speed* (Figure 4.4). Using the normally distributed data ($W = .99, p > .05$), a one-way ANOVA found no effect of METHOD on *SDLP* ($F_{2,105} = 2.00$, $p > .05$) or on *Speed* ($\chi^2_F(2) = 2.6, p > .05$).

4.6.4 *Number of off-screen Glances*

Overall, participants were less distracted performing MIDAIR compared to TAP. There was a main effect of METHOD on number of off-screen *Glances* ($\chi^2_F(2) = 18.2$, $p < .001$) in our non-normally distributed data ($W = .78, p < .001$). Post hoc tests show that MIDAIR, with an average number of off-screen

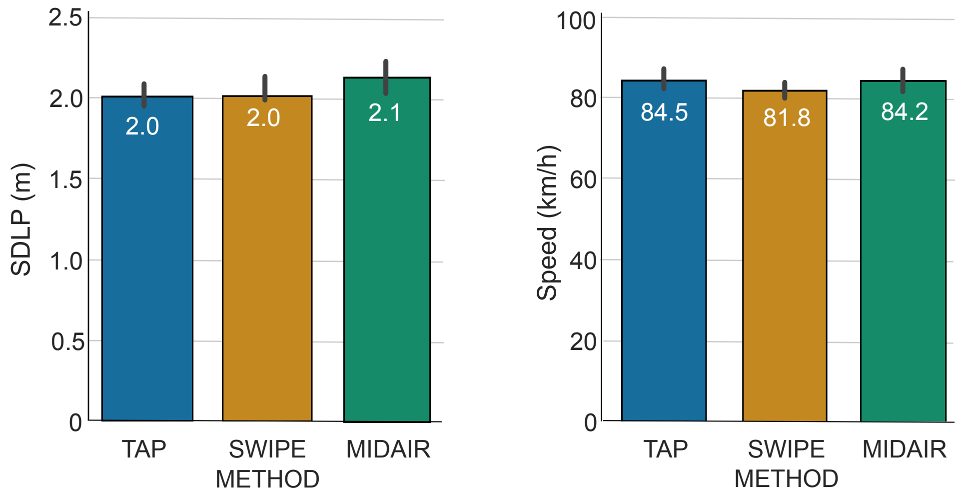

Figure 4.4: *SDLP* and *Speed* by METHOD.

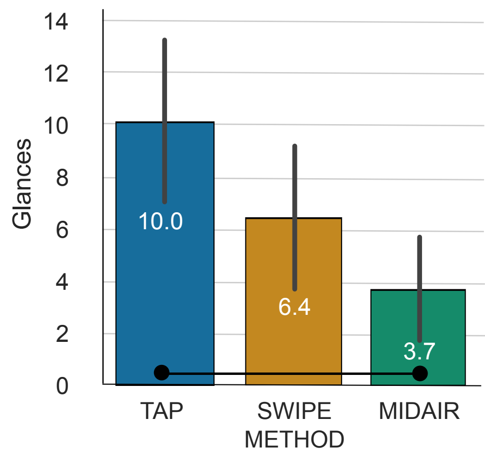

Figure 4.5: Number of off-screen *Glances* by METHOD.

Glances of 3.7, is less distracting than TAP with 10.1 off-screen *Glances* ($p < .05$). However, there was no significant difference between MIDAIR and SWIPE methods. In total MIDAIR has 63% fewer off-screen *Glances* than TAP (Figure 4.5).

4.6.5 *NASA TLX Load Index*

To measure cognitive load, the participant answered a NASA TLX questionnaire after completing tasks all input methods. Overall, all input methods exhibit the same behaviour across all NASA TLX metrics except that midair was perceived as more physically demanding than tap (Figure 4.6). There was a main effect of METHOD on *Physical Demand* ($F_{2,51} = 3.9$, $p < .05$) with post hoc tests showing MIDAIR was perceived as 50% more physically demanding than TAP ($p < .05$).

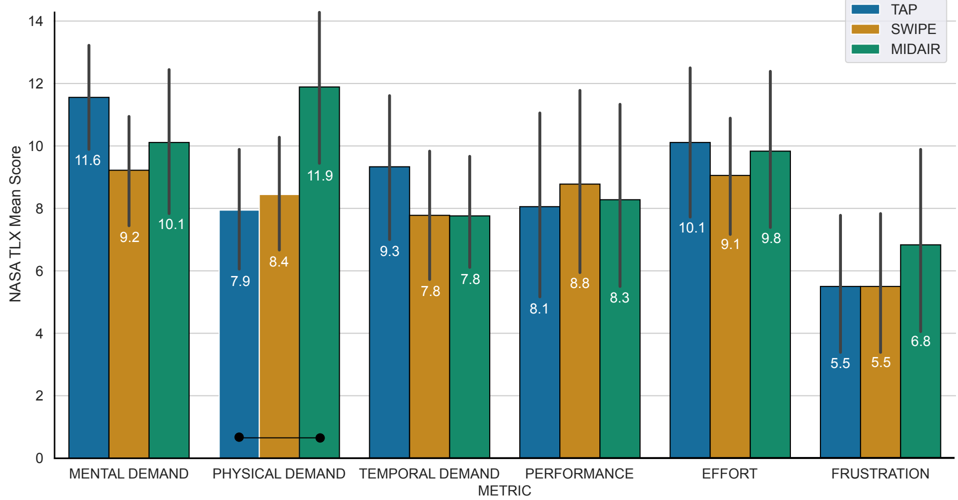

Figure 4.6: *NASA TLX Mean Score* by METHOD.

5 DISCUSSION

After we summarize our findings with a comparison to the most related prior work, we discuss additional insights with secondary analysis, and finally acknowledge the limitations of our work.

Overall, midair directional gestures performed very well. They were 13% to 22% faster than other methods, and achieved a respectable 82% high accuracy considering the demanding 8-direction task. Perhaps most important for our focus on an in-vehicle context, is how midair required fewer off-road glances than touch tapping. This is despite inherent limitations of midair input, no haptic feedback and no well-defined frame-of-reference, as well as greater potential for tracking and recognition issues, and less participant familiarity with this more novel style of interaction. Fatigue is also a potential issue with midair input, our NASA TLX results revealed a single difference across all dimensions: midair input was more physically demanding than touch tapping.

Wu et al. [38] also report a higher workload for midair, but otherwise our results are quite different since they found midair gestures are slower than touch tapping. Considering the limitations, technical challenges, and lack of familiarity with midair input, their results also make sense, especially considering most people are familiar with touchscreen input due to the ubiquity of smartphones and other touch devices. Importantly, our experiment used a distinctly different input task to isolate single atomic actions for input, it is possible that Wu et al.'s more complex, longer duration task was also measuring other factors, not just pure input performance. It is likely that our midair sensing and recognition was more robust, Wu et al. used a Leap Motion consumer depth camera while we used a commercial marker-based motion tracking system with multiple cameras. The difference may also be attributed to our more complex 8-direction (and 8-button) input methods. The time and attention to visually locate and then press one of 8 touchscreen buttons may be more than the time to raise a hand in the air and move it in one of 8 directions. Examining *Response Time* (Figure 4.2) shows that raising the hand before making a midair gesture takes only 65% of the time to reach out and contact the centre console touch screen.

Regarding the need to look at the centre console or not is central to driver safety. Based on recorded number of glances, tapping and swiping both clearly require visual attention to work quickly and reliably. While some participants attempted to perform these two on-screen methods without taking their eyes off the road, they were unable to do so reliably. Participant comments support this, for instance *"It was challenging to get the tapping precisely without looking at the screen, so I typically had to look to make sure I was touching the right button."* [p18] and *"I think I got more distracted and veered off the road when I was doing the tap gestures because I had to look at the tablet."* [p7] To see if how participants glanced affected their error rate, we tested for a correlation between *Glances* and *Error Rate* for each *method*. No statistical correlation was found for any method (all $p > .20$).

5.1 PERFORMANCE DIVERSITY ACROSS INPUT METHODS

We found that participants performed MIDAIR with a greater degree of variation compared to TAP or SWIPE. Figure 5.1 illustrates how combinations of *Error Rate* and *Response Time* for different participants have a wider range of values when performing MIDAIR gestures.

To conduct our analysis, we normalized data using min-max feature scaling, which was necessary in order to have the same effects on data point distance for both *Error Rate* and *Response Time*. Outliers were retained for in order to capture all observed behaviour. Then we plotted the convex hull to show the behaviour space for each METHOD. The resulting plot confirms that the performance space for TAP and SWIPE is relatively smaller (i.e. more consistent) than that of MIDAIR, highlighting how TAP is more accurate and SWIPE is faster. For TAP, there was also a positive correlation between *Error Rate* and *Response Time* ($r(16) = .56, p = .01$), but not for MIDAIR or SWIPE ($p > .20$) On the other hand, MIDAIR gestures carry a wider range of performance with more density for lower *Response Time*. By investigating potential causes of this performance diversity, it may be possible to identify ways to improve MIDAIR gestures in order to interact with in-vehicle centre console systems with the least amount of distraction.

5.2 PERFORMANCE DIVERSITY WITHIN MIDAIR INPUT

To investigate performance diversity within midair input, we cluster the normalized *Error Rate* and *Response Time* points representing the 18 participants. To find the number of clusters, we use the Elbow method [17]. First, Nearest Neighbours calculates point distances, and reaches an optimal $\epsilon = 0.12$. This is fed to the DBSCAN algorithm to determine k, the optimal number of clusters. The method found $k = 3$, which we use with the K-Means algorithm to cluster in the participant points.

The results shows how participants can be classified into three clusters based on their performance (see Figure 5.2). The green cluster of 13 participants have lower *Response Time* and low *Error Rate*. They learned how to balance speed and error, and overall managed to perform with some expertise. The blue cluster of 3 participants have slower *Response Time* but low *Error Rate*. They focused on carefully performing the directional action even if it took a much longer time. These participants are likely less confident about their actions, needing more practice or additional instructions to reach a more optimal performance. Finally, the red cluster of 2 participants have very high *Error Rate*. This indicates either a priority of speed over accuracy, great difficulty in performing the technique, or general lack of care when performing the task. Note that we performed the same analysis for TAP and

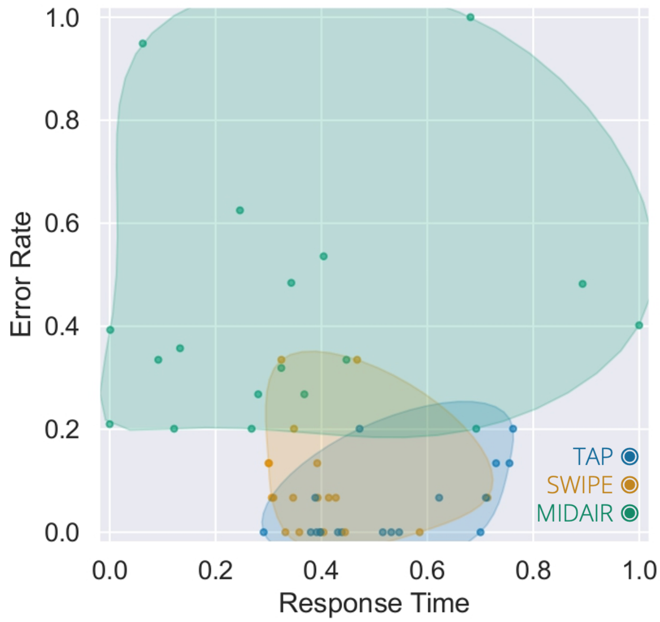

Figure 5.1: Performance behaviour based on normalized *Response Time* and *Error Rate*, each point represents mean of participant data for a METHOD: (green) MIDAIR; (blue) TAP; and (yellow) SWIPE. The convex hull indicates a "performance space" for a method.

SWIPE methods, and those participant data points were each clustered into 1 group. This suggests a homogeneity of performance with tap and swipe, and further confirms additional performance diversity with midair.

5.3 DRIVING EXPERIENCE

Out of 18 participants, P10 and P17 did not hold a valid driving license and had little driving experience. Our analysis below suggests this had no noticeable effect on their performance in the driving task or the input task.

We use the 95% confidence interval (95% CI) for the 16 participants who held a valid drivers licence, and examine if values for the two participants without a drivers license, P10 and P17, fall into that interval. For *SDLP*, the mean is 2.0 m (SD = 0.2) with 95% CI [1.6, 2.4]: both P10 (2.1 m) and P17 (2.2 m) are within. For *Speed*, the mean is 83.3 km/h (SD = 6.0) with 95% CI [71.3, 95.3]: both P10 (82.3 km/h) and P17 (88.0 km/h) fall within. We believe this is because our DRIVING TASK only required simple vehicular control rather than interactions with other vehicles and pedestrians requiring knowledge of the rules of the road and more driving decisions.

We also examine input tasks using mean times combining all three methods. For *Accuracy*, the mean is 92.4 (SD = 10.2) with 95% CI [72.0, 100]: both P10 (90.2) and P17 (93.0) fall within. For *Response Time*, the mean is 940 ms (SD = 359) with 95% CI [222, 1658]: both P10 (1199) and P17 (954) fall within.

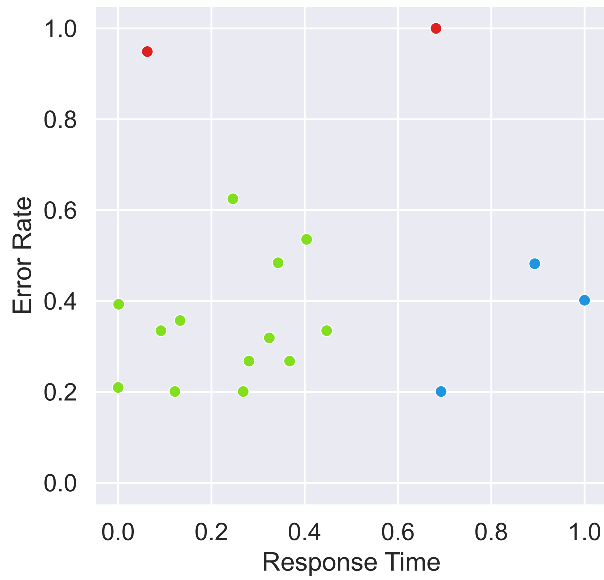

Figure 5.2: Clustering of participant data for MIDAIR using normalized *Response Time* and *Error Rate*. Colours indicate three groups of participants.

It could have been possible that non-drivers were more focused on driving which could have affected their input performance, but the analysis does not show this.

5.4 MIDAIR AND SWIPE KINEMATICS

We investigate how optimal the thresholds were for the midair and swipe gestures, and find the interaction space size that participants typically used. Participants performed swipe gestures with a mean distance of 85 mm (SD = 26) and for midair gestures, a mean distance of 114 mm (SD = 26). These both well exceed the minimum distance thresholds required by the recognizers, 12 mm and 50 mm respectively. In terms of speed, participants performed touch swipes with a mean speed of 321 mm/s (SD = 157) and midair swipes with a mean speed of 451 mm/s (SD = 267). Again, these well exceed the minimum recognizer thresholds of 12 mm/s and 50 mm/s respectively. This suggests the threshold distances and speeds used by the recognizer were more than adequate. It is further observed that the space used to perform gestures follows a similar pattern. We plot the start and end points of each gesture, and use 95% confidence intervals to define the smallest bounding rectangle. We found this rectangle area was 1480 cm^2 (418 mm × 354 mm) for MIDAIR and 331 cm^2 (230 mm × 144 mm) for SWIPE. The rectangle for MIDAIR is approximately 4.47 times larger than SWIPE, confirming the 4.17 (50 mm / 12 mm) ratio between recognizer thresholds was reasonable.

5.5 MIDAIR ACCURACY WITH FEWER DIRECTIONS

In the experiment, participants performed eight MIDAIR directional gestures (up, down, right, left, up-right, up-left, down-right, and down-left), with an overall accuracy of 82% across all directions. However, if the data is restricted to the three groups of gestures, up/down, right/left, and the four main directions, up/down/right/left, then the accuracy improves to 98%, which is similar to that of TAP and SWIPE techniques (Figure 5.3). First, the original slopes of linear regression of midair gesture trajectories were taken. Then, we divided the circular space to identical circular sectors for each direction. Each circular sector was 180° for right/left and up/down and 90° for up/down/right/left. Considering Figure 5.3, it can be concluded that performing MIDAIR gestures toward up-left, up-right, down-left, and down-right have lower accuracy than performing gestures toward the four main directions.

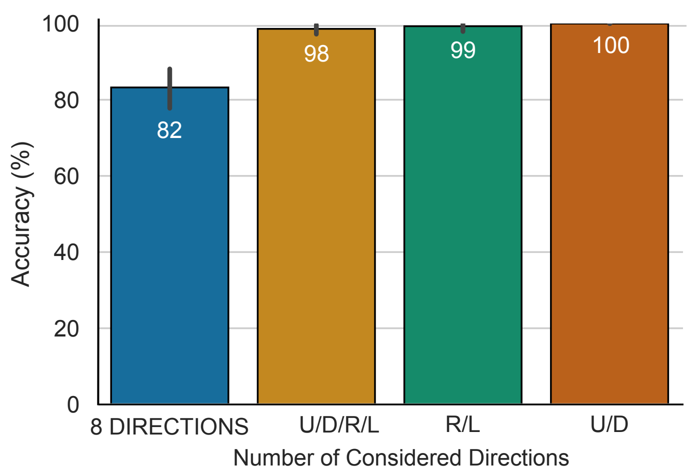

Figure 5.3: Comparison of MIDAIR accuracy when reducing the number of considered directions.

5.6 HAPTIC FEEDBACK AND MIDAIR GESTURES

Due to how barehand midair input does not involve a device or contact with a surface, it lacks natural haptic feedback, which can be considered its limitation. However, there is a large body of work developing and demonstrating ultrasound methods to stimulate a hand in midair to produce synthetic haptic feedback [30]. Midair user interface components, midair gesture feedback, and virtual object representations are primary input-related use cases for this technology. Several automotive manufacturers demonstrated using ultrasound haptics for midair touch input [16] including the "HoloActive Touch" concept demo by BMW which combines the technique with a holographic midair touch display [8]. UltraLeap sells systems for detecting midair gestures and delivering touch-less haptic feedback using arrays of ultrasound emitters [6]. One of their target markets is in-vehicle control, where they demonstrated working systems using gestures like grab-release and hand-twist.

We discussed Harrington et al. [12] more generally in Related Work, here we focus on their results regarding haptics. They found conditions with ultrasound haptic feedback were more accurate, had shorter selection times, and fewer off-road glances. Our directional midair gestures are similar to the open-handed midair gesture they used to manipulate the slide-bar in their task. Their results suggest ultrasound haptic feedback to indicate successful gestures, whether a continuous gradual force or a sudden pulse, could facilitate more accurate, and potentially safer, interactions. Shakeri, Williamson, and Brewster [33] also studied ultrasound haptics for in-vehicle input, comparing uni-modal and multi-modal haptic variations of a midair gesture. Their results found multi-modal ultrasound feedback decreases the time drivers looked away from the road, a further indication that ultrasound haptics can improve midair gestures in a vehicle setting. Our study focused on midair gestures in a more easily deployable form for manufacturers, without the added complexity and cost of ultrasound haptic feedback. However, we anticipate this type of haptic feedback could provide similar improvements to our midair results.

5.7 APPLICATION SCENARIOS

When four or more directional gestures are distributed radially (as we did in our study), there is a natural mapping to circular "pie" menu interfaces [5]. The in-vehicle system commands triggered by such a pie menu could be contextual based on the system state. For example, when there is an incoming call, the menu items could be 'Answer', 'Reject', 'Mute', and 'Call Back' (Figure 5.4a). Or, when playing media, the menu items could be 'Next', 'Previous', 'Volume Up', and 'Volume Down'. A pie menu can also be made hierarchical, where a directional movement on the first level triggers a second level with related commands. For example, the first level could have items for 'Navigation', 'Climate Control', 'Window Control', and 'Vehicle Functions'. Selecting 'Navigation' would move to the second level where items control the navigation system, such as 'Pause', 'Repeat Instructions', 'Zoom In', and 'Zoom Out'. One advantage for a directional pie menu is that it can be easily transformed into a "marking menu", where the first and second level in a hierarchy can be selected in two swift directional movements [18].

Another appropriate interface for directional input is a linear menu. A large number of menu items can be scrolled through using up and down directional gestures, with a right gesture to 'select' the current item and a left gesture to 'go back' to a previous level or dismiss the menu (Figure 5.4b). This could be applied to vehicle systems like selecting a media item among a larger collection (e.g. a playlist or a station) or setting one of several climate control modes (e.g. air, heat, defrost, etc.).

These menu examples could be implemented with directional gestures performed in midair or with a touch swipe. This could enable control by either method as desired by the user or the driving situation. Or, given the different performance characteristics, some menus could be accessed with touch swipes and some with midair input. For example, frequent actions, such as a menu of windscreen wiper functions (e.g. 'Pulse', 'High', 'Low', 'Clean') could be accessed with midair input, and less frequent menus with touch swipes, such as media selection or vehicle settings.

Our study and the examples above are specific to in-vehicle systems, but the general style of directional gestures and associated menu interfaces could be used in other settings, such as controlling a smart television when seated on a sofa (Figure 5.4c). However, a vehicle cabin has the advantage of a very stable configuration: a very specific seated user location, a well-defined space for tracking midair gestures, and a nearby touch surface.

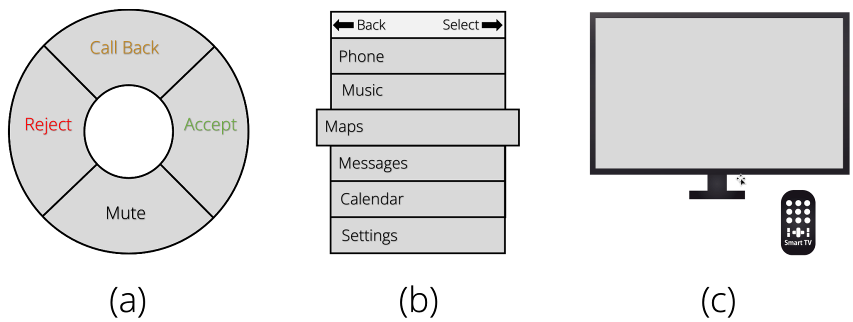

Figure 5.4: Potential application scenarios: (a) pie menu; (b) linear menu; (c) controlling a smart TV.

5.8 LIMITATIONS AND FUTURE WORK

A desktop driving simulator in a lab environment is used to evaluate midair input for in-vehicle control in a similar manner to previous work (Table 2.1). We also use the standard lane change task, as was done in two other previous works [14, 38]. This reflects a trade-off between realism and experiment control. Conducting our study with the same driving task in a real car on a track would undoubtedly intensify the driver's attention, and perhaps reveal changes in how they perform the different input methods. One approach is for participants to perform gestures as a passenger to understand the impact of vehicle motion [26], but the logistics and safety for a study where participants simultaneously drive a real car would be very challenging. A more conservative next step is to replicate this experiment in a more high-fidelity and immersive vehicle simulator. Similarly, using a more open-ended "driving in traffic" task like Angelini et al. [3] or Graichen, Graichen, and Krems [10] would likely change the intensity of the driving task relative to the input task. However, as in any experiment, open-ended tasks are difficult to control, which reduces the internal validity necessary for unbiased comparisons between input conditions.

In order to illustrate the advantages and challenges involved in conducting this type of experiment in a real car, we discuss the work of Ng and Brewster [26]. In their paper, three scrolling input techniques were compared, none of which were midair, but there is a shared goal of comparing in-vehicle input techniques. The experiment was conducted in a moving vehicle driving through a quiet neighbourhood of a city. However, due to safety regulations, the car was driven by the experiment facilitator while the participant sat in the passenger seat holding a mock steering wheel between performing the scrolling methods. The results showed that a direct touch method was less accurate though faster than on-screen buttons or pressure input. They also performed the experiment with the participant driving an actual car in another study [25]. The driving task was in town, carriageways, and highways scenarios in presence of the experimenter for safety which resulted in lower accuracy and longer selection time. In comparison to our focus, the experiment setting and open-ended task prevented a controlled measure of driver distraction.

Although not specifically mentioned by Ng and Brewster, disturbance caused by physical movements in a real car could effect input method performance. The suspension in modern cars maintains a reasonably smooth ride over well maintained streets, but rough or twisting roads would likely be a problem for any kind of input method. We hypothesize that midair input may be more robust to moderate disturbances compared to input performed on a touchscreen console. The touchscreen vibrates and moves with the car which transforms input into a moving target. However, with midair input there is no car-anchored target and the shoulder and elbow joints will dampen some hand movement caused by car movement disturbances. As an anecdotal example, when encountering very rough roads, a person may remove a full cup of coffee from the cup holder and hold it in midair in order to reduce the motion and prevent spilling. Technical prediction tools such as Predictive Touch [1] have been proposed to remove the need for a direct touch and predict targets in initial pointing states, mitigating the perturbations effect.

Projecting the MIDAIR gestures to a different virtual plane changes perceived accuracy. In our study, we projected the MIDAIR gestures to the average plane of the monitor since naturally, people intend to perform gestures in that plane. However, projection on an angled plane may modify the projected data points and possibly alters the final accuracy. Further mathematical averaging of the mean plane that participants made or machine learning analysis are possible to find the optimum plane. As the optimum plane is logically close to our current plane, we may reach subtle improvements.

In the experiment, we asked general demographic questions of our participants. Providing additional demographic information regarding users' driving abilities, the amount and manner in which they drive, as well as their experience with technology, including video games, may provide us with additional insight into why some participants are highly successful in performing MIDAIR gestures and some are not, compared to the other two interaction methods, which have consistent performance across all participants. Although in general, our participants had slightly better performance for the last few blocks of the experiment (0.5% higher *Accuracy* and 22.3% lower *Selection Time*), suggesting that practice could be a crucial component to consider, especially in the case of such a new interaction method as this. A follow-up study may examine how users can adopt this new midair input technique, since it is less distracting than other commercially implemented input methods (for example, on-screen swipes and taps).

There is also a trade off between creating more advanced midair gestures to reach a higher functionality and keeping the driver's distraction at a minimum point. Future works can find a balance for this issue by proposing task specific gesture sets such as assigning different gesture methods for various part of the in-vehicle interaction. With more progress in autonomous driving, the problem of driver's safety and distraction is changing overtime, leaving a good chance for midair gestures to take action in more areas.

6 CONCLUSION

We compared and evaluated the performance of three functionally equivalent methods for in-vehicle input: 8-direction midair gestures, 8-direction touch swipes, and 8-button touch taps. Midair gestures were faster but less accurate compared to touch taps and swipes. In practice, midair achieves a respectable accuracy of 82%, and we show this increases to 98% if only 2 or 4 directions are used. Although touch swipe gestures were slower than midair, they were very accurate. Critically for the context of in-vehicle control, midair gestures were less distracting than touch screen taps in terms of fewer off-road glances.

With additional analysis, we showed surprising diversity for midair gesture performance among our participants. Essentially, some people were very good and some people were not. This suggests a need for techniques to teach people how to perform midair gestures effectively and for automated methods to personalize recognition parameters so midair gestures can adapt to the capabilities and natural tendencies of specific drivers. But overall, our results suggest that manufacturers should consider midair gestures more closely to complement or even replace conventional target touch tapping on the centre console.

BIBLIOGRAPHY

[1] Bashar I. Ahmad, Simon J. Godsill, Patrick M. Langdon, and Lee Skrypchuk. "Predictive Touch: A Novel HMI Technology for Intelligent Displays in Automotive." In: *Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '18. Toronto, ON, Canada: Association for Computing Machinery, 2018, 259–260. ISBN: 9781450359474. DOI: 10.1145/3239092.3267103. URL: https://doi.org/10.1145/3239092.3267103.

[2] Leonardo Angelini, Jürgen Baumgartner, Francesco Carrino, Stefano Carrino, Maurizio Caon, Omar Abou Khaled, Jürgen Sauer, Denis Lalanne, Elena Mugellini, and Andreas Sonderegger. "A Comparison of Three Interaction Modalities in the Car: Gestures, Voice and Touch." In: *Actes de La 28ième Conference Francophone Sur l'Interaction Homme-Machine*. IHM '16. Fribourg, Switzerland: Association for Computing Machinery, 2016, 188–196. ISBN: 9781450342438. DOI: 10.1145/3004107.3004118. URL: https://doi.org/10.1145/3004107.3004118.

[3] Leonardo Angelini, Francesco Carrino, Stefano Carrino, Maurizio Caon, Omar Abou Khaled, Jürgen Baumgartner, Andreas Sonderegger, Denis Lalanne, and Elena Mugellini. "Gesturing on the Steering Wheel: A User-Elicited Taxonomy." In: *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '14. Seattle, WA, USA: Association for Computing Machinery, 2014, 1–8. ISBN: 9781450332125. DOI: 10.1145/2667317.2667414. URL: https://doi.org/10.1145/2667317.2667414.

[4] indiGO Auto Group. *BMW's Innovative Gesture Control Technology Sets the Industry Standard*. BMW. 2021. URL: https://www.indigoautogroup.com/blogs/1364/automotive-news/bmws-innovative-gesture-control-technology-sets-the-industry-standard/ (visited on 09/13/2022).

[5] J. Callahan, D. Hopkins, M. Weiser, and B. Shneiderman. "An Empirical Comparison of Pie vs. Linear Menus." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '88. Washington, D.C., USA: Association for Computing Machinery, 1988, 95–100. ISBN: 0201142376. DOI: 10.1145/57167.57182. URL: https://doi.org/10.1145/57167.57182.

[6] Matt Corrall. *Design guide for gesture control in automotive HMI*. Ultraleap. 2020. URL: https://www.ultraleap.com/company/news/blog/gesture-control-automotive-hmi/ (visited on 09/13/2022).

[7] Zhitong Cui, Hebo Gong, Yanan Wang, Chengyi Shen, Wenyin Zou, and Shijian Luo. "Enhancing Interactions for In-Car Voice User Interface with Gestural Input on the Steering Wheel." In: *13th International Con-*

ference on Automotive User Interfaces and Interactive Vehicular Applications. AutomotiveUI '21. Leeds, United Kingdom: Association for Computing Machinery, 2021, 59–68. ISBN: 9781450380638. DOI: 10.1145/3409118.3475126. URL: https://doi.org/10.1145/3409118.3475126.

[8] Darrell Etherington. *BMW's HoloActive Touch in-car interface offers tactile feedback on a floating display*. en-US. Dec. 2016. URL: https://techcrunch.com/2016/12/15/bmws-holoactive-touch-in-car-interface-offers-tactile-feedback-on-a-floating-display/ (visited on 12/08/2022).

[9] Hessam Jahani Fariman, Hasan J. Alyamani, Manolya Kavakli, and Len Hamey. "Designing a User-Defined Gesture Vocabulary for an in-Vehicle Climate Control System." In: *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. OzCHI '16. Launceston, Tasmania, Australia: Association for Computing Machinery, 2016, 391–395. ISBN: 9781450346184. DOI: 10.1145/3010915.3010955. URL: https://doi.org/10.1145/3010915.3010955.

[10] Lisa Graichen, Matthias Graichen, and Josef F Krems. "Evaluation of gesture-based in-vehicle interaction: user experience and the potential to reduce driver distraction." In: *Human factors* 61.5 (2019), pp. 774–792.

[11] Arman Hafizi, Jay Henderson, Ali Neshati, Wei Zhou, Edward Lank, and Daniel Vogel. "In-Vehicle Performance and Distraction for Midair and Touch Directional Gestures." In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Hamburg, Germany: Association for Computing Machinery, 2023. ISBN: 9781450394215. DOI: 10.1145/3544548.3581335. URL: https://doi.org/10.1145/3544548.3581335.

[12] Kyle Harrington, David R. Large, Gary Burnett, and Orestis Georgiou. "Exploring the Use of Mid-Air Ultrasonic Feedback to Enhance Automotive User Interfaces." In: *Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '18. Toronto, ON, Canada: Association for Computing Machinery, 2018, 11–20. ISBN: 9781450359467. DOI: 10.1145/3239060.3239089. URL: https://doi.org/10.1145/3239060.3239089.

[13] Waseem Hassan, Ahsan Raza, Muhammad Abdullah, Mohammad Shadman Hashem, and Seokhee Jeon. "HapWheel: Bringing In-Car Controls to Driver's Fingertips by Embedding Ubiquitous Haptic Displays into a Steering Wheel." In: *IEEE Transactions on Intelligent Transportation Systems* 23.10 (2022), pp. 18526–18534. DOI: 10.1109/TITS.2022.3160496.

[14] Renate Häuslschmid, Benjamin Menrad, and Andreas Butz. "Freehand vs. micro gestures in the car: Driving performance and user experience." In: *2015 IEEE Symposium on 3D User Interfaces (3DUI)*. New York, NY, USA: Institute of Electrical and Electronics Engineers, 2015, pp. 159–160. DOI: 10.1109/3DUI.2015.7131749.

[15] Jingun Jung, Sangyoon Lee, Jiwoo Hong, Eunhye Youn, and Geehyuk Lee. "Voice+Tactile: Augmenting In-Vehicle Voice User Interface with Tactile Touchpad Interaction." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–12. ISBN: 9781450367080. DOI: 10.1145/3313831.3376863. URL: https://doi.org/10.1145/3313831.3376863.

[16] Haje Jan Kamps. *Look out for Ultrahaptics in a car near you very soon*. en-US. Jan. 2017. URL: https://techcrunch.com/2017/01/03/ultrahaptics-launching/ (visited on 12/08/2022).

[17] Veer Kumar. *Tutorial for DBSCAN Clustering in Python Sklearn*. Machine Learning Knowledge. 2021. URL: https://machinelearningknowledge.ai/tutorial-for-dbscan-clustering-in-python-sklearn/ (visited on 12/12/2022).

[18] Gordon Kurtenbach and William Buxton. "User Learning and Performance with Marking Menus." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '94. Boston, Massachusetts, USA: Association for Computing Machinery, 1994, 258–264. ISBN: 0897916506. DOI: 10.1145/191666.191759. URL: https://doi.org/10.1145/191666.191759.

[19] Stefan Mattes. "The lane-change-task as a tool for driver distraction evaluation." In: *Quality of Work and Products in Enterprises of the Future* 57.60 (2003), p. 7.

[20] Stefan Mattes and Anders Hallén. "Surrogate distraction measurement techniques: The lane change test." In: *Driver distraction: Theory, effects, and mitigation* 1 (2009), pp. 107–121.

[21] 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. Seattle, WA, USA: Association for Computing Machinery, 2014, 1–6. ISBN: 9781450307253. DOI: 10.1145/2667239.2667280. URL: https://doi.org/10.1145/2667239.2667280.

[22] Keenan R. May, Thomas M. Gable, and Bruce N. Walker. "Designing an In-Vehicle Air Gesture Set Using Elicitation Methods." In: *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '17. Oldenburg, Germany: Association for Computing Machinery, 2017, 74–83. ISBN: 9781450351508. DOI: 10.1145/3122986.3123015. URL: https://doi.org/10.1145/3122986.3123015.

[23] Keenan May, Thomas M. Gable, Xiaolong Wu, Ruta R. Sardesai, and Bruce N. Walker. "Choosing the Right Air Gesture: Impacts of Menu Length and Air Gesture Type on Driver Workload." In: *Adjunct Proceedings of the 8th International Conference on Automotive User Interfaces*

and Interactive Vehicular Applications. AutomotiveUI '16 Adjunct. Ann Arbor, MI, USA: Association for Computing Machinery, 2016, 69–74. ISBN: 9781450346542. DOI: 10.1145/3004323.3004330. URL: https://doi.org/10.1145/3004323.3004330.

[24] Mercedes-Benz. *MBUX in the New S-Class | Next Generation of the Infotainment System*. Mercedes-Benz. 2021. URL: https://www.youtube.com/watch?v=7VTD4MzcI3w (visited on 09/13/2022).

[25] Alexander Ng, Stephen A. Brewster, Frank Beruscha, and Wolfgang Krautter. "An Evaluation of Input Controls for In-Car Interactions." In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. CHI '17. Denver, Colorado, USA: Association for Computing Machinery, 2017, 2845–2852. ISBN: 9781450346559. DOI: 10.1145/3025453.3025736. URL: https://doi.org/10.1145/3025453.3025736.

[26] Alexander Ng and Stephen Brewster. "An Evaluation of Touch and Pressure-Based Scrolling and Haptic Feedback for In-Car Touchscreens." In: *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. AutomotiveUI '17. Oldenburg, Germany: Association for Computing Machinery, 2017, 11–20. ISBN: 9781450351508. DOI: 10.1145/3122986.3122997. URL: https://doi.org/10.1145/3122986.3122997.

[27] Eshed Ohn-Bar and Mohan Manubhai Trivedi. "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations." In: *IEEE transactions on intelligent transportation systems* 15.6 (2014), pp. 2368–2377.

[28] Francisco Parada-Loira, Elisardo González-Agulla, and José L. Alba-Castro. "Hand gestures to control infotainment equipment in cars." In: *2014 IEEE Intelligent Vehicles Symposium Proceedings*. New York, NY, USA: Institute of Electrical and Electronics Engineers, 2014, pp. 1–6. DOI: 10.1109/IVS.2014.6856614.

[29] T. K. Philip Hwang, Yao-Tin Huang, and Pin-Chieh Kuo. "Employing Natural Finger Positioning Strategy for Improving Blind-Positioning of Steering Wheel Mounted Switches." In: *Advances in Human Aspects of Transportation*. Ed. by Neville Stanton. Cham: Springer International Publishing, 2020, pp. 85–91. ISBN: 978-3-030-50943-9.

[30] Ismo Rakkolainen, Euan Freeman, Antti Sand, Roope Raisamo, and Stephen Brewster. "A survey of mid-air ultrasound haptics and its applications." In: *IEEE Transactions on Haptics* 14.1 (2020), pp. 2–19.

[31] Graham Rapier. *German court ruled Tesla's touchscreen windshield wiper controls are the same as texting while driving*. en-US. 2020. URL: https://www.businessinsider.com/tesla-touchscreen-wiper-controls-akin-texting-driving-german-court-2020-8 (visited on 07/07/2022).

[32] Thomson Reuters ·. *Biggest trend in new car technology? Super-sized screens | CBC News*. CBC. Canadian Broadcasting Corporation. Jan. 10, 2019. URL: https://www.cbc.ca/news/science/ces-trend-48-inch-screen-car-1.4972574 (visited on 09/08/2022).

[33] 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. Toronto, ON, Canada: Association for Computing Machinery, 2018, 1–10. ISBN: 9781450359467. DOI: 10.1145/3239060.3239081. URL: https://doi.org/10.1145/3239060.3239081.

[34] BMW USA. *Pinch And Drag Gesture | BMW Genius How-To*. BMW. 2015. URL: https://www.youtube.com/watch?v=jzqoTu_q78c (visited on 09/13/2022).

[35] Joris C Verster and Thomas Roth. "Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP)." In: *International journal of general medicine* 4 (2011), p. 359.

[36] Jacob O. Wobbrock, Meredith Ringel Morris, and Andrew D. Wilson. "User-Defined Gestures for Surface Computing." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Boston, MA, USA: Association for Computing Machinery, 2009, 1083–1092. ISBN: 9781605582467. DOI: 10.1145/1518701.1518866. URL: https://doi.org/10.1145/1518701.1518866.

[37] Huiyue Wu, Yu Wang, Jiayi Liu, Jiali Qiu, and Xiaolong Luke Zhang. "User-defined gesture interaction for in-vehicle information systems." In: *Multimedia Tools and Applications* 79.1 (2020), pp. 263–288.

[38] Shawn Wu, Thomas Gable, Keenan May, Young Mi Choi, and Bruce N Walker. "Comparison of surface gestures and air gestures for in-vehicle menu navigation." In: *Archives of Design Research* 29.4 (2016), pp. 65–80.

[39] Yiqi Xiao, Ke Miao, and Chenhan Jiang. "Mapping Directional Mid-Air Unistroke Gestures to Interaction Commands: A User Elicitation and Evaluation Study." In: *Symmetry* 13.10 (2021), p. 1926.

[40] Justin Young, Michael Lin, Alexander Bick, Ammar Sarwar, and Jack Dennerlein. "Gestural Workspaces for Computer Interaction: Configuration and Performance." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 57.1 (2013), pp. 424–428. DOI: 10.1177/1541931213571092. eprint: https://doi.org/10.1177/1541931213571092. URL: https://doi.org/10.1177/1541931213571092.