Multimodal Feedback for Mid-Air Gestures during the Lane Change Task

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

ABSTRACT

An increasing number of cars now allow drivers to interact with the in-vehicle systems using mid-air gestures. Whilst intended to reduce driver distraction by replacing visually-demanding button interfaces with 'simple' hand movements, mid-air gestures are not straightforward and usability issues could make them just as distracting. We investigate the use of bimodal gesture feedback across a number of modalities, to support gesture input while performing the Lane Change Task. These include audio and tactile feedback and visual cues in the periphery. Our results show that the bimodal combination of Audio-Visual has significantly higher impact on lane deviation and eyes-off-the-road time compared to the Audio-Tactile, Audio-Peripheral, and Tactile-Peripheral combinations. Our non-visual feedback techniques did not affect driving performance, gaze behaviour, gesturing performance, nor increased mental demand, suggesting they can support gesture input without increasing distraction and putting safety at risk. By distributing interaction feedback across multiple modalities, drivers can focus more on the primary task of driving.

ACM Classification Keywords

H.5.2. User Interfaces: Haptic I/O

Author Keywords

Automotive HCI; Driving simulator study; Lane deviation; Mid-air gestures; Multimodal feedback.

INTRODUCTION

Many car manufacturers are introducing mid-air gestures as a new interaction modality because they allow drivers to operate in-car systems without reaching for physical controls. This has the potential to minimise distraction [36] and improve safety [32] by replacing complex actions (e.g., precisely selecting buttons on touchscreens) with simple hand movements (e.g., swiping the hand in mid-air). However, gestures are unfamiliar to most drivers, which may negatively impact driving performance and mental workload [13, 30]. Gestures are also prone to usability challenges which may increase frustration and workload. Users need to know *where* to perform gestures [9, 10]; they also need to know *how* to perform them, and the mapping of hand movements to actions is not always clear. Users may experience uncertainty while gesturing and the lack of tactile feedback is a core part of this [4, 7, 27], which can lead to drivers taking their eyes off the road [14]. However, good feedback can overcome these issues.

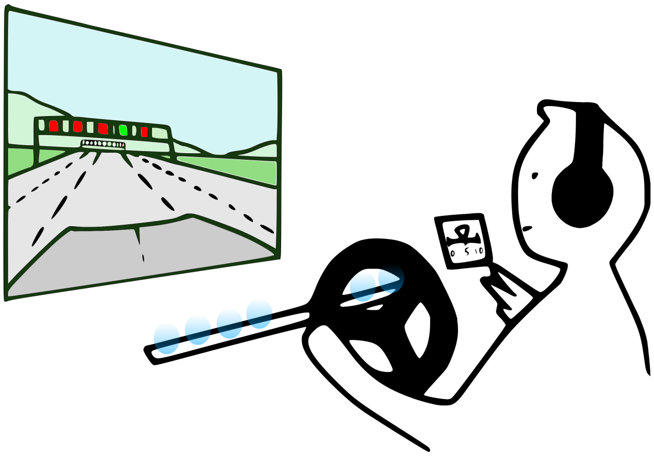


Figure 1. Experiment set-up. From left to right: 5 lane motorway projected onto wall; LED strip for Peripheral lights feedback; steering wheel with Cutaneous Push feedback; 8 inch screen for Visual feedback; headphones for Audio feedback. In this scenario, the participant is gesturing Victory with their right hand above the gear stick.

Current mid-air gesture systems in cars only give limited and mainly visual feedback[1,2]. If drivers need to take their eyes off the road to understand their interactions with the car, then the benefits of mid-air gestures are not being fully realised. Non-visual feedback is ideal for in-car gestures: visual attention can remain on the road whilst information about secondary tasks (i.e., interacting with the in-car systems) is offloaded to other sensory modalities [43]. In this work, we investigate the use of auditory and haptic feedback for in-car gestures, as well as novel visual feedback in the driver's visual periphery [36], intended to allow them to keep their eyes on the road [22].

Others have started to explore the use of non-visual feedback for in-car gestures. Shakeri *et al.* [36] found that auditory cues

---

[1] BMW 7 Series (G11): https://en.wikipedia.org/wiki/BMW_7_Series_(G11)

[2] VW Golf R: http://www.volkswagen.co.uk/new/golf-vii-pa/explore/r

Submitted to AutoUI 2019.

and haptic feedback from the steering wheel allowed drivers to keep their eyes on the road, although since feedback was given in a single modality, feedback content was limited. The limited amount of information was not as effective as visual feedback shown on the console screen, which required eyes off the road. Ultrasound haptic feedback is a new but active feedback mechanism for cars, and has been used for mid-air buttons [14, 33], sliders [14] and dials [12]. This feedback can reduce eyes-off-the-road time (EORT) [14], although drivers still relied on visual feedback for spatial information (e.g., where buttons were located). Shakeri *et al.* [37] also looked at ultrasound haptic feedback, this time for simple gestures and poses, rather than controls like sliders. Ultrasound haptic feedback performed well, although was more effective when paired with audio or visual feedback. These works show the potential of novel feedback modalities for in-car gestures, using new technologies to discourage drivers from looking at the car's console screen.

We build on these by further investigating non-visual feedback for mid-air gestures. We focus on bimodal feedback, in particular, because the previous work suggests that individual feedback modalities only offer limited support for interaction whilst driving. Multimodal information display can be beneficial as information is distributed across multiple sensory modalities. This is especially important for drivers, as driving is a cognitively demanding task [41, 16, 29, 36]. By offloading interaction feedback to other modalities, we reduce the demands of interacting with the in-car systems and increase the noticeability of the feedback.

In this work, we focus on simple mid-air hand movements, likes swipes and poses, rather than user interface controls like buttons and sliders used in previous in-car research [12, 14, 33]. Simple movements are more representative of how production cars utilise gestures and are intended to further reduce distraction, using input to simply invoke actions rather than precisely control quantities (as with sliders and dials).

We conducted a simulated driving study where users performed mid-air gestures whilst performing a Lane Change Task (LCT). This is a standard task (ISO standard 26022:2010) used in automotive studies to investigate the demands of secondary tasks whilst driving (e.g., responding to navigation instructions or interacting with a system). We evaluated the use of bimodal feedback, with combinations of four feedback types: Visual (standard car console screen), Audio (headphones), Peripheral Visual (LED display behind steering wheel) and Tactile (solenoids on the steering wheel) (Figure 1). Our contribution is a detailed investigation of the effects of gesture interaction and feedback whilst driving. We studied the impact of novel bimodal feedback combinations on visual distraction (gaze away from the road), car control (lane deviation), gesturing performance, and cognitive workload. Our results show the potential of multisensory non-visual feedback to support interaction without compromising driving safety.

Our primary contribution is an investigation of bimodal feedback for in-car gestures during the Lane Change Task. We look at four bimodal combinations of gesture feedback: Auditory-Visual, our baseline which is representative of typical in-car systems; Auditory-Peripheral; Auditory-Tactile ('push' feedback from the steering wheel); and Tactile-Peripheral. The latter three feedback types use non-visual and/or peripheral visual to enable drivers to focus on the road. In a simulated driving experiment, we investigate the efficacy of these feedback combinations.

RELATED WORK

Users experience uncertainty when interacting with mid-air gesture systems [9], which often lack a sense of control [4]. This is partly because there are less tactile cues to support interaction: for example, users lose the important feedback from physically touching a keyboard button or a touchscreen [7, 27]. Good feedback can help users overcome these usability problems, by reassuring users that they are interacting correctly [7]. This is especially important for gestures in cars; usability problems not only impede interaction, but may have negative effects on driving as well [13, 28, 30]. This is a timely problem, as some production vehicles now use gesture interaction with their in-car systems.

The gesture systems currently found in cars only give a limited amount of feedback. The feedback is also predominantly visual, shown on the central console screen (e.g., BMW's 7 Series, VW Golf, etc). This is not ideal for an in-car interface, as visual feedback requires users to divide attention between the road and the display. Others have investigated alternative feedback mechanisms for in-car gestures: for example, auditory feedback from loudspeakers, haptic feedback from the steering wheel, and simple visuals presented in the periphery of the windscreen [36]. These feedback modalities reduced eyes-off-the-road time significantly and did not compromise driving performance. However, the feedback significantly increased mental demand compared to visual feedback on the screen.

Feedback modalities that discourage drivers from looking at the console screen are worth investigating. Whilst potentially able to convey less detailed information than a high resolution screen, multimodal combinations may increase their efficacy, supporting successful input without affecting driving. Multimodal feedback is beneficial for gesture systems [9] and in-car information systems [13, 28, 32, 36]. It can reduce mental demand [25], reduce visual distraction [14, 36, 38], and assure users they are interacting correctly [4, 7, 9]. In this work, we investigate multimodal feedback comprised of three types of sensory information: tactile feedback direct from the steering wheel, auditory feedback, and ambient visual feedback presented in the periphery of the driver's visual attention.

Tactile Feedback in Cars

Visual attention is taken up by the primary driving task and the auditory modality is also often overloaded, with navigation systems, entertainment systems and passengers demanding attention. This leaves the tactile sense as an under-utilised modality, making it ideal for feedback about interaction with in-car systems.

Research has investigated various methods for presenting tactile feedback to drivers. Vibration is the most common form used in interactive devices and has also been integrated into steering wheels. Kern *et al.* [18] found that vibration from the steering wheel could accurately convey information through vibrotactile patterns. However, vibration may not be suitable for driving; even in laboratory settings, users find it difficult to distinguish the location of vibration on the steering wheel [2]. Vibration may also be masked by natural in-car vibrations from the road [18].

An alternative to vibration is to change the shape of the steering wheel, so that it 'pushes' the driver's hands for feedback [3, 36]. This is more distinctive than vibration from the surface. Shape-changing steering wheels have been used to inform drivers about events in autonomous cars [3] and to give feedback about gesture interaction [36]. The latter system consisted of soft pins that extended from the steering wheel surface to convey information. Different combinations of pins could be used to convey different information. This was used to give feedback about gestures, to the non-interacting hand that was holding the steering wheel. The evaluation of this 'cutaneous push' feedback found that it supported interaction, but increased mental demand.

Ultrasound haptic feedback has been used for feedback directly on the hand performing gestures [12, 14, 33, 37]. One study evaluated the use of ultrasound haptic feedback for movement-based gestures [37]. They found that ultrasound haptic feedback was good, but more effective when used with other types of feedback. On its own, ultrasound haptics still demanded visual attention away from the road. Harrington *et al.* [14] compared ultrasound haptics to visual feedback on a screen for two types of in-car controls: sliders and buttons. They found that users still looked at the screen when interacting with buttons, but haptic feedback was ideal for the slider controls, allowing drivers to keep attention on the road. In this work, we investigate the use of steering wheel 'cutaneous push' feedback, described above. However, ultrasound haptic feedback also shows potential and is worth investigating in future work.

Peripheral Visual Feedback

Screens are the predominant way of presenting information to drivers in cars, e.g., the central console screen and navigation devices placed near the dashboard. These are commonly used to present interaction feedback, although this is not ideal for driving scenarios: a driver's visual attention should be on the road. An alternative to screens is to present visual information using low fidelity displays within the driver's visual periphery, so that it can be seen whilst focusing on the road.

Peripheral displays are effective in demanding environments because they do not interfere with the primary task [26] of driving. Despite their simplicity, peripheral visual cues are also effective at conveying information [17, 39]. Such displays have been investigated for visual communication in cars, e.g., the *AmbiCar* system [21]. Peripheral visual cues have been used to inform drivers about safety distance violations [24], lane change decisions [23], current travel speed [42], and intentions of an automated car [21]. They have also been used for mid-air gesture feedback [36, 10]. In this work, we combine peripheral visual feedback with other feedback modalities, to see if this increases its effectiveness.

Summary

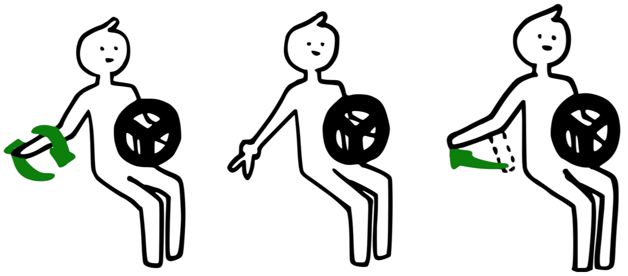
Mid-air gestures are now being used by drivers on the road, as well as in the active research area of automotive HCI. Gesture usability is especially important in cars, as usability problems could cause frustration and distraction. Good feedback can help users overcome these usability concerns, but feedback is particularly challenging in cars, as drivers already have high demands placed on their vision and audition. This has inspired research into novel feedback types that aim to discourage drivers from taking their eyes off the road. These have had some success in user studies but have limitations that may be overcome through multimodal presentation. Redundantly presenting feedback about interaction across multiple modalities can reduce the demands placed on drivers, allowing them to make use of the most appropriate information.

METHODOLOGY

We designed a simulated driving study to evaluate the effectiveness of different bimodal feedback techniques for three different mid-air gestures. The aim was to gain insight into the modality combination which best supported interaction without compromising driving safety.

Gestures

The set of gestures used for this study was based on mid-air gesture design guidelines [44, 11] and some already available for in-car interaction (BMW, VW). We used three gestures: swipe, circle, and victory (Figure 2). VW introduced the swipe left/right motion in their gesture enabled user interface [3]. BMW use a circular motion to increase/decrease a setting. They also introduced the victory gesture to turn the centre console screen on/off; this gesture is executed by extending the index and middle finger parallel to the gesture sensor device.



(a) Circle (b) Victory (c) Swipe

Figure 2. The three types of gestures used in our study. Participants were asked to gesture with their right hand above the gear stick.

Feedback

We combined the following four modalities into bimodal feedback techniques: Visual (standard car centre console), Auditory (headphones to cancel out the solenoid noises), Tactile (solenoids embedded into the steering wheel; Figure 4), and Peripheral Visual (LED strip behind the steering wheel; Figure

---

[3] https://www.volkswagen.co.uk/technology/comfort-and-convenience/gesture-control Accessed 2019-04-11

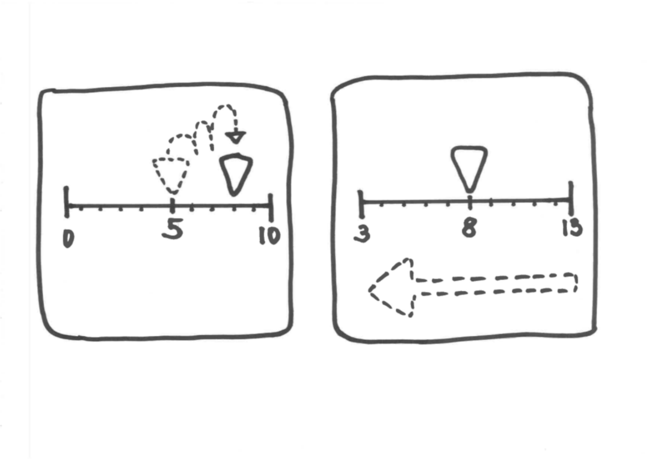
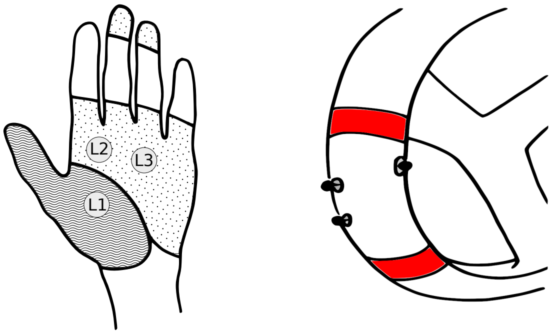


Figure 3. The 8 inch centre console screen for visual feedback. Left: change in cursor position after three Circle Clockwise motions. Right: change in scale labels after three Swipe Left motions.



(a) (b)

Figure 4. (a) Inner left hand with its most sensitive regions to pressure input via a $1cm^2$ round metal pin. Wavy: thenar/thumb region; dotted: median palmar region is 10-20% less sensitive than thenar/thumb region; White: fingers are more than 20% less sensitive to pressure input than thenar/thumb region [6]. The dots indicate where the solenoids were placed. (b) Steering wheel with three extruding cutaneous push pins. Red markers indicate where the participants' hand was placed.

1). Feedback was presented for each of the above gestures. In addition, whenever the hand entered the gesture interaction area feedback was given to assure the user that the system was attentive and ready for input [11, 8, 10]. There was no feedback provided on the hand exiting the interaction-box. The newly set system state (e.g. increased value) provided functional feedback for the successful interaction.

Visual Feedback: this was presented on the centre console to the right of the participant (Figure 3). The GUI design was based the Landrover Discovery Sport's centre console in terms of size of screen display, size of menu items, size of letters, etc. and Shakeri *et al.*'s [36] centre console display for mid-air gestures. The GUI was a single horizontal scroll bar (from 0 to 10) with the cursor set at 5. A circle gesture caused the cursor to move up or down on the scale depending whether it was a clockwise or anticlockwise motion. A swipe left caused the entire scale to shift from 0 - 10 to 1 to 11, with the cursor remaining at the centre position (as if the user "drags" the scale to the left). A swipe right moved the scale in the other direction. The scale shifted for the swipe motions because we wanted it to be different from the circle gesture feedback (cursor moving along the scale). The victory gesture turned the screen on / off. Feedback for entrance of the hand in the interaction box was a brightening of the screen from "standby" to "active" mode. The screen turned darker again after the hand exited the interaction box.

:-

Auditory Feedback: was presented through both speech (Spearcon i.e. fast speech) and non-speech feedback and lasted 500 ms (Table 1). Auditory feedback was provided via headphones. Feedback for the clockwise motion was the increase of a tone by an octave, and decrease by an octave for the anticlockwise motion. Speech feedback followed non-speech feedback representing an internal count from 0 to 10. Whenever the driver's hand entered the interaction box, they heard the c5 tone for 150 ms. The non-speech tones were generated in Audacity[4] and guided by Shakeri *et al.*'s [36] design (Table 1). The speech feedback was spoken by a male US American voice (www.cereproc.com/ Voice: Nathan. Accessed 2016-01-31).

Gesture	Non-speech	Duration	Non-Speech
V on	g#4 → c5	300 ms	on
V off	c5 → g#4	300 ms	off
SL	c4 → c4	250 ms	$\uparrow \{0-10\}$
SR	c5 → c5	250 ms	$\downarrow \{0-10\}$
CW	c4 → b4	250 ms	$\downarrow \{0-10\}$
CAW	b4 → c4	250 ms	$\uparrow \{0-10\}$

Table 1. This table shows the auditory feedback used for each gesture. The arrow in Non-Speech describes the transition from one note to the next. Duration describes the length of each non-speech unit. Speech stands for spoken gesture feedback; e.g. feedback for SL would be an increment of a number between $\{0-10\}$, and a decrement for SR. In total the feedback lasted 500ms.

Tactile Feedback: was presented via three pins under the driver's left palm (Figure 4). Pin P1 presented feedback to the thenar/thumb region, P2 and P3 provide feedback to the median palmar region (P2 behind the index finger; P3 behind the little finger). Feedback for the clockwise motion was the sequential presentation of pins P3, P2, and then P1 (Table 2). Presenting the pins in this fashion resulted in a "circular" motion mimicking the gesturing hand; anti-clockwise was represented by the sequential display of P1, P2, and P3. Each presentation lasted 500 ms in total. When the hand entered the interaction box, P3 was presented for 150 ms.

Peripheral Light Feedback: was presented on an LED strip from the A-pillar on the left side of the driver to the beginning of centre console (Figure 1). The strip was placed behind the steering wheel where the car instrument cluster would be (as proposed by Löcken *et al.* [22]). This placement enables feedback in the periphery of visual attention, allowing drivers to keep their eyes on the road ahead.

---

[4] Audacity Version 2.1.2 http://www.audacityteam.org/

Gesture	Pins	Duration	Time b/w Pins
V	all	150 ms	-
SL	P1 → P2	150 ms	50 ms
SR	P2 → P1	150 ms	50 ms
CC	P3 → P2 → P1	166 ms	-
CCC	P1 → P2 → P3	166 ms	-

Table 2. This table shows the tactile feedback used for each gesture. The arrow in Pins describes the transition from one feedback location on the palm to the next. Duration describes the length of each pin presentation.

Feedback for the swiping motions left and right was a yellow light animation mimicking the direction of the gesturing hand. Duration of the animation was 500 ms. Successful (anti-)clockwise motion was indicated by blue lights either incrementing to the right or decrementing to the left. As long as the hand was inside the interaction box, the blue lights remained on. Victory gesture feedback was presented with an animation of blue lights moving from the ends of the LED strip to the centre — to turn the system on. The duration was 500 ms. To turn the system off, feedback for the Victory gesture was an outward animation of red lights (from centre to the ends of the LED strip) (i.e. turn the system off). We chose red and blue colours to avoid issues for users who are colour blind. On entrance of the hand in the interaction box, the strip would pulse briefly (350 ms) in a dim white light.

Apparatus

The experiment was conducted in a usability laboratory equipped with 1) a computer, on which the OpenDS Version 3 simulation[5] was run, 2) a 81 inch projected driving simulator, 3) an 8 inch screen to the right of the driver mimicking a car's centre console screen, 4) a Leap Motion tracker to sense the user's gesturing hand, 5) a Logitech webcam located on the mount of the steering wheel in front of the driver, 6) three solenoid powered pins protruding from the steering wheel providing feedback to the driver's left palm [35], 7) a capacitive sensor attached to the steering wheel under the driver's right hand (to measure when driver's hand leaves steering wheel), and 8) a 107 cm long LED light strip. The placements of the individual devices were guided by the measurements of a Land Rover Discovery Sport. We placed the Leap Motion device where the gear stick would be such that the interaction area is a cube on the right of the steering wheel, above the gear stick. This ensured that the gesture execution area was close to the steering wheel and gear shift, as recommended by Riener *et al.* [31]. The measurements of the interaction box are: width: 14 cm, height: 20 cm, and depth: 14 cm.

In order to be able to analyse visual distraction, the webcam recorded the participants' eye gaze while performing the driving and input tasks. Gaze and head pose data were extracted using OpenFace[6], an open source tool for eye-gaze and head pose estimation. An SVM classifier with a linear kernel was trained on 11,845 images obtained during a pilot study. Input data for the classifier were 3D vectors for each eye and head pose rotation. The SVM model classified 94.56% eyes-off-the-road time correctly (10-fold cross validation).

Procedure

During pilot studies and as found by previous research [25, 36], participants were prone to make accidental gestures by entering the interaction box and causing unwanted system response. To account for falsely classified gestures which can increase mental demand [11], we implemented our system such that it only provided feedback to the expected gesture. If a circular motion was expected, only circular gestures caused system response. We chose this solution because we were not evaluating the quality of the gesture recogniser (we used the default recogniser for LeapMotion) but the feedback.

Participants were asked to perform gestures in following ways: Swipe Left 2, 3, 4 times (SL2, SL3, SL4), Swipe Right 2, 3, 4 times (SR2, SR3, SR4), Victory (V), Circle Clockwise 2, 3, 4 times (CW2, CW3, CW4), and Circle Anti Clockwise 2, 3, 4 times (CAW2, CAW3, CAW4). We differentiate the gestures depending on motion and direction, not number of execution; this results in five gesture types overall: SL, SR, CW, CAW, and V. As suggested by previous research [5, 31, 1] we kept the gesture set smaller than eight. However, the number of instructions to gesture (e.g. SR 3; SL 4 etc.) increased the count of performed gestures to 16.

The literature recommends [11, 8] a system should provide information whenever the hand enters the interaction box. In our study, feedback for this lasted 150 ms. The duration of a single gesture lasted for 750 ms gesture execution and 500 ms gesture feedback. If a participant is instructed to swipe left 4 times (SL4), the entire interaction can last up to 5150 ms (150 ms on interaction box entrance plus 4*x*750ms gesture execution and 4*x*500ms feedback). For the swipe and circular motions we used the built in Leap Motion classifiers. The victory gesture was recognised by extending the index and middle finger for at least 750 ms.

The feedback was presented *functionally*. It was presented *after* gesture execution in a *discrete* manner instead of presenting it during the execution and continuously. Continuous feedback is important for usability [11, 9]. However, it might overload the driver and increase distraction.

Gestures were performed with the right hand (as if driving in Right-Hand-Traffic).

Lane Change Task

The Lane Change Task (LCT) (ISO standard 26022:2010) aims at measuring the decrease in driving performance while conducting a secondary task. Therefore, the result of the LCT serves as an estimate for the demand of the secondary task [19]. From the instruction to change lane until the car reaches the target lane, we measure its position for LCT analysis (Figure 5). The average time to complete a single lane change task on a motorway is 5.8 seconds [40], which correlates with our findings. Similar to Lee *et al.* [20], we define lane-change initiation by the time when the wheel crosses the lane line. On average, the participants used the first $\tilde{4}$.1 seconds to comprehend and initiate the lane change; during this time, lane

---

[5] OpenDS Version 3, https://www.opends.eu/ Accessed 2017-04-25

[6] OpenFace, https://github.com/TadasBaltrusaitis/OpenFace Accessed 2017-04-17

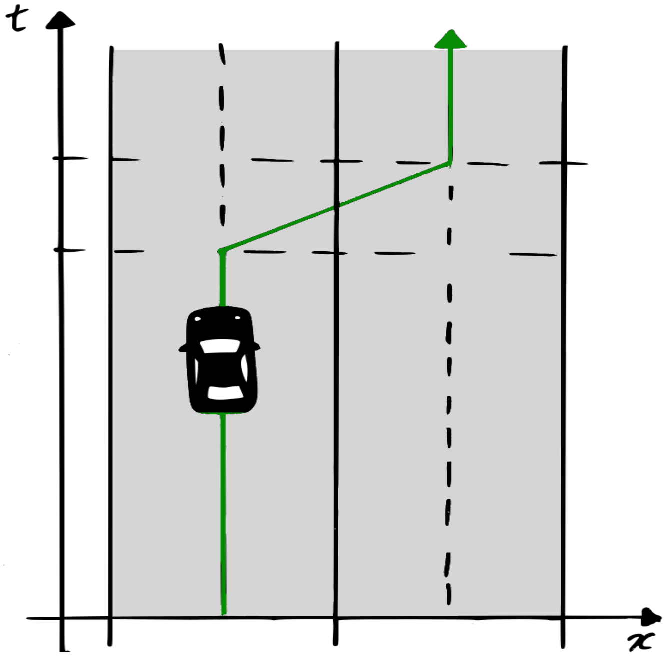


Figure 5. Lane Change Task from the left lane to the right lane. The (green) arrow denotes the optimal driving path (i.e. zero lane deviation). Instruction to change is given at 0 seconds; after 4.1 seconds on average, the lane change is initiated and completed by 5.8 seconds.

deviation equals zero in the centre of the *current lane* (left lane on Figure 5). During the next $\tilde{1}$.7 seconds, the car transitioned towards the target lane (right lane); zero lane deviation is on the *transitioning vector*. These two phases result in the optimal lane change path, where lane deviation is zero (green path on Figure 5. During the study, participants were prompted to change lanes every 24 seconds. Either simultaneously or up to four seconds after the instruction to gesture was given. We asked them to execute both commands as quickly and safely as possible.

STUDY

The aim of this study was to measure the impact of gesture feedback modalities on demanding driving. Therefore, participants performed the Lane Change Task along with the mid-air gesturing task.

Conditions

We presented following combinations of feedback modalities: Auditory-Visual (AV), Auditory-Tactile (AT), Auditory-Peripheral (AP), and Tactile-Peripheral (TP). Auditory-Visual feedback functions as baseline for our study since it has already been used in the literature for mid-air gesture interaction [25, 38]. We did not test for Visual-Peripheral because both techniques use the same sensory channel for information throughput; we did not want to overload a single channel thus decided against this combination. We also omitted Tactile-Visual as feedback technique since it resulted in worst performance and was ranked as least preferred by participants in a pilot study.

Hypotheses

H1: There will be no significant difference in visual distraction from the primary driving task between the conditions;
H2: There will be no significant difference in lane deviation between the feedback conditions;
H3: No significant difference in perceived mental workload will occur across the conditions;
H4: Participants will prioritise Lane Change Task over gesture execution when prompted simultaneously.

Experimental Variables

The Independent Variable was mid-air gesture feedback. There were four levels: Auditory-Visual, Auditory-Peripheral, Auditory-Tactile, and Tactile-Peripheral. The Dependent Variables were: lane deviation (metres), visual attention to primary task (number of glances at centre console), number of correct of gestures (% correct), task duration (ms), prioritisation of LCT over gesture task (start time of execution), perceived workload (NASA TLX), and our own questionnaire (demographics, handedness, preferences of feedback).

Participants

Nineteen participants (nine females) ranging from 19 to 53 years of age (μ 26.68 σ 9.23) were recruited via our University's student online forum. Of these 19, XXX participants had a Left-Hand-Traffic driving license and XXX a Right-Hand-Traffic license. A total of 14 participants indicated that they had no prior experience with mid-air gesture interfaces. All participants were right handed.

Procedure

Upon arrival, participants were briefed about the study and given an introductory training session. This session was structured like the experiment, but it was shorter. In the training, participants executed each gesture once during each feedback condition. Participants started on the out most left lane on a five lane motorway and had to steer towards the middle lane. After 30 seconds a trigger was fired to instruct the driver to change a single lane (left/right) and execute a gesture simultaneously. The instructions to change lane were indicated with green arrows on bridge panels over the motorway. The arrows pointed down onto the target lane. The lane-change and gesture instructions were presented after random intervals; instructions to gesture were prompted $0-4$ seconds after a lane-change instruction; participants had 30 second to complete both instructions before the next instructions were prompted. The participants were asked to complete both tasks as quickly and safely as possible. The direction of changes was balanced.

During the main experiment, each bimodal feedback block was presented twice, resulting in 8 blocks in total. Each block lasted 8 minutes (16 gestures x 30 seconds per lane-change/gesture instruction). Each gesture was executed once resulting in 16 gestures. After each block, participants were asked to fill in a NASA TLX workload questionnaire. At the end of the experiment, participants were asked to fill in a demographics questionnaire we designed. The experiment lasted 90 minutes with briefing and questionnaire, participants were reimbursed with £10.

RESULTS

Lane Deviation

We used the Root Square Mean Error to measure how close a the user's driving path is to the ideal driving path [34]. We realigned the resulting non-normal distribution using the Aligned-Rank-Transform. A repeated-measures Anova showed a significant difference of condition on lane deviation $\chi^2(3) = 4.545, p = 0.003$, however no impact of gesture on lane deviation $\chi^2(12) = 1.131, p = 0.330$ (Figure 6(a)). A pairwise comparison test revealed that there is a significant difference between the feedback conditions AudioVisual and AudioTactile, $p = 0.001$. There were no significant differences between any other pairwise comparison.

A Spearman's correlation was run to asses the relationship between driving license obtained in a Left-Hand-Traffic country and driving performance. There was no statistically significant correlation, $r_s = 0.009, p = 0.952$. A Spearman's corrrelation was run between gesturing performance and driving performance with $r_s = 0.007, p = 0.965$. Neither gesturing nor driving performance influenced each other. A Spearman's correlation was run between gender and driving performance, no statistically significant correlation was found, $r_s = 0.173, p = 0.478$.

If instruction to gesture and to change lane were given at the same time, a gesture was labelled as "prioritised" if it was executed before the lane change task was initiated. A total of 20.57% gestures were prioritised over lane change. Binomial logistic regression was performed to ascertain the effects of gesture prioritisation over lane change task. None of the gestures (CW3 $p = 0.703$, CW4 $p = 0.76$, CAW2 $p = 0.949$, CAW3 $p = 0.239$, CAW4 $p = 0.49$, SL2 $p = 0.346$, SL3 $p = 0.528$, SL4 $p = 0.282$, SR2 $p = 0.245$, SR3 $p = 0.792$, SR4 $p = 0.885$, V $p = 0.849$) had a significant impact on decision to prioritise it over lane change maneuver.

Gaze

A Shapiro-Wilk normality test showed that our collected gaze data is non-normal, $W = 0.401, p \leqslant 0.001$. An Aligned Rank Transform ANOVA was conducted to determine the effects of condition ($\chi^2(3) = 30.301, p \leqslant 0.001$) and gesture ($\chi^2(12) = 6.278, p \leqslant 0.001$) on gaze behaviour (Figure 6); both have a significant impact on eyes-off-the-road time. Pairwise comparison test revealed that Audio-Visual has significantly higher impact on eyes-off-the-road time than the other conditions ($p \leqslant 0.001$). Further, Audio-Peripheral and Tactile-Peripheral feedback have significantly differing impact on gaze behaviour ($p \leqslant 0.001$). Pairwise comparison test revealed that gestures the V, CAW2, CAW3, CW2 (EORT <= 0.075 seconds) had a significantly lower impact on looking away time than SR2, SR3, SR4, and SL4 (EORT >= 0.125 seconds).

Friedman's test revealed a significant difference between the types of gestures on gaze behaviour, $\chi^2(4) = 15.898, p = 0.003$. *Post hoc* analysis with Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.001$. Median levels for the gesture types CC, CAW, SL, SR, and V were 0.200, 0.166, 0.262, 0.264, and 0.069 respectively. A multiple comparison test revealed that following significant differences between the gesture types: CC - SR $p = 0.005$; CAW - SR $p = 0.035$; SL - V $p = 0.041$; and SR - V $p \leqslant 0.001$. There were no significant differences between the other types of gestures.

A Spearman's correlation was run to assess the relationship between eye gaze behaviour and gesture performance. There is no statistically significant correlation ($r_s = -0.304, p = 0.062$. An ANOVA on the regression model found no significant main effect of gaze behaviour on driving performance during gesture execution: $chi^2(1) = 0.966, p = 0.324$.

Secondary Task Performance

Overall gesture performance across all conditions and participants was 73.45%. An ANOVA on the regression model found no significant main effect of condition on gesture performance: $chi^2(3) = 3.31, p = 0.346$. Average gesture performance during each condition is AV with 74.624%, AT with 74.237%, AP with 74.496%, and TP with 70.743%.

Gesture	Success
CIRCLE CLOCKWISE 2	73.826
CIRCLE CLOCKWISE 3	67.333
CIRCLE CLOCKWISE 4	66.216
CIRCLE COUNTER CLOCKWISE 2	87.417
CIRCLE COUNTER CLOCKWISE 3	89.404
CIRCLE COUNTER CLOCKWISE 4	84.564
SWIPE LEFT 2	68.000
SWIPE LEFT 3	65.101
SWIPE LEFT 4	52.667
SWIPE RIGHT 2	58.278
SWIPE RIGHT 3	44.295
SWIPE RIGHT 4	35.811
VICTORY	95.477

Table 3. Average gesture performance across conditions and participants.

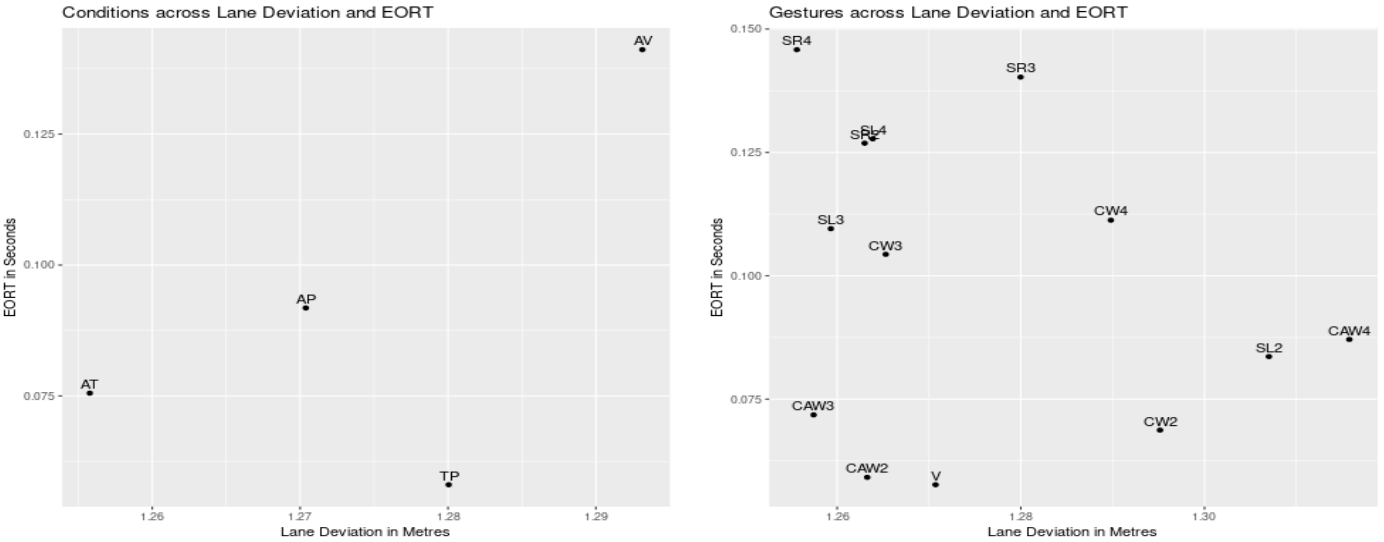
Workload Measure

Manova analysis shows that there is no impact of multimodal feedback (condition) on any of the perceived workload measures $F(3, 148) = 0.875, p = 0.94; Wilk's\lambda = 0.36$.

Users preferred Audio-Visual feedback over the other feedback conditions (Figure 7(b)).

DISCUSSION

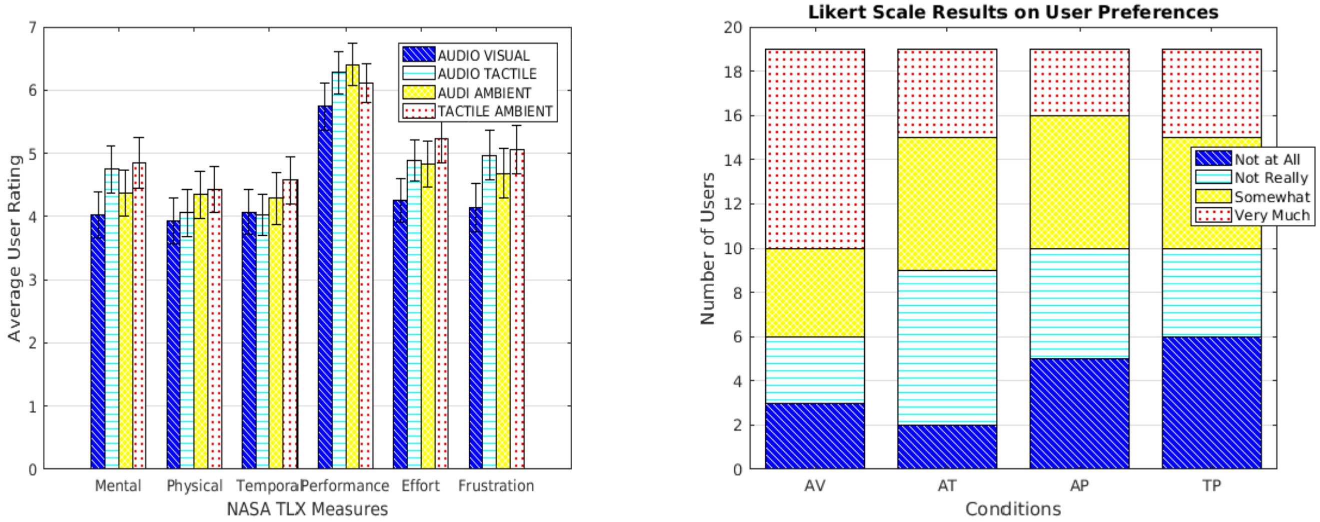
We investigated the efficacy of bimodal feedback for gesture input during a simulated driving task, comparing three novel modality combinations to typical Audio-Visual feedback. For in-car gesture feedback to be successful, it needs to support interaction whilst importantly not having a negative impact on the driver's awareness of the road and control of the vehicle. Audio-Visual feedback had a significantly higher effect on gaze off-the-road and lane deviation. These findings support the use of the other presented novel feedback modalities used in this work. The other feedback conditions — Audio-Peripheral, Audio-Tactile, and Tactile-Peripheral — did not provide as much information as the Audio-Visual. However, the lower information feedback modalities have less impact on the primary driving task, whilst still supporting the driver sufficiently for input without glancing at the input sensor or the



(a) Eyes-Off-the-Road Time (EORT) and lane deviation across conditions AudioVisual (AV), AudioTactile (AT), AudioPeripheral (AP), and TactilePeripheral (TP).

(b) Eyes-Off-the-Road Time (EORT) and lane deviation across gestures. The gestures: Circle Anti Clockwise (CAW), Circle Clockwise (CW), Swipe Left (SL), Swipe Right (SR), and Victory (V). The numerical value on the abbreviations indicate the number of times the instructed gesture execution.

Figure 6. Lane deviation and gaze behaviour analysis across presented conditions and gestures.



(a) Results of the NASA TLX questionnaire.

(b) User preferences of bimodal feedback. (AV: Audio-Visual, AT: Audio-Tactile, AP: Audio-Peripheral, TP: Tactile-Peripheral).

Figure 7. Questionnaires and user feedback.

car console screen. The four feedback combinations did not impact gesturing performance, nor influence the participant's choice to prioritise gesture execution over lane change. Participants generally prioritised lane change over gesture execution, when both instructions were given at the same time.

Two of our feedback combinations used Peripheral Visual output: low-fidelity light animations presented in the visual periphery, behind the steering wheel. This type of display can only give a limited amount of feedback and is limited to one axis for spatial information. Yet this was enough to support successful gesture input. Such peripheral displays are worth investigating more, to reduce the reliance on a centre console screen that necessitates taking eyes off the road. Our peripheral display prototype was half the length of the dashboard in front of the driver, but this could be reduced further, reducing the amount of information in the visual periphery.

A previous study [36] with unimodal gesture feedback found that unfamiliar output modalities (e.g., peripheral visuals and haptic feedback on the steering wheel) increased cognitive demand whilst driving, in comparison to visual feedback on the centre console. This has negative safety implications. Our motivation for investigating bimodal feedback was to see if redundantly presenting information across two modalities could reduce the mental workload associated with interaction. Our results show that mental workload was similar for the four feedback combinations, supporting the use of combinations of Audio, Peripheral Visual, and Tactile feedback. These novel combinations had similar cognitive demand to the baseline Audio-Visual pair, with the advantage of eliminating the need to glance at the console screen.

Secondary Task Performance: Our analysis revealed a significant difference in secondary task performance across gestures. The V gesture yielded highest performance with 95.47% accuracy. The Victory gesture consists of a single static and discrete pose, unlike Circle or Swipe. The worst performing gesture type was Swipe with an overall performance of 44.66% (SR4 with 35.81%). During the Swipe Right motion, the arm is moving away from the driver and this "away" motion causes greater arm and shoulder fatigue [15]. Further, the swiping motion has to be "reset" — the hand has to be returned to the starting point to swipe again. This "resetting" motion might have caused misclassification of intent. This is supported by user feedback reporting frustration with the swiping. We therefore argue that swipe gestures are not suitable for the limited interaction area of a car cockpit where the driver cannot "reset" the hand outside of the interaction area without causing unwanted reverse actions.

User Preferences Nine out of nineteen users preferred the Audio-Visual condition over the other feedback combinations. This high rate may be due to the users being more familiar with audio and visual than tactile feedback. This familiarity results in less time and effort needed to learn the system messages. However, it would be beneficial for drivers to learn these messages since there is a clear potential for gestures with tactile feedback in the automotive context.

CONCLUSION

We investigated novel bimodal feedback combinations for mid-air gestures in cars, which are now being used by an increasing number of drivers on the road. Usability concerns mean good feedback is important, but a balance needs to be found between supporting interaction and reducing distraction in an already demanding environment. Visual feedback is most commonly used, but takes visual attention away from driving. We investigated three alternatives that lack the expressive capabilities of high resolution screens, but are intended to allow drivers to focus on what is happening around them. Our experiment found that these feedback modalities offered just as much support for interaction without negatively affecting driving performance, visual attention and cognitive demand. These results provide compelling support for using non-visual feedback from in-car systems, supporting input whilst letting drivers focus on driving.

REFERENCES

1. Micah Alpern and Katie Minardo. 2003. Developing a car gesture interface for use as a secondary task. *CHI EA* (2003), 932. `DOI:` **http://dx.doi.org/10.1145/766077.766078**
2. Frank Beruscha, Lei Wang, Klaus Augsburg, and Hartmut Wandke. 2010. Do drivers steer towards or away from lateral directional vibrations at the steering wheel?. In *Proc. 2nd European Conference on Human Centred Design for Intelligent Transport Systems*. 227–236.
3. Shadan Sadeghian Borojeni, Torben Wallbaum, Wilko Heuten, and Susanne Boll. 2017. Comparing Shape-Changing and Vibro-Tactile Steering Wheels for Take-Over Requests in Highly Automated Driving. In *AutoUI '17*. ACM, Oldenburg, Germany, 221–225. `DOI:` **http://dx.doi.org/10.1145/3122986.3123003**
4. Patricia Ivette Cornelio Martinez, Silvana De Pirro, Chi Thanh Vi, and Sriram Subramanian. 2017. Agency in Mid-air Interfaces. In *CHI*. `DOI:` **http://dx.doi.org/10.1145/3025453.3025457**
5. Nelson Cowan. 2001. The magical number 4 in short term memory. A reconsideration of storage capacity. In *Behavioral and Brain Sciences*, Vol. 24. 87–186. `DOI:` **http://dx.doi.org/10.1017/S0140525X01003922**
6. Charlotte Fransson-Hall and Asa Kilbom. 1993. Sensitivity of the hand to surface pressure. *Applied Ergonomics* 24, 3 (1993), 181–189. `DOI:` **http://dx.doi.org/10.1016/0003-6870(93)90006-U**
7. Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2014. Tactile Feedback for Above-Device Gesture Interfaces: Adding Touch to Touchless Interactions. In *ICMI*.
8. Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2015. Interactive light feedback: Illuminating Above-Device gesture interfaces. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 9299. 478–481. `DOI:` **http://dx.doi.org/10.1007/978-3-319-22723-8{_}42**
9. Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2016. Do That , There : An Interaction Technique for Addressing In-Air Gesture Systems. In *CHI*. ACM, 2319–2331. `DOI:` **http://dx.doi.org/10.1145/2858036.2858308**
10. Euan Freeman, Dong-Bach Vo, and Stephen Brewster. 2019. HaptiGlow: Multimodal Feedback to Help Users Position their Hands for Mid-Air Gestures and Ultrasound Haptic Feedback. In *Proceedings of the IEEE World Haptics Conference 2019, the 8th Joint Eurohaptics Conference and the IEEE Haptics Symposium*. IEEE, to appear.
11. Thomas M Gable, R May May, and Bruce N Walker. 2014. Applying Popular Usability Heuristics to Gesture Interaction in the Vehicle. In *AutoUI*. 1–7. `DOI:` **http://dx.doi.org/10.1145/2667239.2667298**
12. Orestis Georgiou. 2017. Haptic In-Vehicle Gesture Controls. In *AutoUI*. Oldenburg, Germany. `DOI:` **http://dx.doi.org/10.1145/3131726.3132045**

13. Paul Green. 2000. Crashes Induced by Driver Information Systems and What Can Be Done to Reduce Them. *Society of Automotive Engineers Conference Proceedings* (2000), 27–36.

14. Kyle Harrington, David R. Large, Gary Burnett, and Orestis Georgiou. 2018. Exploring the Use of Mid-Air Ultrasonic Feedback to Enhance Automotive User Interfaces. In *AutomotiveUI '18*. ACM, Toronto, ON, Canada, 11–20. DOI: `http://dx.doi.org/10.1145/3239060.3239089`

15. Juan David Hincapié-ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance : A Metric to Quantify Arm Fatigue of Mid - Air Interactions. (2014), 1063–1072.

16. Cristy Ho, Hong Z. Tan, and Charles Spence. 2005. Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behaviour* 8, 6 (2005), 397–412. DOI: `http://dx.doi.org/10.1016/j.trf.2005.05.002`

17. John Jonides. 1981. Voluntary Versus Automatic Control Over the Mind'S Eye'S Movement. *Attention and performance IX* 9 (1981), 187–203.

18. Dagmar Kern, Paul Marshall, Eva Hornecker, Yvonne Rogers, and Albrecht Schmidt. 2009. Enhancing navigation information with tactile output embedded into the steering wheel. 5538 LNCS, 1 (2009), 42–58. DOI: `http://dx.doi.org/10.1007/978-3-642-01516-8`

19. Thomas Kopinski, Jan Eberwein, Stefan Geisler, and Uwe Handmann. 2016. Touch versus mid-Air gesture interfaces in road scenarios - Measuring driver performance degradation. In *IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC*. DOI:`http://dx.doi.org/10.1109/ITSC.2016.7795624`

20. Suzanne E Lee, Erik C B Olsen, Walter W Wierwille, and Michael Goodman. 2004. *A Comprehensive Examination of Naturalistic Lane-Changes 6. Performing Organization Code Unclassified*. Technical Report.

21. Andreas Löcken, Wilko Heuten, and Susanne Boll. 2016. AutoAmbiCar: Using Ambient Light to Inform Drivers About Intentions of Their Automated Cars Motivation. In *AutoUI*. DOI:`http://dx.doi.org/10.1145/3004323.3004329`

22. Andreas Löcken, Heiko Müller, Wilko Heuten, and Susanne Boll. 2013. AmbiCar: Towards an in-vehicle ambient light display. In *AutoUI*.

23. Andreas Locken, Heiko Muller, Wilko Heuten, and Susanne Boll. 2015. An experiment on ambient light patterns to support lane change decisions. In *IEEE Intelligent Vehicles Symposium*. DOI: `http://dx.doi.org/10.1109/IVS.2015.7225735`

24. Andreas Loecken, Wilko Heuten, and Susanne Boll. 2015. Supporting Lane Change Decisions with Ambient Light. *AutoUI '15* (2015), 204–211. DOI: `http://dx.doi.org/10.1145/2799250.2799259`

25. Keenan R May, Thomas M Gable, and Bruce N Walker. 2014. A Multimodal Air Gesture Interface for In Vehicle Menu Navigation. In *AutoUI*. 1–6. DOI: `http://dx.doi.org/10.1145/2667239.2667280`

26. Mark I. Nikolic and Nadine B. Sarter. 2001. Peripheral visual feedback: a powerful means of supporting effective attention allocation in event-driven, data-rich environments. *Human factors* 43, 1 (2001), 30–38. DOI: `http://dx.doi.org/10.1518/001872001775992525`

27. Antti Oulasvirta, Sunjun Kim, and Byungjoo Lee. 2018. Neuromechanics of a Button Press. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. DOI: `http://dx.doi.org/10.1145/3173574.3174082`

28. Carl a. Pickering, Keith J. Burnham, and Michael J. Richardson. 2007. A research study of hand gesture recognition technologies and applications for human vehicle interaction. *3rd Conf. on Automotive …* (2007), 1–15. `http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.135.7688&rep=rep1&type=pdf`

29. Ioannis Politis. The Effects of Modality, Urgency and Message Content on Responses to Multimodal Driver Displays. In *AutomotiveUI 2014 Adjunct Proceedings*. Seattle, WA, USA, 1–5. DOI: `http://dx.doi.org/10.13140/2.1.4592.3842`

30. Bryan Reimer, Bruce Mehler, J. Dobres, and J.F. Coughlin. 2013. *The Effects of a Production Level "Voice - Command" Interface on Driver Behavior : Summary Findings on Reported Workload , Physiology , Visual Attention , and Driving Performance*. Technical Report 17A. Massachusetts Institute of Technology, Cambridge, MA.

31. Andreas Riener, Alois Ferscha, Florian Bachmair, Patrick Hagmüller, Alexander Lemme, Dominik Muttenthaler, David Pühringer, Harald Rogner, Adrian Tappe, and Florian Weger. 2013. Standardization of the in-car gesture interaction space. *AutoUI* (2013), 14–21. DOI: `http://dx.doi.org/10.1145/2516540.2516544`

32. Meghan Rogers, Yu Zhang, David Kaber, Yulan Liang, and Shruti Gangakhedkar. 2011. The effects of visual and cognitive distraction on driver situation awareness. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. DOI: `http://dx.doi.org/10.1007/978-3-642-21741-8{_}21`

33. Sonja Rümelin, Thomas Gabler, and Jesper Bellenbaum. 2017. Clicks are in the Air: How to Support the Interaction with Floating Objects through Ultrasonic Feedback. In *AutoUI*. DOI: `http://dx.doi.org/10.1145/3122986.3123010`

34. Dario D. Salvucci. 2001. Predicting the effects of in-car interfaces on driver behavior using a cognitive architecture. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01* (2001), 120–127. DOI:`http://dx.doi.org/10.1145/365024.365064`

35. Gözel Shakeri, Alexander Ng, and Stephen A Brewster. 2016. Evaluating Haptic Feedback on a Steering Wheel in a Simulated Driving Scenario. In *CHI EA '16*. `DOI:` `http://dx.doi.org/10.1145/2851581.2892497`
36. Gözel Shakeri, John H Williamson, and Stephen Brewster. 2017. Novel Multimodal Feedback Techniques for In-Car Mid-Air Gesture Interaction. In *AutoUI*. `DOI:` `http://dx.doi.org/10.1145/3122986.3123011`
37. Gözel Shakeri, John H Williamson, and Stephen Brewster. 2018. May the Force Be with You: Ultrasound Haptic Feedback for Mid-Air Gesture Interaction in Cars. (2018). `DOI:http://dx.doi.org/10.1145/3239060.3239081`
38. Jason Sterkenburg, Steven Landry, Myounghoon Jeon, and Joshua Johnson. 2016. Towards an in-vehicle sonically-enhanced gesture control interface: a pilot study. In *ICAD*. 0–3. `DOI:` `http://dx.doi.org/10.21785/icad2016.015`
39. Jan Theeuwes. 1991. Exogenous and endogenous control of attention: the effect of visual onsets and offsets. *Perception & Psychophysics* 49, 1 (1991), 83–90. `DOI:` `http://dx.doi.org/10.3758/BF03211619`
40. L. Tijerina, W. R. Garrott, M. Glecker, D. Stoltzfus, and E. Parmer. 1997. *Van and Passenger Car Driver Eye Glance Behavior During Lane Change Decision Phase, Interim Report*. Technical Report. NHTSA, U.S. Department of Transportation.
41. Jan B F Van Erp and Hendrik a H C Van Veen. 2004. Vibrotactile in-vehicle navigation system. *Transportation Research Part F: Traffic Psychology and Behaviour* 7, 4-5 (2004), 247–256. `DOI:` `http://dx.doi.org/10.1016/j.trf.2004.09.003`
42. Hanneke Hooft van Huysduynen, Jacques Terken, Alexander Meschtscherjakov, Berry Eggen, and Manfred Tscheligi. 2017. Ambient Light and its Influence on Driving Experience. In *AutoUI*. `DOI:` `http://dx.doi.org/10.1145/3122986.3122992`
43. Christopher D Wickens. 2008. Multiple resources and mental workload. *Human factors* 50, 3 (2008), 449–455. `DOI:http://dx.doi.org/10.1518/001872008X288394`
44. Yu Zhang and Linda Angell. 2014. Pointing Towards Future Automotive HMIs: The Potential for Gesture Interaction. *AutoUI* 22, 3 (2014), 22–29. `DOI:` `http://dx.doi.org/10.1016/j.apergo.2013.10.013`