Novel Multimodal Feedback Techniques for In-Car Mid-Air Gesture Interaction

Gözel Shakeri, John H. Williamson, Stephen Brewster
Glasgow Interactive Systems Group
University of Glasgow, Scotland
g.shakeri.1@research.gla.ac.uk, {johnh.williamson,stephen.brewster}@glasgow.ac.uk

ABSTRACT
This paper presents an investigation into the effects of different feedback modalities on mid-air gesture interaction for infotainment systems in cars. Car crashes and near-crash events are most commonly caused by driver distraction. Mid-air interaction is a way of reducing driver distraction by reducing visual demand from infotainment. Despite a range of available modalities, feedback in mid-air gesture systems is generally provided through visual displays. We conducted a simulated driving study to investigate how different types of multimodal feedback can support in-air gestures. The effects of different feedback modalities on eye gaze behaviour, and the driving and gesturing tasks are considered. We found that feedback modality influenced gesturing behaviour. However, drivers corrected falsely executed gestures more often in non-visual conditions. Our findings show that non-visual feedback can reduce visual distraction significantly.

CCS Concepts
•**Human-centered computing** → **Haptic devices; Auditory feedback; Gestural input;**

Author Keywords
multimodal feedback; mid-air gestures;

INTRODUCTION
In-car interfaces can negatively impact safety if they increase mental workload or distract the driver [10]. The third most common cause of car crashes and near crash events is driver distraction [4]. Rogers *et al.* [33] investigated the effects of driver distraction and found that visual distraction significantly affected all aspects of situation awareness. In order to reduce driver distraction it is important to minimise eyes-off-the-road time when interacting with the infotainment system. Mid-air gesture systems are becoming more common with popular cars such as the 2017 BMW 5 Series and the 2017 VW Golf. These systems have the potential to reduce eyes-off-the-road time compared to traditional touch based interactions [15, 25]. Car manufacturers like BMW, VW, Cadillac, and Hyundai are investing in mid-air gesture interfaces[1] for in-car interaction. Despite a lack of scientific literature and a full understanding of the effects of mid-air gestures on interaction, driving performance, visual attention to the driving task and perceived workload, mid-air gestures are used in commercially available cars. Thus, it is important to investigate the impact of in-car mid-air gesture interaction on driving performance and mental and visual demand.

The advantage of gesturing in mid-air is that it is monitored by a human's proprioception and can in many cases be performed without paying any visual attention [23]. Despite a range of possible feedback types (e.g. auditory, tactile), mid-air gesture output in commercially available cars is generally limited to visual displays[2]. Driving is a highly visual task and input on infotainment systems normally compete for the driver's visual attention. Even if no visual attention is required (such as with voice entry) users tend to look towards the loudspeaker or microphone, awaiting system response [9, 31]. Thus, it may be more suitable to support mid-air gestures with non-visual feedback.

Using mid-air gestures can also minimise biomechanical interference [37] which may decrease the driver's ability to operate car controls. Biomechanical interference in this context is considered as a state when the driver's is referred to parts of the body is shifting out of the natural driving position, e.g. when reaching for dials, knobs, touch screen displays, etc. A reduction in biomechanical interference can be achieved with mid-air gestures since they only require casual motion in air rather than accurate point movements at the centre of the dashboard. This makes mid-air gestures potentially less physically and mentally demanding than direct touch.

Beyond distraction, mid-air gesture interaction is still a relatively new technology and users may not be familiar with such applications. This unfamiliarity may affect driving performance and mental workload. Since driving is a highly visual task, the distribution of information to other available

[1] USA Today https://www.usatoday.com/story/money/cars/2013/01/10/cartech-gestures-ces/1820453/ Accessed 2017-05-17

[2] BMW 7 Series (G11), Accessed 2017-05-16 https://en.wikipedia.org/wiki/BMW_7_Series_(G11), VW Golf R, Accessed 2017-05-16 http://www.volkswagen.co.uk/new/golf-vii-pa/explore/r

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

channels, such as auditory or tactile, may help to alleviate interference and potential resultant performance decrements.

Wickens' Multiple Resource Theory [40] suggests that interference can occur between primary and secondary tasks when they call upon the same resources. There is a body of research [39, 13, 29] showing that distributing information to non-visual modalities does not increase reaction time to cues when mental workload is increasing. Auditory feedback [29], tactile stimuli [38, 39, 13], and peripheral visual feedback [27, 22] have shown to successfully cue driver attention without dividing visual attention. An additional advantage of providing feedback is that it is needed for gesturing to help the user understand: (1) system attention [6]; (2) the in/correctness of gesture execution; and (3) provide greater user satisfaction [26].

Using different sensory channels for mid-air gesture feedback has the potential to enhance in-car interaction, however research has only considered visual and audio-visual feedback [25, 35], leaving other modalities unstudied. Thus, this paper presents an investigation into the effects of different types of feedback for in-car mid-air gestures: (1) auditory feedback; (2) peripheral lights feedback; and (3) cutaneous push feedback. Performance of the feedback is invested through eyes-off-the road, driving and gesture performance, and perceived mental workload.

RELATED WORK

Visual attention is a measure used in driving studies [24, 11, 33] to evaluate the amount of distraction caused by secondary tasks. Liang *et al.* [19] have shown that duration of glances reliably predicts the risk of crashes. Donmez *et al.* [5] have shown that drivers who received non-visual feedback had shorter average glance durations to the infotainment system screen than driver who received visual feedback. Peng *et al.* [28] investigated driver's lane keeping ability with eyes off road in a naturalistic study and found that it significantly increased lane deviation. Thus, it is imperative to investigate a wide range of non-visual feedback types for mid-air gestures. Potential types include auditory, peripheral visual, and tactile feedback. The effects of eyes-off-the-road are clear, however have not been tested for a range of multimodal mid-air gesture feedback yet.

Auditory Displays

Sterkenburg *et al.* [35] observed significant reductions in eyes-off-the-road time their visual and auditory conditions. They conducted a pilot study on the potential of auditory feedback for in-air gesture control in vehicles. A grid of letters was presented to participants with either visual or auditory-visual feedback. They did not find any significant differences in driving performance between the conditions nor in perceived driver workload. May *et al.* [25] also found no significant difference in driving performance between visual and auditory feedback in their mid-air gesture study. They provided visual (centre console screen) feedback for direct touch input and auditory (speech and non-speech) feedback for mid-air gesture input. Interestingly, the authors did not observe any difference in glance duration between the conditions. Both studies report different observations regarding eyes-off-the-road time and perceived mental workload. Sterkenburg *et al.* [35] found a significant difference in eyes-off-the-road time and mental demand between the visual and visual-auditory condition, which is contrary to May *et al.'s* [25] findings. Thus, research is still required in order to address these questions sufficiently. Further, auditory feedback has been shown to reduce looking away time significantly for in-car interaction [5].

Peripheral Visual Displays

It has been shown that *peripheral* visual feedback in event driven and data-rich environments (i.e. aeroplane cockpits [27]) does not interfere with the performance of the primary visual task [27], and peripheral visual cues are highly effective in conveying information [14, 36]. These findings have led to growing interest in ambient lighting in vehicles. AmbiCar [21, 20, 22] presents ambient light displays in the vehicle to inform the driver about driving related events. The authors tested lane changing behaviour when information was prompted with ambient lighting for rear safety distance violations [22]. They found that ambient lights demanded significantly less visual attention than in-car navigation systems.

Peripheral lights have not been tested as feedback for mid-air gesture interaction in a driving environment. Freeman *et al.* [6] showed that ambient lighting in combination with tactile feedback can successfully surpass the shortcomings of each feedback type and provide an additional modality for mid-air gesture feedback [6] for mobile phones. These findings are promising for an in-car gesture application.

Haptic Displays

Another potential feedback type is haptics. Previous research has examined various types of haptic feedback in cars but very few have utilised haptic feedback for mid-air gestures [30]. Since the hands are generally in constant contact with the steering wheel, research has investigated on providing haptic (mainly vibrotactile) feedback patterns on the steering wheel. However, vibrotactile steering wheels have limitations: even in laboratory conditions, participants struggle to correctly identify the location of the vibration on the wheel [16], especially when tasked with simulated driving [2].

To mitigate the influence of natural car vibrations on the tactile messages, Shakeri *et al.* [34] successfully used cutaneous push feedback which allows for messages to be presented to the palms of drivers. They embedded three solenoid pins in each side of the rim of the steering wheel and presented cutaneous push feedback to the median palm. Cutaneous push feedback recognition accuracy was at 88.7% if presented ipsilaterally and did not cause any significant increase in lane deviation. When gesturing in-car, one hand will remain on the steering wheel thus this is promising for mid-air gesture feedback in cars.

Summary

There is limited research on multimodal feedback for mid-air gestures in cars. The studies we found presented contradictory findings in terms of eyes-off-the-road time and perceived mental workload. It is necessary to find a conclusive answer such that driving can be made safer. Further, this paper presents a study to bring together a novel set of feedback modalities to compare and contrast their effectiveness for in-car gesture feedback for the first time.

EXPERIMENT

We designed a study to investigate the effectiveness of four different types of feedback for four different mid-air gestures in a simulated driving environment. The key aim was to gain insight into the feedback modality which distracts the driver least from the primary driving task and provides the most effective feedback.

Gestures

We used an existing set of gestures based on in-car mid-air gesture design guidelines [41, 7] and ones already available for in-car use (BMW, VW). VW introduced the mid-air swipe left/right gesture in their gesture enabled user interface[3], which we used in our study. BMW use a circular motion to increase/decrease a selection. Our participants had to perform a circular motion with an extended index finger either clockwise (increase) or anticlockwise (decrease). The victory gesture (extended index and middle fingers) was also introduced by BMW and adopted by us. Its purpose is to turn the system / display screen on or off.

In our study, the following types mid-air gestures were presented to participants: 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 (CC2, CC3, CC4), and Circle Anti Clockwise 2, 3, 4 times (CCC2, CCC3, CCC4). We differentiate the gestures depending on hand posture and arm movement, not the number of executions, thus the gesture set consists of five different gestures (SL, SR, V, CC, CCC). Previous research suggests the use of four to eight gestures for in-car gesture interaction [3, 32, 1].

A gesture interaction consists of three parts: the gesture, the execution time, and feedback time. The duration of a single gesture consists of 750 ms execution time and up to 500 ms for gesture feedback. For example, a single swipe motion has to last for at least 750 ms. If the participants were instructed to execute a swipe left four times (SL4), the duration of the interaction is at least 3600 ($4x$750ms gesture execution and $4x(150+50)$ms feedback) ms long (see Table 1). We used the gesture recognition provided by the Leap Motion SDK for swipe and circle gestures. The victory gesture was detected by extending the index and middle finger for 750ms dwell time.

In this study, the gestures were performed with right hand only (as if the car was driving on the right). The gestures were distinct motions completed above the area where the gear stick is located (see Figure 1).

Feedback Types

As found by May *et al.* [25] and during our pilot studies, participants were prone to making accidental gestures by entry of sensing area above the Leap Motion (i.e. interaction box). Thus we implemented our system such that feedback was only

[3]http://www.volkswagen.co.uk/new/golf-vii-pa/explore/r

Gesture	Tone	Duration	Time b/w notes
V on	g#4 → c5	225 ms	50 ms
V off	c5 → g#4	225 ms	50 ms
SL	c4 → c4	150 ms	50 ms
SR	c5 → c5	150 ms	50 ms
CC	c4 → b4	500 ms	-
CCC	b4 → c4	500 ms	-

Table 1. This table shows the auditory feedback used for each gesture. The arrow in Tone describes the transition from one note to the next. Duration describes the length of each note.

provided to the expected gesture types for the current task. If a circular motion was expected, only CC and CCC motions caused system reactions. This allowed participants to make mistakes but did not trigger unwanted system responses [7].

Visual feedback (VF) was provided on the centre console screen to the right of the participant (see Figure 1). VF functioned as baseline for the other feedback conditions. We chose VF because it is the most common feedback type in combination with mid-air gestures in cars (e.g. BMW[4], VW[5]). The GUI design was adapted from the Jaguar Landrover Discovery Sport's GUI in terms of the size of the screen, size of letters, etc. The GUI consisted of a horizontal single scroll bar (from -5 to $+5$). Zobl *et al.* [42] used a horizontal alignment of the bar since swiping motions are performed horizontally. Further, the horizontal bar resembles VW's mid-air gesture GUI in which a swipe left motion moves a song cover from the right to the left to the next song. In our GUI, SL and SR shift the scale of the bar in either direction (SR1: -6 to $+4$, SR2: -7 to $+3$), maintaining the cursor in the centre of the screen (mimicking VW's song swipe); successful CC and CCC motions result in increase/decrease of the cursor on the scale; and V turns the screen on / off.

Auditory Feedback (AF) was presented via earcons [**?**]. The tones used were generated in Audacity[6] and guided by Freeman's audio feedback [6] (see Table 1). After a gesture was recognised by the system, audio feedback was presented in real time. Each gesture had a distinct auditory feedback. Feedback for the V gesture were Windows XP hardware insert/remove sounds. The entire duration of the feedback was 500ms. SR feedback was a double beep at C5 frequency which lasted a total of 350ms, and feedback for SL was a double beep at C4. Feedback for the clockwise circular motion was the increase of the note by an octave, and decrease by an octave for the anti clockwise motion. The feedback for the circular motions lasted 500ms.

Tactile Feedback (TF) was presented via three pins protruding from the steering wheel and provided feedback to the driver's left palm [34]. P1 presents cutaneous push 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 V gesture presented all pins to

[4]https://en.wikipedia.org/wiki/BMW_7_Series_(G11)
[5]http://www.volkswagen.co.uk/new/golf-vii-pa/explore/r
[6]Audacity Version 2.1.2 http://www.audacityteam.org/ Accessed summer 2016

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

the palm for 150ms. Feedback for the SL motion was a sequential display of P2 with a gap of 50ms (total 350ms). Feedback for SR was the sequential display of P1. This feedback resembles the double beep feedback from the audio condition. Feedback for the circular motion mimicked the circling hand: if the driver circled clockwise, the cutaneous push feedback presented P3 $\rightarrow$ P2 $\rightarrow$ P1. Each presentation lasted 166ms, totalling 500ms (see Table 2).

Peripheral Light Feedback (PLF)

The feedback displayed on a single LED strip from the A-pillar on the left side of the driver to the beginning of centre console. The strip was placed behind the steering wheel where the car instrument cluster would be (as proposed by Löcken *et al.* [21]).

V on gesture feedback was presented with an animation of blue lights moving from the ends of the LED strip to the centre. The entire animation lasted for 500 ms. V off feedback was an outward animation of red lights (from centre to the ends of the LED strip). We chose red and blue colours to avoid issues for users who were colour blind.

Feedback for SL and SR gestures mimicked their movements with animations moving from right to left, and left to right respectively. Duration of the entire animation was 350 ms and with each motion, the colours transitioned from red to blue or blue to red, respectively.

Feedback for the CC and CCC gestures were 500 ms long pulses of the entire strip. The colours transitioned from blue to red with a clockwise motion and from red to blue with a anticlockwise motion.

METHOD AND PROCEDURE

The usability laboratory was equipped with 1) a computer, on which the OpenDS simulation was run, 2) a 24 inch screen on which the driving simulator was displayed, 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 web camera located on top of the main screen, 6) three solenoid powered pins protruding from the steering wheel providing feedback to the driver's left palm [34], 7) a capacitive sensor attached to the steering wheel under the driver's right hand, and 8) a 107 cm long LED light strip (see Figure 1). The placements of the individual devices were guided by the measurements of a Jaguar 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.

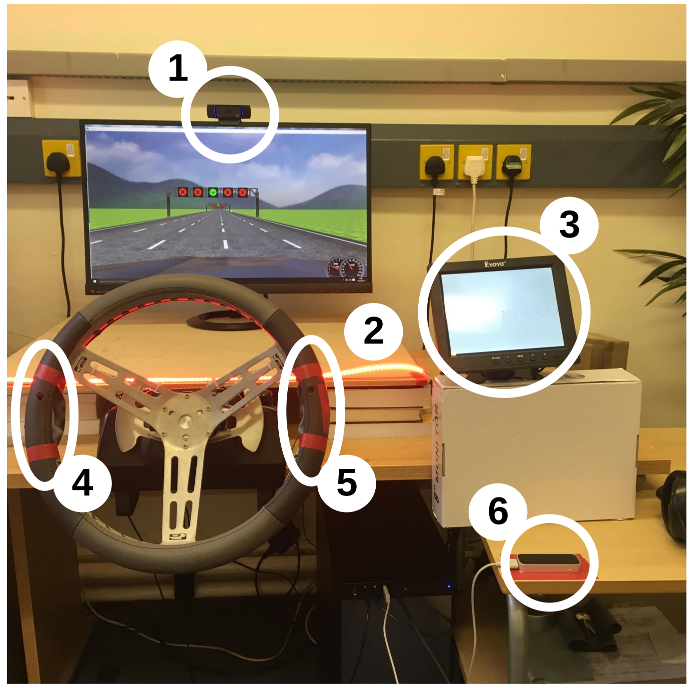

Figure 1. Experimental Setup. The placements of the individual devices were guided by the measurements of a Jaguar Landrover Discovery Sport. 1: webcamera, 2: LED strip, 3: centre console monitor, 4: cutaneous push feedback [34], 4: capacitive sensor, and 6: Leap Motion sensor.

This ensured that the gesture execution area was close to the steering wheel and gear shift, as recommended by Riener *et al.* [32]. The measurements of the interaction box are the Leap Motion's default settings: width: 235.24 mm, height: 235.24 mm, and depth: 147.75 mm.

The webcam recorded the participants' eye gaze while performing the driving and input tasks. Eye gaze direction analysis showed the amount of eyes-off-the-road time caused by the interaction.

Gaze and head pose data were extracted using OpenFace[7], an open source tool for eye-gaze and head pose estimation. An SVM classifier with a linear kernel was trained on 7078 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 91.54% eyes-off-the-road time correctly (10-fold cross validation).

We used OpenDS Version 3[8] to simulate a lane-keeping driving scenario. Participants had to drive the car in the middle lane. The centre of the motorway (middle of middle lane) was used as the zero point for measuring lane deviation.

Hypotheses

H1: Visual distraction from the primary driving task will be significantly decreased in the non-visual conditions;

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

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

H2: There will be a significant difference in lane deviation across the conditions, with lane deviation highest in the visual feedback condition;
H3: Secondary gesture task performance will be significantly higher in the visual feedback condition;
H4: There will be no significant difference in perceived mental workload across the conditions;
H5: Users will prefer the non-visual feedback types over the visual feedback.

Experimental Variables

The Independent Variable was mid-air gesture type. There were four levels: visual, auditory, haptic and peripheral visual. The Dependent Variables were: lane deviation (metres), visual attention to primary task (number of glances at centre console, average duration per glance, average time between glances), number of correct of gestures (% correct), task duration (ms), perceived workload (NASA TLX), and our own questionnaire (demographics, handedness, preferences of feedback).

Participants

Nineteen participants (10 females) ranging from 19 to 35 years of age (μ 24.42 σ 5.79) were recruited via our university's student online forum. Of these 19, nine participants had a UK driving license and 10 a license from elsewhere. A total of 13 participants indicated that they had no prior experience with mid-air gesture interfaces.

Procedure

On arrival, participants were provided with an introduction to the experiment. This included two executions of each mid-air gesture in every condition (4 gestures per condition *x* 2 executions per condition and *x* 4 feedback modalities = 32 gesture executions). After approximately 20 seconds of stabilised driving in the middle lane, the experiment and recordings of the data started. The experiment consisted of four blocks, one block for each feedback condition. During each block, participants executed 30 mid-air gestures (10 *x* SL/SR, 10 *x* V, 10 *x* CC/CCC). Each block lasted approximately 10 minutes. To counter balance for any learning effect, the conditions were ordered via a Balanced Latin Square.

Mid-air gesture instructions were presented via a pop-up message box at the centre top of the main screen above the road. The instructions were also provided via speech through headphones the participants were wearing at all times. The message box was displayed for 3 seconds and the accompanying auditory instructions lasted up to 2 seconds. The auditory instructions were "swipe left/right 2-4", "(counter) clockwise 2-4", "victory" (and during the introduction phase there was an additional "find sweet spot"). The speech instructions were read aloud by a male US American voice (www.cereproc.com/ Voice: Nathan. Accessed 2016-01-31). This reduced the chances of participants missing an instruction.

Once an instruction to gesture was provided, the participants took their right hand off the wheel (whilst steering the simulated car with their left), executed the requested gestures as fast and as accurately as possible, and returned the hand back to the steering wheel. Once the hand was placed back, there was a random interval of 5 - 10 seconds before the next gesture execution. This interval provided an opportunity to return the car to the middle lane and regain stabilised driving, if necessary. After each feedback condition block, participants were asked to fill in a NASA TLX workload questionnaire. Participants were reimbursed with £6 for an hour of their time.

RESULTS

H1: Eye Gaze Behaviour

For all conditions, mean eyes-off-the-road time across conditions and participants (1616 ms) was within the NHTSA guidelines (< 2000 ms). Collected glance data was non-normal. Kruskal-Wallis analysis of eyes-off-the-road time per gesture revealed a significant difference in total glance duration per gesture ($p < 0.001, \chi^2(12) = 86.67$). The V gesture required on average the least time 0.19 sec and CC4 the longest at 0.59 sec. There was a significant difference in gaze duration per condition ($p < 0.001, \chi^2(3) = 188.70$) being highest during the visual condition and lowest in auditory (VF: 0.88 sec, AF: 0.25 sec, PLF: 0.45 sec, TF: 0.32 sec) (see Figure 2).

Effects on Primary Driving Task

There was a significant influence of total duration of glances on primary task ($p < 0.0001, \chi^2(693) = 2009.00$), number of glances per gesture on primary task ($p < 0.001, \chi^2(24) = 1847.223390$), mean glance duration on primary task ($p = 0.000000, \chi^2(739) = 2007.813773$), and mean duration of time between glances on primary task ($p < 0.0001, \chi^2(739) = 2007.813773$).

A multiple regression was run to predict lane deviation from average duration of glances and number of glances. These variables statistically significantly predicted lane deviation, $F(3,95) = 7.1514, R^2 = 0.0105, p < 0.001$. Both variables added statistically significantly to the prediction, $p < .05$. A Kruskal-Wallis analysis revealed significant effects of the combination <*duration of glances, number of glances*> on lane deviation ($p < 0.001, \chi^2(2) = 22.90$) with few and short glances resulting in least lane deviation ($\mu = 0.4696m, \sigma = 0.0113m$), many and short glances in the middle ($\mu = 0.5963m, \sigma = 0.0389m$), and few and long glances resulting in most lane deviation ($\mu = 0.7040m, \sigma = 0.0683m$). Drivers did not glance away from the road more than 5 times per gesture or for more than 2 seconds per glance.

Effects on Secondary Gesture Task

We used the Root Mean Square Error to measure the lane deviation and thus lane deviation data is non-normal. A Kruskal-Wallis analysis of eye gaze behaviour on the secondary task showed that there was no significant difference of total duration of glances on secondary task performance ($p = 0.45, \chi^2(693) = 697.01$), average glance duration on secondary task performance ($p = 0.30, \chi^2(739) = 758.36$), nor average time between glances on secondary task performance ($p = 0.79, \chi^2(490) = 463.74$). However, there was a significant difference in number of glances on secondary task performance ($p = 0.05, \chi^2(24) = 36.53$).

There is a correlation between total glance duration and type of gesture ($p < 0.001$) and number of glances and type of

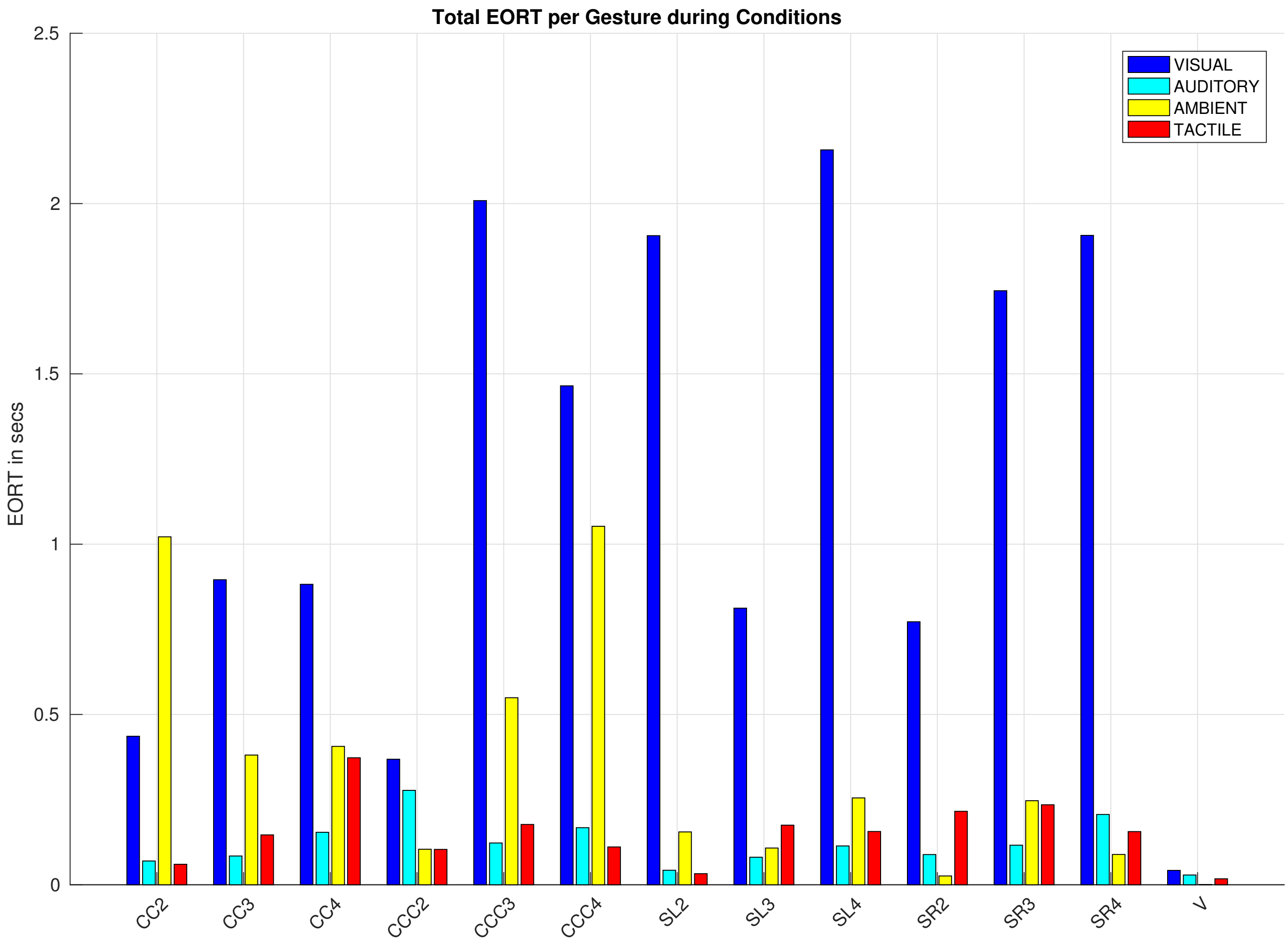

Figure 2. Eyes-Off-the-Road Time per condition grouped by gestures across all conditions.

gesture ($p < 0.001$), i.e. the more time the gesture execution required, the more glances were made towards the gesture sensing device. A multiple regression was run to predict gesture performance from total duration of glances and number of glances. These variables statistically significantly predicted gesture performance, $F(3,95) = 9.41, R^2 = 0.01, p < 0.001$. A Mann-Whitney test showed an impact of gesture duration on lane deviation ($p < 0.001$).

H2: Lane Deviation

A Kruskal-Wallis H test showed that there was no statistically significant difference in lane deviation across conditions, $\chi^2(3) = 2.54, p = 0.4689$ (see Figure 3). Further analysis of lane deviation showed that there was a statistically significant difference in lane deviation across gestures, where $\chi^2(12) = 117.75, p < 0.0001$. Dunn's post-hoc comparison showed there that the V gesture influenced lane deviation least and CCC4 most.

H3: Secondary Task Performance

A gesture was classified as correct if the executed gesture was performed as instructed. Overall, 69.06% of instructed gestures were executed correctly with the V gesture the best at 96.25% (see Figure 5). A Kruskal-Wallis H test showed that there was a statistically significant difference in performance across gestures, ($\chi^2(12) = 330.33, p < 0.0001$) (see Figure 5) and conditions ($\chi^2(3) = 29.61, p < 0.0001$) (see Figure 4). Visual and auditory feedback yielded highest correct gesture executions.

There was no statistically significant difference in gesture duration depending on feedback type, $\chi^2(3) = 9.34, p = 0.02505$. There was no significant difference in duration of gestures execution across the conditions ($p = 0.4210, \chi^2(3) = 2.81$). Average task durations were below the 15s rule [8].

Further, there was a significant difference in correct gesture execution between the swipe and circle motions across conditions ($p = 0.0155, \chi^2(3) = 10.38$), with the SR gestures performing worse (44.26% correct) than SL (59.21%) or CC (69.01%) / CCC (76.54%) gestures.

Participants performed a correction gesture for 21.70% of instructions. A correction gesture was executed by the participant if an unintended SR motion was recognised by the system during an SL trial, for example resetting the hand to a

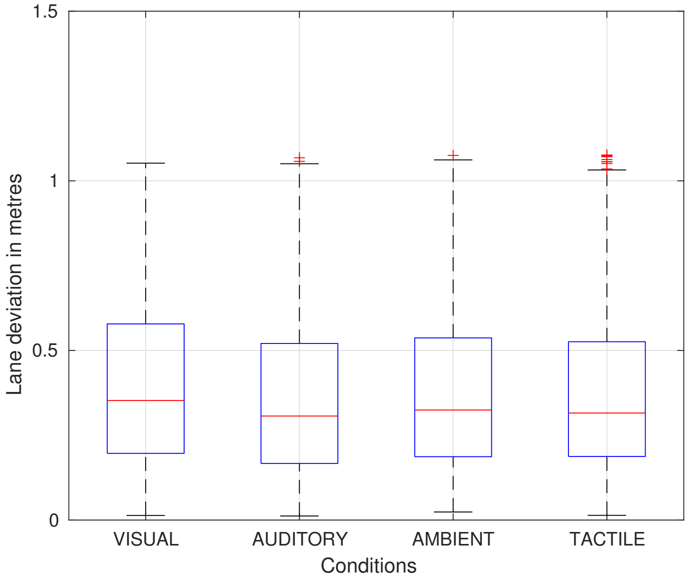

Figure 3. Lane deviation across conditions in metres. VF: μ 0.63 m $\pm$ 0.5 m; AF: μ 0.44 m $\pm$ 0.27 m; PLF: μ 0.65 m $\pm$ 0.63 m; TF: μ 0.48 m $\pm$ 0.29 m.

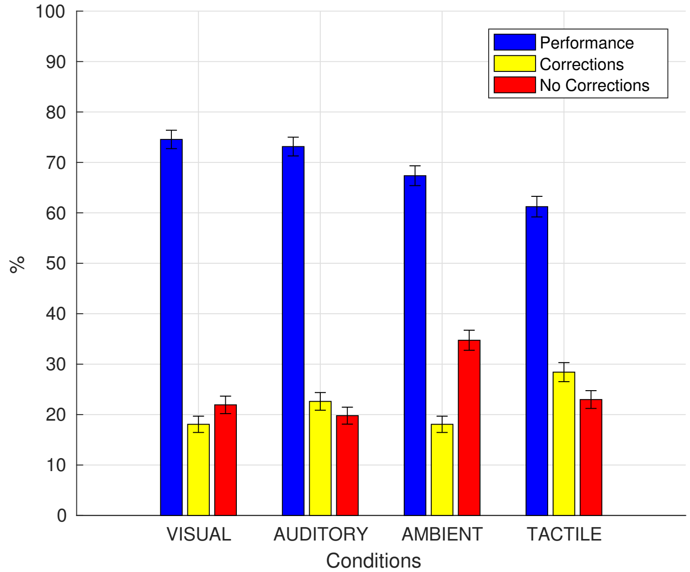

Figure 4. Secondary task performance per condition. Three bars for each condition: left: average performance of gestures, middle: average percentage of attempted corrections, right: average percentage of no attempts at correcting.

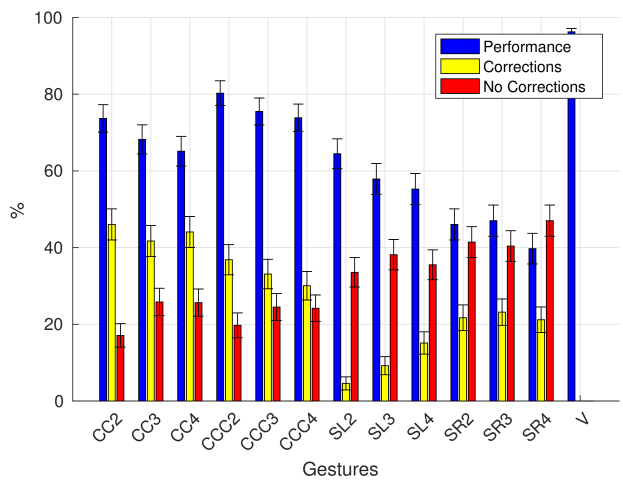

Figure 5. Secondary task performance per gesture.

position to execute the next SL gesture required moving the hand from left to right; if this positioning motion was executed within the interaction box of the sensing device, an SR gesture was recognised; therefore, a correction gesture was necessary. Of these correcting gestures, 75.51% resulted in successful gesture instruction completion. Thus, a total of 6.07% of gestures remained unsuccessful after performing a correction gesture. Feedback condition significantly influenced correcting behaviour ($p < 0.001, \chi^2(3) = 25.15$) (see Figure 4), as does gesture type ($p < 0.001, \chi^2(12) = 337.80$) (see Figure 5). Dunn's post-hoc test showed that TF influenced correcting behaviour the most and VF the least. Further, feedback had no influence on correction behaviour of V, but more effect on swiping gestures than circular motions. However, 24.87% of instructions were not executed correctly and not corrected. As a Kruskal-Wallis test shows, not correcting gestures is also feedback dependent ($p < 0.001, \chi^2(3) = 41.23$). Dunn's post-hoc comparison shows that circular motions remained more un-corrected than swiping motions.

H4: Subjective Workload

Analysis of the NASA TLX questionnaire revealed a significant difference in mental demand ($\chi^2(2) = 6.57, p = 0.04$), with TF having the highest level. There were no significant differences in the remaining measures: physical demand ($\chi^2(2) = 1.00, p = 0.60$), temporal demand ($\chi^2(2) = 1.65, p = 0.44$), performance ($\chi^2(2) = 1.71, p = 0.42$), effort ($\chi^2(2) = 0.70, p = 0.40$), and frustration ($\chi^2(2) = 7.01, p = 0.07$).

H5: Preferences

Each participant could rank the feedback types from most to least preferred. Analysis of our questionnaire showed that 38.9% of participants preferred audio feedback, followed by 33.3% preferring both peripheral visual and tactile feedback. Visual feedback was ranked as least preferred feedback type by 44.4% of the participants.

DISCUSSION AND CONCLUSIONS

In this paper, we investigated the effects of different types of feedback on mid-air gesture interaction when users were driving in a simulator. The results provide insights into the effects of feedback type on the primary driving task and the secondary gesturing task.

Our results suggest that providing non-visual feedback for mid-air gesture input is promising since it reduces eyes-off-the-road time significantly (see Figure 2), with auditory and tactile feedback resulting in the least time looking away from the road. Therefore, we accept hypothesis H1. However, due to our design choices to only provided feedback to the expected gesture, our results are a best case scenario as eyes-off-the-road time can increase in the case of error states. A follow up study will look into providing feedback to any recognised gesture. We plan to implement a 500ms dwell time after entry of the interaction to avoid accidental gestures [25]. The duration of glances off the road had a significant impact on lane deviation across all conditions. Further, we found that the duration of glances had a greater effect on lane deviation than the number of glances. This means that many short glances away from the primary task have less negative impact on lane deviation than fewer but longer glances. Interestingly, feedback type had no significant effect on lane deviation, thus hypothesis H2 is rejected. This result may be due to:

1. the secondary task duration being too short (average glance duration was less than 2000ms). Research has shown that if drivers' glances off the road are shorter than 2 seconds, it has no significant effect on lane deviation [17];

2. the primary task was not challenging enough and over time participants found the optimal steering wheel position for least lane deviation.

It will be interesting to see how a more challenging primary task (e.g. changing lanes) might influence lane deviation during a gesturing task. We will test this in a future study.

Our analysis of type of gesture revealed that gestures which require more time to execute (e.g. CCC4), influenced eyes-off-the-road time significantly across all conditions. The more time the gesture required for execution, the greater the glance duration and number of glances off the road. Further, the longer a gesture required for execution, the less successful it was. The wider the time window for gesture operation, the more units of movement are executed, and the more mistakes can be made. In other words, the longer the duration, the more consistently accurate the movement has to be. A long term effect of mid-air gesture execution might be fatigue of the arm and shoulder since participants had to move their entire arms [12]. This means that in-car mid-air gestures should be designed such that they require little arm and shoulder movements and are short, such as the V gesture.

Hypothesis H3 is accepted since we found significant differences between secondary task performance depending on feedback condition. Gestures were performed best during the VF condition (see Figure 4). We believe the high performance rate during VF is due to participants being familiar with visual feedback in general. Gesture performance during AF was second highest, which is in accordance with the literature suggesting that auditory feedback is a suitable alternative for visual feedback. Interestingly, feedback type influenced the willingness of participants to correct a gesture that was wrongly executed/classified, being highest in the tactile condition. Ambient feedback caused significantly higher rates of non-correction gestures. We will investigate in future work whether this effect is due to the used patterns and colour combinations.

Further analysis revealed a significant difference in secondary task performance across gestures. The V gesture yielded the highest performance accuracy with 96.26%. This might be due to the V gesture consisting of a single discrete and static motion. Other gestures consisted of two or more motions (e.g. SL2, CC2). Swipe motions performed worst, especially SR. CC motions performed worse than CCC gestures. This might be due to the SR and CC gestures being motions where the arm is moving away from the torso of the driver. This "away" movement might have caused greater arm and shoulder fatigue [12]. The difference between the circular motion and the swipe motion was the nature of their continuity. A CC2 motion is one continuously performed gesture. With SL2, the user has to return the hand to the start point and swipe again. This interruption of rhythm — the new alignment of the hand inside the interaction box — might have caused the different performance rates between the gestures.

We found a significant difference in perceived mental demand across the conditions, thus H4 is rejected. TF caused highest perceived mental demand, and AF caused least. With haptic feedback presented to the palm, the sensors for the feedback also need to execute driving manoeuvres simultaneously, which also gives haptic/mechanical feedback to drivers (e.g. wheel turning). This feedback mechanism directly competes with driving in the sensory stage, which might make this feedback mechanism high in cognitive demand. In addition, haptic patterns and motion on the steering wheel are not commonly used which caused uncertainty in interaction. The unfamiliarity of the feedback mechanism with users — and the resulting uncertainty that was cause — was expressed as "it was not as easy to differentiate between the different swipe types than the others [feedback types]". Finally, we accept hypothesis H5 since VF was ranked least preferred. AF was ranked most preferred followed by PLF. Research has shown that multimodal feedback reduces perceived mental workload [18]. Thus, we will investigate the effects of multimodal mid-air gesture feedback on eye gaze behaviour, driving behaviour, gesturing behaviour, and perceived mental workload.

This paper contributes two new feedback techniques for mid-air gesture interaction in a driving scenario: peripheral visual feedback and cutaneous push feedback. Our results show that the presented feedback techniques cause significantly less eyes-off-the-road time than the use of visual feedback thus have the potential to make driving safer.

ACKNOWLEDGMENTS
This research was funded by the Horizon 2020 HAPPINESS project.

REFERENCES

1. Micah Alpern and Katie Minardo. 2003. Developing a car gesture interface for use as a secondary task. *CHI'03 extended abstracts on Human factors in ...* (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. N. 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

4. T. a. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. a. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R.R. Knipling. 2006. The 100-Car naturalistic driving study phase II - Results of the 100-Car field experiment. *Chart* April (2006), No. HS–810 593. DOI: http://dx.doi.org/DOTHS810593
5. Birsen Donmez, Linda Ng Boyle, and John D Lee. 2010. Differences in Off-Road Glances: Effects on Young Drivers' Performance. *Journal of Transportation and Engineering* 136, 5 (2010), 403–409. DOI: http://dx.doi.org/10.1061/(ASCE)TE.1943-5436.0000068
6. Euan Freeman, Stephen Brewster, and Vuokko Lantz. 2016. Do That , There : An Interaction Technique for Addressing In-Air Gesture Systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, 2319–2331. DOI: http://dx.doi.org/10.1145/2858036.2858308
7. Thomas M Gable, R May May, and Bruce N Walker. 2014. Applying Popular Usability Heuristics to Gesture Interaction in the Vehicle. In *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications 2014*. 1–7. DOI:http://dx.doi.org/10.1145/2667239.2667298
8. Paul Green. 1999. The 15-second rule for driver information systems. *Proc of the ITS America* Cd (1999), 1–9. http://www.umich.edu/~driving/publications/ITSA-Green1999.pdf
9. 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.
10. Paul Green. 2004. Driver distraction, telematics design, and workload managers : safety issues and solutions. *Proc. Int. Congr. Transp. Electron* (2004), 165 –180. DOI: http://dx.doi.org/10.1167/3.9.157
11. Marigo Heijboer, Elise van den Hoven, Bert Bongers, and Saskia Bakker. 2016. Facilitating peripheral interaction: design and evaluation of peripheral interaction for a gesture-based lighting control with multimodal feedback. *Personal and Ubiquitous Computing* 20, 1 (2016), 1–22. DOI:http://dx.doi.org/10.1007/s00779-015-0893-5
12. 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.
13. 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
14. John Jonides. 1981. Voluntary Versus Automatic Control Over the Mind'S Eye'S Movement. *Attention and performance IX* 9 (1981), 187–203.
15. Raine Kajastila and Tapio Lokki. 2013. Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human Computer Studies* 71, 5 (2013), 627–640. DOI: http://dx.doi.org/10.1016/j.ijhcs.2012.11.003
16. Dagmar Kern, Paul Marshall, Eva Hornecker, Yvonne Rogers, and Albrecht Schmidt. 2009. Enhancing navigation information with tactile output embedded into the steering wheel. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5538 LNCS, 1 (2009), 42–58. DOI: http://dx.doi.org/10.1007/978-3-642-01516-8
17. S.G. Klauer, T. A. Dingus, V. L. Neale, J.D. Sudweeks, and D.J. Ramsey. 2006. DOT HS 810 594 The Impact of Driver Inattentionnn On Near-Crash/Crash Risk::: An Analysis Using the 100-Car Naturalistic Driving Study Data. *National Highway Traffic Safety Administration* (2006), 226. http://hdl.handle.net/10919/55090
18. Ju-Hwan Lee and Charles Spence. 2008. Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions. *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction-Volume 1* (2008), 185–192.
19. Yulan Liang, John D. Lee, and Lora Yekhshatyan. 2012. How Dangerous Is Looking Away From the Road? Algorithms Predict Crash Risk From Glance Patterns in Naturalistic Driving. In *Human Factors*. 1104–1116. DOI: http://dx.doi.org/10.1177/0018720812446965
20. Andreas Löcken. 2015. AmbiCar : Ambient Light Patterns in the Car. In *AutomotiveUI*. Nottingham, UK.
21. Andreas Löcken, Heiko Müller, Wilko Heuten, and Susanne Boll. 2013. AmbiCar: Towards an in-vehicle ambient light display. In *AutomotiveUI 2013*. 107–108.
22. 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
23. Sebastian Loehmann, Martin Knobel, Melanie Lamara, and Andreas Butz. 2013. Culturally independent gestures for in-car interactions. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8119 LNCS, PART 3 (2013), 538–545. DOI: http://dx.doi.org/10.1007/978-3-642-40477-1{_}34
24. Keenan R May, Thomas M Gable, and Bruce N Walker. 2013. Comparing Three Novel Multimodal Touch Interfaces for Infotainment Menus. *Auto UI 2013* (2013), 100–107. DOI: http://dx.doi.org/10.1145/2516540.2516559

25. Keenan R May, Thomas M Gable, and Bruce N Walker. 2014. A Multimodal Air Gesture Interface for In Vehicle Menu Navigation. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '14*. 1–6. DOI:http://dx.doi.org/10.1145/2667239.2667280

26. S Morrison-Smith and J Ruiz. 2014. Using Audio Cues to Support Motion Gesture Interaction on Mobile Devices. *CHI'14 Extended Abstracts on Human Factors . . .* (2014), 1621–1626. DOI: http://dx.doi.org/10.1145/2559206.2581236

27. 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

28. Yiyun Peng, Linda Ng Boyle, and Shauna L. Hallmark. 2013. Driver's lane keeping ability with eyes off road: Insights from a naturalistic study. In *Accident Analysis and Prevention*. DOI: http://dx.doi.org/10.1016/j.aap.2012.06.013

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. Asm Rahman, Jamal Saboune, and a. El Saddik. 2011. Motion-Path Based in Car Gesture Control of the Multimedia Devices. *Proceedings of the First ACM International Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications* (2011), 69 – 76. DOI: http://dx.doi.org/10.1145/2069000.2069013

31. 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.

32. 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. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '13)* (2013), 14–21. DOI: http://dx.doi.org/10.1145/2516540.2516544

33. 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

34. 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

35. Jason Sterkenburg, Steven Landry, Myounghoon Jeon, and Joshua Johnson. 2016. Towards an in-vehicle sonically-enhanced gesture control interface: a pilot study. In *The 22nd International Conference on Auditory Display (ICAD-2016)*. 0–3. DOI: http://dx.doi.org/10.21785/icad2016.015

36. 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

37. Louis Tijerina. 2001. Issues in the Evaluation of Driver Distraction Associated with In-Vehicle Information and Telecommunications Systems. (2001), 1–13.

38. J.B.F. Van Erp and H.a.H.C. Van Veen. 2001. Vibro-tactile information presentation in automobiles. *Proceedings of Eurohaptics* (2001), 99–104. DOI: http://dx.doi.org/10.1186/1743-0003-10-21

39. 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

40. Christopher D Wickens. 2008. Multiple resources and mental workload. *Human factors* 50, 3 (2008), 449–455. DOI:http://dx.doi.org/10.1518/001872008X288394

41. Yu Zhang and Linda Angell. 2014. Pointing Towards Future Automotive HMIs: The Potential for Gesture Interaction. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '13* 22, 3 (2014), 22–29. DOI: http://dx.doi.org/10.1016/j.apergo.2013.10.013

42. Martin Zobl, Ralf Nieschulz, Michael Geiger, Manfred Lang, and Gerhard Rigoll. 2004. Gesture components for natural interaction with in-car devices. *The Hand* (2004), 448–459. DOI: http://dx.doi.org/10.1007/978-3-540-24598-8{_}41