Feel the Noise: Mid-Air Ultrasound Haptics as a Novel Human-Vehicle Interaction Paradigm

David R. Large[1], Kyle Harrington[1], Gary Burnett[1], Orestis Georgiou[2]

[1]Human Factors Research Group, University of Nottingham, UK

[2]Ultrahaptics Ltd., Bristol, UK

ABSTRACT

Focussed ultrasound can be used to create the sensation of touch in mid-air. Combined with gestures, this can provide haptic feedback to guide users, thereby overcoming the lack of agency associated with pure gestural interfaces, and reducing the need for vision – it is therefore particularly apropos of the driving domain. In a counter-balanced 2×2 driving simulator study, a traditional in-vehicle touchscreen was compared with a virtual mid-air gestural interface, both with and without ultrasound haptics. Forty-eight experienced drivers (28 male, 20 female) undertook representative in-vehicle tasks – discrete target selections and continuous slider-bar manipulations – whilst driving. Results show that haptifying gestures with ultrasound was particularly effective in reducing visual demand (number of long glances and mean off-road glance time), and increasing performance (shortest interaction times, highest number of correct responses and least 'overshoots') associated with continuous tasks. In contrast, for discrete, target-selections, the touchscreen enabled the highest accuracy and quickest responses, particularly when combined with haptic feedback to guide interactions, although this also increased visual demand. Subjectively, the gesture interfaces invited higher ratings of arousal compared to the more familiar touch-surface technology, and participants indicated the lowest levels of workload (highest performance, lowest frustration) associated with the gesture-haptics interface. In addition, gestures were preferred by participants for continuous tasks. The study shows practical utility and clear potential for the use of haptified gestures in the automotive domain.

Author Keywords

Ultrasound; mid-air haptics; gestures; visual demand; touchscreen; driving.

INTRODUCTION

Touchscreens are currently recognised as the de facto interface for in-vehicle, secondary controls. This is largely because they offer flexibility in design, and can be 'easy' and 'intuitive' to use, providing direct access to diverse information and functionality from complex systems [1, 2]. Moreover, there is evidence to suggest that touchscreens may be more effective when undertaking certain tasks, such as simple menu selection, in the automotive domain, and some drivers have indicated a strong preference for touchscreens, when they have been evaluated alongside to other physical devices [3]. However, touchscreens inherently demand visual attention, and concerns have also been raised regarding their potential to elevate physical and mental workload when used concurrently with driving [1, 4, 5]. This is because touchscreens typically necessitate precise finger-touch inputs directly onto

the display, yet present a uniform surface to users that is absent of any tactile guidance. Consequently, a driver's eyes are required to see what their hand is doing. This naturally results in an increase in eyes-off-road time, which in turn is likely to detriment driving performance and vehicle control, ultimately contributing to the elevation of accident risk, particularly when off-road glances extend beyond 2.0-seconds [6, 4]. In addition, the utility and usability of touchscreens in the automotive domain may be further undermined by the vibrations and lateral accelerations occurring while the vehicle is in motion, especially on non-ideal road conditions [7]; these problems may be exacerbated for older drivers due to changes in perceptual and motor skill capabilities that accompany the aging process [2]. As such, a more robust, inclusive solution, that demands less visual attention and provides scope to overcome road-induced perturbations, is required.

Background

There is a corpus of literature pertaining to novel interaction techniques and artful designs that aim to minimise the visual demand, reduce workload and increase the usability and utility of touchscreens in vehicles (for example, see: [1, 8, 9]). In addition, attempts have been made to augment touchscreens with haptic information, so that drivers may be able to 'feel' the location of the on-screen targets [10]. While successes have been reported for all of these approaches, they remain fundamentally limited in that ultimately they still rely on the use of the touchscreen (or an equivalent interactive display) as the primary user interface. The reliance on using a visual stimulus and visually-oriented controls means that even with highly proficient interaction design, drivers may continue to direct visual attention towards the screen or physical device, whether it is required to or not. Interface designers and user-experience professionals have therefore explored more radical approaches that have the potential to eliminate a physical interface completely – and in so doing, remove the need (or temptation) for vision. For example, recent efforts have explored the use of hand gestures to control in-vehicle secondary devices and systems.

As an interaction technique, gestures are an attractive solution – they are not bound to any surface and do not normally require vision. They are also generally considered to be 'intuitive' – in that they are often inspired by natural, everyday interactions between humans – and are therefore popular for novice users or infrequently used applications. As such, the use of gestures as an input modality has been explored in the automotive domain, with human-vehicle interaction (HVI) gesture systems already attracting research efforts [11, 12]. Moreover, commercially-available systems are now available in some high-end marques (e.g. BMW's 7-series, Jaguar XF Sportbrake, Mercedes-Benz S-Class). In this context, gestures have been employed in a number of distinct ways, for example, to indicate intent, control functions, signify a response to a question or prompt, or to initiate a dialogue with the vehicle [13]. Automotive gesture-related research has tended to be positive, suggesting that drivers like gesture control, with gestural interfaces being rated more highly (i.e. more 'enjoyable') than conventional controls [14, 15, 16]. However, task success rate and task-time are typically highly variable, particularly for more complex gestures or tasks, with the highest successes and shortest task times associated with 'simple' and 'natural' gestures. Indeed, scholars have warned about the additional time required to learn more complex gestures [17], as well as potential differences in cultural connotations associated with their use [14]. A further problem is that pure gestural interfaces lack tactile feedback, a key ingredient towards creating a sense of agency – the subjective experience of voluntary control over one's actions [18], and this increases the potential for inadvertent operation. This paper explores one proposed solution to mitigate this problem, that is, to provide a 'mid-air' haptic sensation to help guide the user's hand and provide task feedback.

Whereas early attempts to provide haptic sensations typically required the user to wear an additional device to detect or guide their hands or fingers [19, 20], Hoshi et al. [21] proposed the use of ultrasound to stimulate receptors in the human body. This non-evasive technique, which has been further developed by Carter et al. [22], uses the principle of acoustic radiation force, that is, the force generated when ultrasound is reflected. Therefore, no physical contact or additional wearable device is required. Instead, ultrasound is focused directly onto the surface of the skin, where it induces vibrations in the skin tissue. The displacement triggers mechanoreceptors within the skin generating a haptic sensation [23], that has been described as feeling like gentle, pressurised airflow on the palm of the hand [24]. By focussing the acoustic radiation force and modulating this at 200Hz (the frequency at which vibrations are perceptible by human skin mechano-receptors), a sensation perceivable as a virtual, mid-air three-dimensional shape or 'target' can be created, with its 'physical' size defined by the ultrasound wavelength used to create it. Moreover, by synchronising the phase delays of multiple transducers, different focal points can be generated which are perceivable as discrete mid-air targets, or a multiple-item array [22].

Given its mid-air location, selecting a virtual, haptic target therefore naturally lends itself to hand gestures. Moreover, by assigning different functions to each target, the illusion of a virtual, mid-air interface (or 'buttonscape') can be created, whereby the user could feel multiple, discrete elements. A further advantage is that users are not required to remember the semantic meaning (or 'operational language') of multiple gestures, or the locations of different interface elements, but are able to use their sense of touch to locate a particular target within an array, based on focussed haptic feedback, and then use a simple, intuitive gesture, such as a 'press' or 'grab' to activate or select it. As such, existing interface layouts (e.g. a 2×3 structured array of buttons) can be created as a virtual buttonscape within 3-dimensional space.

Using ultrasound to stimulate receptors in the human body has already been successfully and commercially deployed in entertainment and gaming [25], and preliminary results support its application in the driving domain [24, 26, 27]. Nevertheless, there is currently a dearth of knowledge regarding the overall impact of mid-air ultrasound haptics within the driving domain and its potential use as a novel human-vehicle interaction paradigm. The current study aims to address this by comparing the use of gestures augmented with mid-air haptic feedback with a traditional touchscreen interface, and posits the following hypotheses:

1. Overall, using gestures to execute secondary in-vehicle tasks will require less visual demand, provide a more positive user experience and demand lower workload, compared to a traditional touch-surface interface.
2. The addition of mid-air ultrasound haptic feedback will reduce the visual demand, enhance the user experience and reduce workload associated with both gesture and touch-surface interfaces.
3. The lowest visual demand, most positive user experience and lowest workload will be associated with gestures augmented with mid-air ultrasound haptic feedback.
4. There will be no detriment to primary (driving) and secondary task performance when using gestures augmented with mid-air ultrasound haptic feedback.

METHOD

Overview and Aims

The aim of the current study was to explore the use of gestures augmented with mid-air haptic feedback created using ultrasound as a novel human-vehicle interaction paradigm in a driving context. In a counter-balanced 2×2 experimental design, the study therefore explored independent variables of 'gesture' and touch-surface ('touch'), both with and without ultrasound haptic feedback. Following training and familiarisation for each of the four techniques, each participant was asked to undertake four experimental drives with the order of exposure counterbalanced between participants (Table 1). During each drive, participants were presented with one of the different interaction techniques for the entirety of the drive, and were required to conduct a series of representative, in-vehicle tasks, comprising discrete target selections (i.e. button selections) and continuous target manipulations using a graduated slider-bar.

Table 1. Four conditions used in the 2x2 experimental design

Condition	Description
Touch No Haptics (TN)	Tasks were completed using a conventional touchscreen with no haptic feedback.
Touch with Haptics (TH)	Tasks were completed using the touchscreen enhanced with ultrasonic haptic feedback, aiming to guide the participant's hand towards the touchscreen (i.e. haptic feedback was initiated when their hand was in close proximity to the screen).
Gesture No Haptics (GN)	Tasks were completed using simple gestures (identified using a Leap Motion sensor) but without haptic feedback.
Gesture with Haptics (GH)	Tasks were completed using the same gestures as GN, enhanced by haptic sensations.

Participants

Forty-eight people took part in the study: 28 male, 20 female, with ages ranging from 23 to 70 years (mean age: 35.4). An approximately representative proportion of participants (8) were left-handed, although this was not a controlled independent variable. All participants were experienced and active drivers in the UK (mean number of years with licence: 13.8, average annual mileage: 7091, range: 10k-20k). Participants were self-selecting volunteers, who responded to advertisements placed around the University of Nottingham campus, and were reimbursed with £10 (GBP) of shopping vouchers as compensation for their time. All participants provided written informed consent before taking part.

Apparatus

Driving Simulator

The study took place in a medium-fidelity, fixed-based driving simulator at the University of Nottingham (Figure 1-left). The driving simulator comprises a right-hand drive Audi TT car positioned

within a curved screen, affording a 270 degrees forward and contiguous side image of the driving scene via three overhead HD projectors, together with rear and side mirror displays. A Thrustmaster 500RS force feedback steering wheel and pedal set are integrated with the existing Audi primary controls, and a bespoke dashboard presented on a 7-inch LCD screen, replacing the original Audi instrument cluster. The simulated driving environment was created using STISIM (v3) software, and comprised a three-lane UK motorway with both sides of the carriageway populated by moderate levels of traffic, authentic road signage and geo-typical roadside terrain.

Figure 1. Medium-fidelity driving simulator showing motorway scenario (left) and participant interacting with mid-air haptics (right), with transducer array (beneath hand), Leap Motion camera (beneath wrist) and visual (touchscreen) display with button 'one' of three selected (in red).

Touchscreen

A 7-inch resistive touchscreen was installed in the centre console of the car (Figure 1-right), where it acted as the primary interface during the Touch conditions. The touchscreen also provided visual feedback during Gesture conditions, where it was arguably of greater value during the gesture-only interactions (i.e. in the absence of mid-air haptic feedback). However, when gestures were augmented with mid-air ultrasound haptics, the visual feedback (and therefore the touchscreen) was not strictly required. Nevertheless, the touchscreen remained in all conditions to ensure a balanced comparison between gesture and touch conditions. Target selections were made by touching and releasing the required target for button selections, and by touching and dragging (while maintaining contact) an on-screen pointer during the continuous, slider-bar tasks. For button selection tasks, target arrays were presented as 2, 3 or 4-item structured menus with targets numbered consecutively, left to right and top to bottom (e.g. Figure 2). For continuous tasks, a single slider bar was presented with the pointer initially placed at the centre of the scale; scale increments and decrements were also shown numerically below the slider bar (Figure 2).

Gestures and Haptics

Hand gestures, interaction techniques, and haptic textures were developed in collaboration with Ultrahaptics Ltd. using the Ultrahaptics Touch Development Kit (TDK) integrated with Unity (a software development platform – see: https://unity.com/). The TDK utilises a 14x14 array of ultrasonic transducers operating at 40 kHz to create an effective ultrasonic focusing region approximately 90 degrees wide and up to 50cm high. A LEAP Motion controller (comprising two monochromatic infrared

cameras, three infrared LEDs, and machine vision proprietary software) was used to detect and track the driver's palm position, orientation, posture and motion, and to localise sensations. Sensations were focussed at the centre of the user's palm – this was to improve the accuracy of detecting the presence of the user's hand, rather than relying on a smaller target, such as their fingertip. The TDK was installed in the centre of the car (between driver and passenger seats), as might reasonably be expected for such a device based on discussions with automotive design experts (Figure 1-right). This ensured that the effective haptified hand gesture interaction region was comfortably within the expected trajectory of the driver's arm and hand and not physically obscured by other in-vehicle devices – in practice, the gear stick was removed for the study to further enable this.

For *button selection tasks*, active targets utilised four focal points to create the four vertices of square targets that were fixed in the z-plane (i.e. sensations were focussed at a specific height, approximately 30cm above the array), but not bound to a single plane in x-space. Mid-air haptic feedback was activated when the user's hand entered the active target area and this also changed the colour of the related target on the touchscreen (see Figure 1); the edges and spaces between targets were consequently defined by the absence of haptic feedback. In the absence of haptic feedback (i.e. the gesture-only condition), users were required to identify the active region through proprioception, but could also confirm success using the visual feedback provided on the touchscreen. To make a target selection during the gesture conditions, participants were required to make a simple downward movement of their open hand to simulate a button press.

For *slider-bar tasks*, a pointer was initially placed in the centre of the slider bar. When using gestures, participants initially selected the pointer by making an open-palm gesture. This also determined the centre of the slider-bar in 3-dimensional space. Thereafter, the slider-bar interaction enabled approximately 20cm lateral movement in either direction along the x-plane. Participants were thus required to move their open hand right or left to increase or decrease the value, and then 'grab' the target (by making a fist) to make a selection. Where appropriate, haptic feedback was provided to signify incremental changes (i.e. separate pulsed sensations were generated as the participant's hand passed each consecutive value).
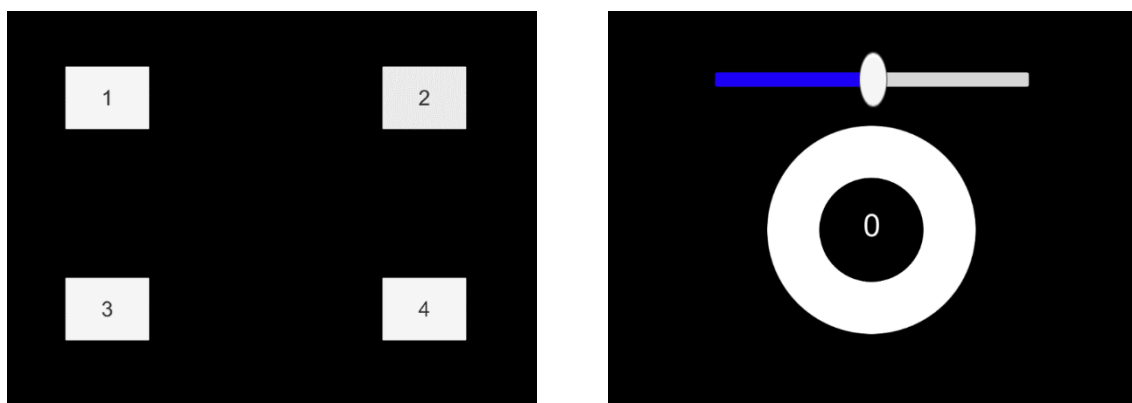

Figure 2. On-screen representations for discrete button-selection (left) and continuous slider-bar tasks (right)

Experimental Design

Each participant was provided with training and familiarisation using the touchscreen and gestures both with and without haptic feedback. This occurred firstly whilst stationary seated in the simulator buck, and secondly, while driving. For each technique, the participant was required to demonstrate three consecutive successful interactions (i.e. using the correct behaviour and selecting the correct target without any false activations), before they were deemed to be competent.

The 'car following' paradigm was adopted as the primary driving task [28]. At the start of the scenario, a yellow car emerged ahead of the participant's vehicle on the motorway. This initially accelerated in unison with the participant's vehicle, and then travelled at a variable speed (between 65 and 75mph) for the rest of the drive. Participants were instructed to follow the lead car, which remained in lane one, at a distance that they deemed to be safe and appropriate. While following the lead vehicle ('the primary task'), participants were asked to interact with the in-vehicle interface ('the secondary task') using each of the four different techniques (Table 1).

For *button selection tasks*, participants were required to select and 'press' a specific target item (either by touching the screen or using a gesture) in response to a pre-recorded voice message. For example, "On this three-item menu, select two". For *slider-bar tasks*, participants were required to select the pointer and then increase or decrease the value by a specified amount, up to five increments in either direction (for example, "Please increase the value by three").

Participants were asked to behave naturally when undertaking the tasks using the different interaction techniques – they were not told to attempt to complete tasks with the gestural interface without looking, for example. Participants completed two repeats for all possible targets and configurations during each condition, for both button selection and slider bar tasks, culminating in 18 button presses and 20 slider tasks per drive. Each drive therefore lasted approximately 8-10 minutes, and the entire study took about 1.5 hours for each participant.

Measures and Analysis Approach

During each drive, participants wore SMI eye-tracking glasses (ETG) (visible in Figure 1-right), with their gaze data subsequently analysed using semantic gaze mapping. Off-road ('in- vehicle') glances comprise all visual behaviour directed towards the interface (i.e. the screen and/or transducer array), and were defined from the moment the driver's gaze started to move away from the road, to the time it returned to the road scene. Thus, a single off-road glance could comprise several fixations towards the physical devices inside the vehicle. Visual behaviour is subsequently determined at an aggregate level (for example, the total number of glances per drive/task type), and reported as mean values per task for total glance time (TGT), mean glance duration (MGD) and number of glances (NoG), in line with existing guidelines [4, 29]. The number of off-road glances longer than 2.0-seconds are also presented (NoG>2).

Primary task (driving) performance data were captured from the STISIM simulation software at tenth-of-a-second intervals. These data were used to calculate standard deviations of lane position (SDLP) and headway (SDHW) – recognised performance metrics associated with driving and driver distraction [30] – which were subsequently compared between conditions. Secondary task performance was determined through measures of accuracy (percentage of correct button/slider-bar selection, and cumulative slider-bar 'overshoots') and task-time. For slider-bar tasks, the task-time comprised the

reaction time and interaction time. Reaction time is defined as the time from the delivery of the task instruction to the start of the interaction, that is, when the Leap Motion sensor detected the hand and initiated the haptic sensations, or the participant made contact with the touch surface. Interaction time is defined as the time that each participant took to manipulate the interface and make their selection using either the appropriate mid-air gesture or touch. For button selection tasks, only total task time was recorded given that for button selections using the touch surface, there was no discernible 'interaction' time. Task time is reported as the mean value for each task-type and interaction technique across each drive.

Participants provided subjective ratings of workload using the NASA-Task Load Index (NASA-TLX) questionnaire [31] associated with each technique. To measure the affective experience ('user-experience'), the self-assessment manikin questionnaire (SAM) [32] was employed. NASA-TLX and SAM are reported per drive and therefore encompass ratings associated with both button-selection and slider-bar tasks for each technique. Participants also assigned a numerical ordering to the four conditions (Table 1) (from 1 to 4, where 1=most preferred and 4=least preferred), based on their overall experience with each – rankings therefore take into account all aspects of the interactions and both task types, and are subsequently interpreted as overall preferences. Rankings and ratings were illuminated with comments captured during an informal post-study interview.

RESULTS

For each measure, results were analysed across conditions, i.e. task types (buttons, slider), using general linear model, repeated-measures ANOVAs, conducted in SPSS, unless otherwise specified. Where appropriate, Bonferroni corrections were applied for multiple comparisons.

Visual Behaviour – Button Selection ('Discrete') Tasks

For *button-selection tasks*, there were no significant differences identified in number of glances (NoG) for either Touch or Haptics. However, there were significantly fewer glances over 2.0-seconds ('long glances' or NoG>2) associated with the touchscreen ($F_{(1,126)}$ = 15.3, $p < .001$). There was also a significant interaction for Touch*Haptics ($F_{(1,126)}$ = 4.22, $p = .042$), indicating that adding haptics to gestures *decreased* NoG>2, whereas the NoG>2 associated with the touchscreen *increased* when haptics were added. A similar trend can be observed for the number of glances overall. Similarly, total off-road glance time (TGT) was significantly lower for Touch compared to Gesture ($F_{(1,126)}$ = 4.93, $p = .028$). Nevertheless, adding haptics reduced TGT for the gesture interface, but extended TGT for the touchscreen. Finally, the mean off-road glance duration (MGD) was significantly shorter for Touch compared to Gesture ($F_{(1,126)}$ = 5.69, $p = .019$). There was also a significant interaction for Touch*Haptics ($F_{(1,126)}$ = 6.97, $p = .009$), indicating that adding haptics to gestures decreased MGD, whereas MGD increased when haptics were added to the touchscreen (Figure 3). Considering the distribution of glances by task (i.e. the percentage of tasks that were achieved with zero, one, or two or more glances – see Table 2) rather than the mean number of glances across each condition, further significant differences are revealed. In particular, a significantly larger proportion of button selection tasks were achieved with no glances when using gestures, and in particular, when gestures were augmented with ultrasound haptics ($F_{(3,141)}$ = 26.35, $p < .001$). Conversely, when using the touchscreen (both with and without haptics), the majority of tasks required one glance and this was significantly more than with gesture-haptics ($F_{(3,141)}$ = 49.30, $p < .001$), although there were more tasks requiring 2 or more glances when using gestures generally ($F_{(3,141)}$ = 36.94, $p < .001$).

Visual Behaviour – Slider Bar ('Continuous') Tasks

For *slider-bar tasks*, there was a significant difference in NoG associated with Touch (F(1,125) = 4.0, *p* = .047), indicating that there were fewer off-road glances made when using gestures. Adding gestures had no significant effect on the number of off-road glances overall, but tended to reduce NoG associated with the gesture interface. There were no significant differences in NoG>2 associated with slider-bar tasks, although there was a trend for fewer long glances when haptics was added. For TGT, there were no significant differences identified for either Touch or Haptics, although adding haptics to the gestures tended to reduce total off-road glance time. Adding haptics also significantly reduced MGD (F(1,125) = 4.40, *p* = .038) for both the touchscreen and gestures, whereas no significant differences in MGD were found associated with Touch for slider-bar tasks (Figure 3). Considering the distribution of glances by task (Table 3), there were significantly more tasks achieved with no glances when using gestures (no touchscreen slider bar tasks were completed without vision), and the most zero-glance tasks were achieved when using gestures-with-haptics (F(3,141) = 14.68, *p* < .001). While a comparable number of tasks between conditions required one glance, significantly fewer tasks required multiple glances (two or more) when using gestures with haptics (F(3,141) = 7.79, *p* < .001).

Table 2. Mean percentage (standard deviation) of button-selection tasks completed with zero, one and two or more glances.

	Zero glances	One glance	Two or more glances
Touch-No Haptics	1.27 (5.76)	84.95 (27.38)	5.44 (6.73)
Touch-With Haptics	2.43 (11.16)	88.19 (22.91)	5.21 (7.72)
Gesture-No Haptics	13.43 (19.54)	56.60 (26.40)	25.81 (18.67)
Gesture-With Haptics	26.39 (26.62)	47.00 (21.47)	24.54 (21.31)

Table 3. Mean (standard deviation) percentage of slider-bar tasks completed with zero, one and two or more glances.

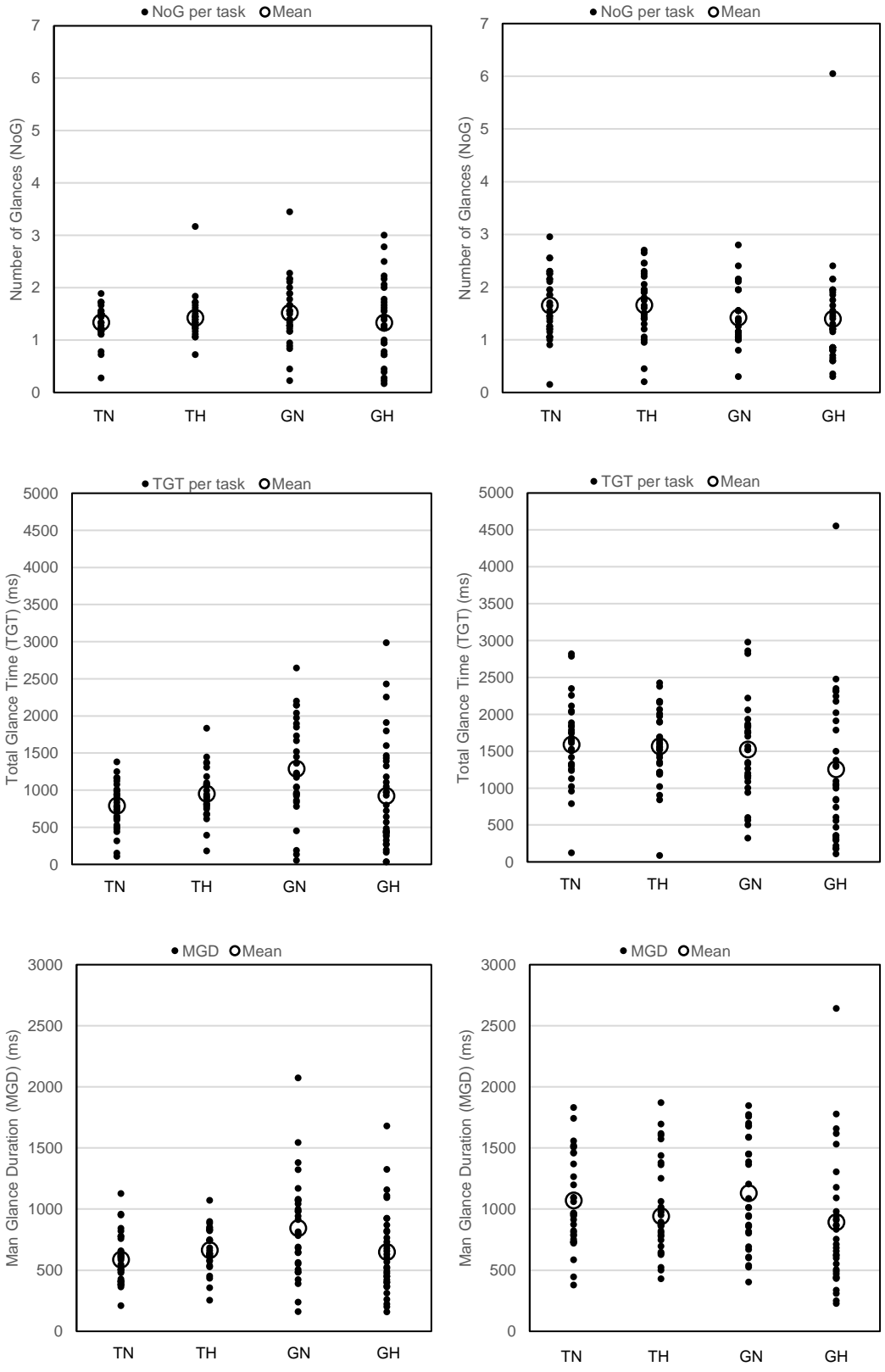|	Zero glances	One glance	Two or more glances
Touch-No Haptics	0	46.35 (30.92)	47.40 (31.03)
Touch-With Haptics	0	49.58 (27.79)	46.25 (27.53)
Gesture-No Haptics	1.46 (3.09)	55.21 (27.54)	39.17 (26.28)
Gesture-With Haptics	10.78 (18.65)	56.15 (24.91)	31.04 (24.39)

Figure 3. Visual performance measures (number of glances (NoG)-top, total glance time (TGT), mean glance duration (MGD)-bottom) for button selection (discrete) tasks (left) and slider bar ('continuous') tasks (right), showing mean values and distribution, where T = touch, G = Gesture, H = with haptics, and N = without (no) haptics
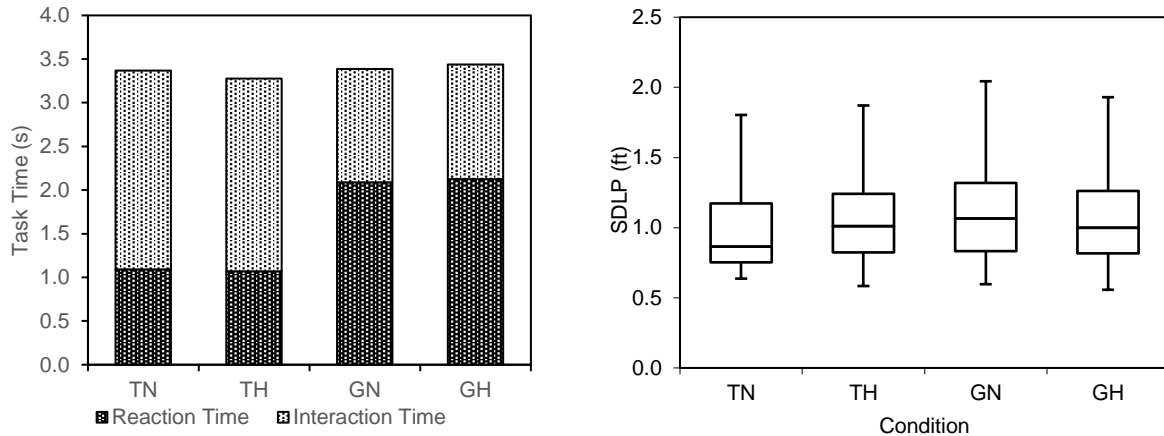
Figure 4. Task performance, showing mean secondary task-time for slider bar tasks split by reaction time and interaction time (left), and standard deviation of lane position (right), where T = touch, G = Gesture, H = with haptics, and N = without (no) haptics

Primary Task (Driving) Performance

There was a significant difference for standard deviation of lane position (SDLP) ($F_{(1,45)}$=407.0, $p <$ .001), with pairwise comparisons indicating that SDLP was lower during Touch-only compared to Gesture-only ($p = $ .039). However, when this was corrected for multiple comparisons, differences were no longer significant (Figure 4-right). There were no significant differences revealed in standard deviation of headway (SDHW) between conditions.

Secondary Task Performance

There was a significant differences between Touch and Gesture for slider-bar 'overshoots' ($F_{(1,46)}$=202.8, $p <$ .001), showing that fewer errors of this type were made when using gestures (both with and without haptics). In addition, when interactions were enhanced with haptics, benefits were more evident during the Touch condition than with Gesture ($F_{(1,46)}$=1848.0, $p <$ .001). There were no significant differences identified for percentage of correct selections for the slider-bar tasks between Touch and Gesture, although again, haptics tended to benefit the Touch condition more than Gesture for this measure. When task-time was broken down into 'reaction time' and 'interaction time' for *slider-bar tasks*, it was evident that the time taken to undertake the task (i.e. the interaction time) was significantly shorter when using gestures ($F_{(1,46)}$=43.15, $p <$ .001), whereas reaction time (i.e. the time to respond to the task instruction) was quicker for Touch compared to Gesture ($F_{(1,46)}$=232.0, $p <$ .001) (Figure 4-left).

Workload – NASA-TLX

NASA-TLX ratings were captured after each drive and therefore relate to each interaction technique, but reflect both types of tasks – button selections and slider-bar tasks. Significant differences were revealed in Total Workload ($F_{(3,141)} = $ 12.652, $p <$ .001), with pairwise comparisons showing that the Total Workload associated with the gesture-haptics condition was significantly lower than all other conditions (all $p <$ .001) (Figure 5). Considering the individual subscales (Figure 6), it is evident that ratings for Total Workload were influenced by significant differences in Performance ($F_{(3,138)} = $ 3.522, $p = $ .017) and Frustration ($F_{(3,138)} = $ 4.861, $p = $ .003). Specifically, participants indicated that they

performed better with Gesture-Haptics compared to Gestures-without-Haptic (p = .017) and Touchscreen-with-Haptics (p = .053). In addition, ratings of frustration associated with the touchscreen (both with and without Haptics) were higher than the Gesture-Haptics interface (p = .035 (TN) and p = .008 (TH)).
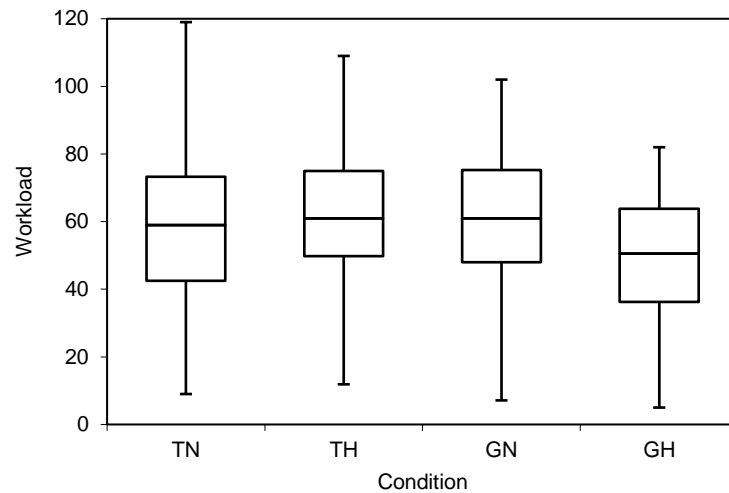

Figure 5. NASA-TLX total workload ratings, where T = touch, G = Gesture, H = with haptics, and N = without (no) haptics
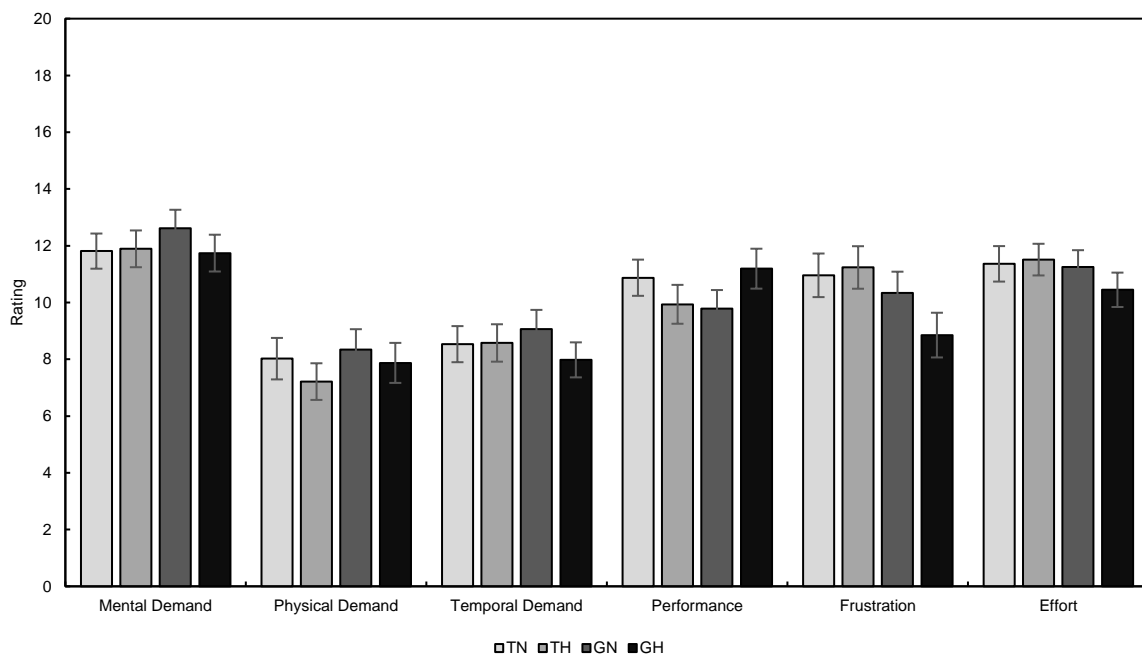

Figure 6. NASA-TLX subscale ratings, with standard error bars, where T = touch, G = Gesture, H = with haptics, and N = without (no) haptics
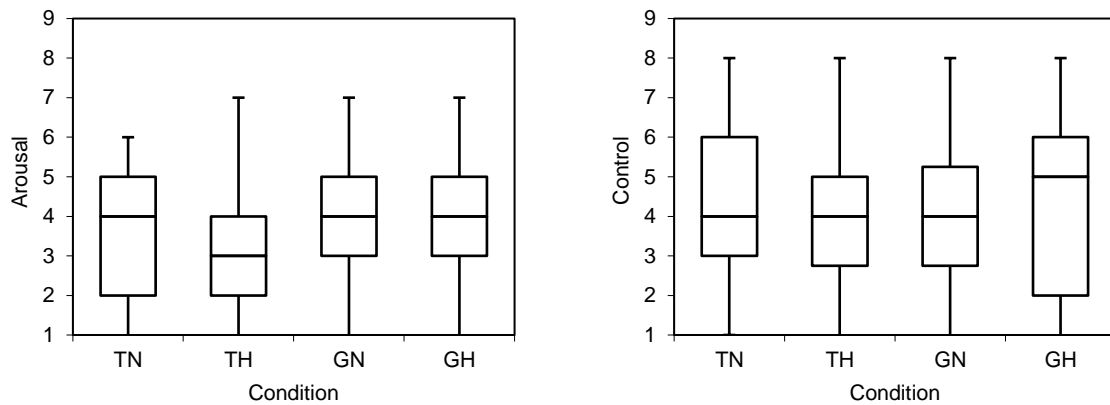
Figure 7. Affective experience, showing ratings of arousal (left) and control (right), where T = touch, G = Gesture, H = with haptics, and N = without (no) haptics

Affective Experience – Self-Assessment Manikin

To measure the affective experience ('user-experience'), participants completed the self-assessment manikin questionnaire (SAM) [32]. There was a significant difference for Arousal between the Touch interfaces (TN, TH) and the Gesture interfaces (GN, GH) ($F(3,138) = 3.165$, $p = .027$) (Figure 7-left). Pairwise comparisons showed that the Gesture interfaces attracted higher (more favourable) ratings compared to Touch. However, when these were corrected for multiple comparisons using Bonferroni corrections, the differences were not significant. There were no significant differences identified between ratings of Control between the four conditions ($F(3,138) = 1.186$, $p = .317$) (Figure 7-right). Nevertheless, there was a trend towards greater control associated with the gesture-haptic interface (based on mean values), although the variability of responses varied significantly.

Preference Ratings and Comments from Participants

Participants were asked to rank the four conditions in order of preference. Pairwise ranking was used to systematically compare each condition with each other. Figure 8 shows pairwise scores and rankings for button-selection tasks and slider-bar tasks, respectively. For button selection tasks, participants tended to prefer the touchscreen, whereas gestures-with-haptics was preferred for slider-bar tasks. These ratings were also supported by comments made by participants. For example, regarding button-selection tasks, a participant commented: "The touchscreen was good because you could see what you pressed." Even so, some participants recognised the limitations associated with the touchscreen: "The touchscreen was difficult to actually get...and you do have to actually look at the number...it meant I had to look at the screen for longer, which was more distracting." Support for gestures-with-haptics was evident for both button-selection and slider-bar tasks: "You don't even need to look, you can kind of feel. So the haptic feedback helps with that"; "It gave me more confidence that I choose the correct number"; "The haptics slider was probably the easiest, because as you moved it gave you a click-click-click feedback." It was also clear from participants' comments that gestures alone (i.e. without haptic feedback) were more challenging to use: "Gesture control without haptics was difficult because you couldn't tell what you were activating."
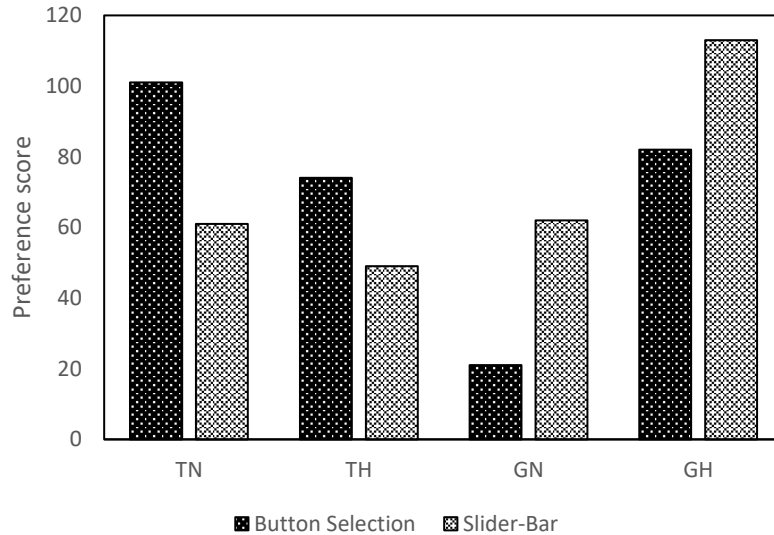
Figure 8. Preference scores (absolute values) for discrete button selection and continuous slider-bar tasks, where T = touch, G = Gesture, H = with haptics, and N = without (no) haptics

DISCUSSION

Given the inherent visual demand associated with using a touchscreen, and the potential impact that this has on primary task performance while driving, a more robust human-vehicle interaction solution that demands less visual attention is required. One proposed solution is to use gestures augmented with focussed ultrasound to create the sensation of touch in mid-air [22]; this overcomes the lack of agency associated with pure gestural interfaces. Utilising a counter-balanced 2×2 experimental design, this study compared the use of gestures with a traditional touch-surface interface in a driving simulator, with each interface evaluated both with and without mid-air ultrasound haptic feedback. Employing measures of primary and secondary task performance, visual behaviour and user experience, the current study aimed to evaluate the use of mid-air ultrasound haptics as a novel human-vehicle interaction paradigm, and in particular evaluated the following hypotheses:

1. Overall, using gestures to execute secondary in-vehicle tasks will require less visual demand, provide a more positive user experience and demand lower workload, compared to a traditional touch-surface interface.
2. The addition of mid-air ultrasound haptic feedback will reduce the visual demand, enhance the user experience and reduce workload associated with both gesture and touch-surface interfaces.
3. The lowest visual demand, most positive user experience and lowest workload will be associated with gestures augmented with mid-air ultrasound haptic feedback.
4. There will be no detriment to primary (driving) and secondary task performance when using gestures augmented with mid-air ultrasound haptic feedback.

It was evident that using gestures with ultrasound haptics enabled a significant percentage of tasks to be completed with no vision – in practice, 26.39% of button-selection tasks and 10.78% of slider bar tasks. In contrast, the majority of button-selection tasks required one glance or more when using the touchscreen, and for slider-bar tasks, using the touchscreen *always* demanded at least one glance.

There was recognition for the potential 'eyes-free' benefit of using gestures and haptics revealed through the post-study interviews, with many participants suggesting that gestures and haptics could enable completely vision-free interactions as their familiarity with, and the usability and integration of, the technology improves; similar aspirations were not associated with the touchscreen. This is particularly illuminating given that participants were not specifically instructed to attempt to complete the tasks without vision, and therefore suggests that mid-air haptics naturally invited and afforded this behaviour and users recognised this potential, whereas the touchscreen did not.

In situations where participants *did* make visual reference to the device, it was possible to reduce the visual demand through the addition of ultrasound haptics. However, this was only true for some tasks. For example, when using gestures for button selections, the addition of mid-air haptics resulted in fewer long glances, and led to a significant reduction in total off-road glance time (TGT) and mean glance duration (MGD). In addition, for slider-bar tasks, the gesture interface attracted the fewest number of glances, and again, the addition of haptics tended to reduce the number of glances, in particular, the number of glances over 2.0-seconds, as well as TGT and MGD. Nevertheless, adding haptics to the touchscreen had the opposite effect. It increased the number of long glances made to the touchscreen, as well as TGT and MGD, although TGT and MGD were shorter and there were fewer long glances (>2.0-seconds) overall when participants used the touchscreen compared to the gesture interface for button selections. There is also some evidence based on the visual performance measures to suggest that the touchscreen alone performed better than gestures for button selections. However, it was not possible to reduce the visual demand further by providing haptics. Instead, providing haptic feedback to the touchscreen actually increased visual demand. In contrast, the provision of mid-air ultrasound haptics to gestures significantly reduced the visual demand associated with this technique, and this was true for all task-types.

The gesture-haptics interface was also associated with the best secondary task performance overall, in terms of the percentage of correct responses and minimising target overshoots. However, there were also notable benefits in terms of accuracy associated with the slider-bar task when the touchscreen was enhanced by haptic feedback. Conversely, utilising gestures for slider-bar tasks appeared to extend total task-time. By splitting task-time into reaction time and interaction time, it was clear that the additional time was associated with the former activity, that is, the time from the delivery of the task instruction to the start of the interaction. Considering the interaction time in isolation (i.e. the time taken to manipulate the slider bar and make selections), there were notable benefits in terms of reduced response time when using gestures-with-haptics.

There were some differences apparent in the primary, driving performance measures (in particular, the standard deviation of lane position), with better lateral vehicle control evident when participants used the touch-surface, compared to gestures *without* haptics. However, when the gestures were haptified with ultrasound, vehicle control was comparable to the touch-surface. This suggests that while gestures on their own have the potential to detriment driving performance, the additional provision of mid-air haptic feedback could negate any deleterious effects on vehicle control. However, these effects were small and therefore further longer-term driving studies are recommended to explore this further.

Subjectively, there were differences identified in overall workload between all four conditions, with the gesture-haptics interface attracting the lowest ratings. This was due to significant differences in ratings for performance and frustration, with both subscales suggesting that the gesture-haptics

interface enabled the best performance, and invited the lowest frustration during use. In addition, results from the self-assessment manikin show a trend towards higher arousal when using the gesture interfaces, compared to touch. While the higher ratings of arousal associated with the gesture interfaces may have been inspired by the novelty of the interaction, it is also interesting to note that there were no differences associated with ratings for control between interfaces, and even some suggestion of improved control based on the subjective ratings. This indicates that participants generally felt that they were able to control the interface equally well when using gestures compared to the more familiar touchscreen technology, with many participants actually indicating better control with the gesture-haptics, as also reflected by the NASA-TLX results. Participants tended to indicate a reduction in perceived control when haptics were added to the touchscreen, whereas adding haptics to gestures tended to increase ratings of control. Nevertheless, ratings of workload and affective experience were captured for each drive and therefore included participants' ratings for both button-selections and slider-bar tasks for each technique. Thus, it is feasible that participants' subjective perception of workload and their affective experience varied between task type as well as interaction technique, potentially confounding overall ratings and increasing the range of responses, and there is some anecdotal evidence from participants' comments post-study to support this. Support for the gestures and haptics was also evident in the preference ratings, with 'gesture with haptics' identified as the most popular for the slider-bar task, and the second most popular for selecting buttons. The fact that the touch-surface was most popular for button selections, and achieved the shortest task times for these, is unsurprising given that touchscreens are now common in many contexts, and the interaction itself (touching the screen) remains perceptively quicker and easier than locating and activating a virtual mid-air button – which was a novel experience to most of our participants. However, it is also worth noting that the Audi TT vehicle used as the simulator buck is very compact and has a characteristically small interior. As such, the touchscreen was located close to the driver in the centre console (see Figure 1-right), and therefore generally within easy reach. This might not be the case in larger vehicles, such as SUVs etc., where touchscreens may be placed outside of easy-reach zones.

Limitations and Future Work

It is worth highlighting that the gestures, associated haptic sensations and experimental interfaces were developed specifically for the study and retrofitted to the simulator vehicle buck. Thus, some of the performance metrics (for example, those associated with the combined touchscreen and ultrasound haptics, which performed poorly for some metrics), may have been influenced by the bespoke implementation, experimental set-up and post-hoc in-vehicle installation. In addition, to enable gesture detection, the transducer array was located in the centre of the vehicle, between the driver and passenger seats. While this might be a reasonably likely location for such a device as it ensures that the effective haptified hand gesture interaction region is comfortably within reach of the driver's arm and hand, it is recognised that locating such a device is a non-trivial issue and should be considered as a focus for future research.

In addition, as part of the training, participants were encouraged to carefully locate their hand above the transducer array before commencing each gesture interaction. This ensured that their hand was detectable by the Leap Motion camera and that it was at the focus of the ultrasound sensations. It also ensured that there was sufficient room to complete slider bar manipulations – not too close to be obstructed by the steering wheel at higher values of the scale, and not too far away to cause difficulties in reach for lower extremities of the scale. However, this self-imposed formality would

likely have required additional time and effort to achieve during the investigation, thereby extending task-time (reaction time), even though participants were trained and deemed to be competent in using the technique. In practice, this would likely reduce as drivers' familiarity with using the technology increases. Moreover, the *(x,y,z)* coordinates for creating the focal point (i.e. at the centre of the driver's palm) were obtained from the LEAP Motion API. It is feasible there was an inherent hardware or software latency between this and the transducer array, meaning that the generation of the acoustic radiation force was delayed. This was in fact evident in situations where haptics were used in conjunction with the touchscreen, meaning that potentially, some selections were made before the haptic feedback was even generated. This is also likely to improve in future implementations as the flexibility, capability and integration of the sensing technology, and design of interaction elements, improves. Finally, participants were required to undertake multiple repetitions of each task for all interaction techniques, and therefore fatigue may have influenced their performance.

A further potential limitation in the approach is that the activities under investigation were selected as task-oriented and not goal-oriented. Therefore, participants were required to make specific selections and manipulations repeatedly using a rudimentary interface with limited real-world functionality or appeal. This was a necessary experimental constraint to increase the confidence of the response data and avoid any confounding effects, but it is recognised that in practice, drivers would likely have a specific, defined goal in mind when interacting with an in-vehicle device, such as increasing music volume. This would not only provide environmental feedback (the music gets louder), but might also not necessitate the precise accuracy demanded during the study, in which participants were required to move a slider-bar by three increments, for example.

Finally, it is worth re-iterating that the touchscreen remained present during all conditions. This was not strictly required when participants were using gestures augmented with haptics, and it is thus feasible that the presence of the touchscreen may have inadvertently attracted superfluous visual attention during this condition. Nevertheless, it was decided that the display should remain present during all conditions in the interests of scientific rigour – it would hardly be surprising if the gesture-haptics condition demanded no visual attention, if no visual stimulus was provided, although it is recognised that a real-world implementation of the system may differ.

Despite these limitations, the study shows practical utility and clear potential for the use of haptified gestures in the automotive domain, and generally supports our hypotheses. This is because the approach overcomes problems associated with the lack of agency, typically experienced with isolated gestural interfaces, and subsequently has the potential to remove the need for vision completely. It is therefore of particular relevance in an automotive context, where a driver's visual attention should ideally remain directed at the primary driving task. Moreover, there is evidently further scope to develop novel gestures and distinct haptic sensations to help drivers differentiate and select different targets. Indeed, future gestures-with-haptics interfaces need not be bound by the traditional restrictions of a visual interface, such as a limited physical space to present a finite number of interaction elements, and this could dramatically increase the scope for novel, multifarious interactions and engaging user experiences.

CONCLUSION

The study evaluated the novel use of ultrasound to emulate discrete mid-air buttons and a graduated slider-bar, activated using gestures, in a driving context. By comparing this with a traditional touch-surface interface, the study shows clear potential and practical utility for the use of ultrasound-haptified gestures in the automotive domain, with reductions in visual demand, shorter interaction times, and improved accuracy for continuous, slider-bar tasks evident when haptic feedback was provided. The combined gesture-ultrasound haptics interface was also very popular amongst participants, who rated it particularly highly for continuous slider-bar manipulations, and indicated better performance and lower levels of frustration associated with its use, compared to the more familiar touch-surface interface. Further work is required to optimise sensations and interactions, for example, to improve the speed of response during discrete button-selection tasks. In addition, future work could consider the optimal location for such a device in a vehicle, as well as seeking to locate the technology within more representative dynamic driving contexts, where vibrations, vehicle movements, and other demands of real-world driving may also impact on usability.

ACKNOWLEDGMENTS

REFERENCES

[1] S. Rümelin and A. Butz, "How to make large touch screens usable while driving," in *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* , 2013.

[2] A. Taveira and S. Choi, "Review study of computer input devices and older users.," *Intl. Journal of Human–Computer Interaction,* vol. 25, no. 5, pp. 455-474, 2009.

[3] D. R. Large, G. Burnett, E. Crundall, G. Lawson and L. Skrypchuk, "Twist It, Touch It, Push It, Swipe It: Evaluating Secondary Input Devices for Use with an Automotive Touchscreen HMI," in *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2016.

[4] National Highway Traffic Safety Administration (NHTSA), "Visual-Manual NHTSA Driver Distraction Guidelines For In-Vehicle Electronic Devices," 2013.

[5] International Organization for Standardization, "Road vehicles -- Transport information and control systems -- Detection-response task (DRT) for assessing attentional effects of cognitive load in driving.," International Organization for Standardization, 2016.

[6] S. Klauer, T. Dingus, V. Neale, J. Sudweeks and D. Ramsey, "The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data," 2006.

[7] B. Ahmad, P. M. Langdon, S. J. Godsill, R. Hardy, L. Skrypchuk and R. Donkor, "Touchscreen usability and input performance in vehicles under different road conditions: an evaluative

study," in *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2015.

[8] B. Ahmad, P. Langdon, S. Godsill, R. Hardy, E. Dias and L. Skrypchuk, "Interactive displays in vehicles: Improving usability with a pointing gesture tracker and Bayesian intent predictors," in *6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2014.

[9] A. L. Eren, G. Burnett and D. R. Large, "Can in-vehicle touch screens be operated with zero visual demand?: An exploratory driving simulator study," in *Proceedings of the 4th International Driver Distraction and Inattention Conference*, Sydney, Australia, 2015.

[10] M. Pitts, G. Burnett, L. Skrypchuk, T. Wellings, A. Attridge and M. Williams, "Visual-haptic feedback interaction in automotive touchscreen use," *Displays,* vol. 33, no. 1, pp. 7-16, 2012.

[11] K. May, T. Gable and B. Walker, "A multimodal air gesture interface for in vehicle menu navigation," in *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2014.

[12] E. Ohn-Bar, C. Tran and M. Trivedi, " Hand gesture-based visual user interface for infotainment," in *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* , 2012.

[13] C. Pickering, K. Burnham and M. Richardson, "A research study of hand gesture recognition technologies and applications for human vehicle interaction," in *3rd Institution of Engineering and Technology Conference on Automotive Electronics*, 2007.

[14] A. van Laack, Measurement of Sensory and Cultural Influences on Haptic Quality Perception of Vehicle Interiors, BoD–Books on Demand, 2014.

[15] A. Laack, O. Kirsch, G. Tuzar and J. Blessing, "Controlling vehicle functions with natural body language," in *Mensch und Computer 2016–Workshopband*, 2016.

[16] S. Rümelin, T. Gabler and J. Bellenbaum, "Clicks are in the Air: How to Support the Interaction with Floating Objects through Ultrasonic Feedback.," in *9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Oldenburg, Germany, 2017.

[17] L. Garber, "Gestural Technology: Moving Interfaces in a New Direction," *Computer,* vol. 46, no. 10, pp. 22-25, 2013.

[18] P. Cornelio Martinez, S. De Pirro, C. Vi and S. Subramanian, "Agency in mid-air interfaces," in *CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, 2017.

[19] M. Weiss, C. Wacharamanotham, S. Voelker and J. Borchers, " FingerFlux: near-surface haptic feedback on tabletops.," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011.

[20] C. Wusheng, W. Tianmiao and H. Lei, "Design of data glove and arm type haptic interface," in *Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2003. HAPTICS 2003.*, 2003.

[21] T. Hoshi, M. Takahashi, T. Iwamoto and H. Shinoda, "Noncontact tactile display based on radiation pressure of airborne ultrasound," *IEEE Transactions on Haptics,* vol. 3, no. 3, pp. 155-165, 2010.

[22] T. Carter, S. Seah, B. Long, B. Drinkwater and S. Subramanian, "UltraHaptics: multi-point mid-air haptic feedback for touch surfaces.," in *Proceedings of the 26th annual ACM symposium on User interface software and technology*, 2013.

[23] L. Gavrilov and E. Tsirulnikov, "Mechanisms of Stimulation Effects of Focused Ultrasound on Neural Structures: Role of Nonlinear Effects.," in *Nonlinear Acoust. at the Beginning of the 21st Cent.*, 2002, p. 445–448.

[24] K. Harrington, D. Large, G. Burnett and O. Georgiou, " Exploring the Use of Mid-Air Ultrasonic Feedback to Enhance Automotive User Interfaces," in *10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, Toronto, Canada, 2018.

[25] Ultrahaptics Ltd., "Ultrahaptics' mid-air haptic technology commercialised in IGT's TRUE 4D Slot Machines," 4 January 2018. [Online]. Available: https://www.ultrahaptics.com/news/announcements/igt-4d/. [Accessed 29 April 2019].

[26] O. Georgiou, V. Biscione, A. Harwood, D. Griffiths, M. Giordano, B. Long and T. Carter, "Haptic In-Vehicle Gesture Controls," in *9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications Adjunct* , Oldenburg, Germany, 2017.

[27] G. Shakeri, J. Williamson and S. Brewster, "May the Force Be with You: Ultrasound Haptic Feedback for Mid-Air Gesture Interaction in Cars.," in *The 10th International ACM Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutoUI2018)*, Toronto, Canada, 2018.

[28] K. Brookhuis, D. Waard and B. Mulder, "Measuring driving performance by car-following in traffic.," *Ergonomics,* vol. 37, no. 3, pp. 427-434, 1994.

[29] International Organization for Standardization, "Road vehicles — Measurement of driver visual behaviour with respect to transport information and control systems — Part 1: Definitions and parameters," ISO, Geneva, Switzerland, 2014.

[30] Society of Automotive Engineers, "Operational Definitions of Driving Performance Measures and Statistics," SAE International, Warrendale, PA, 2015.

[31] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," *Advances in psychology,* vol. 52, pp. 139-183, 1988.

[32] M. Bradley and P. Lang, "Measuring emotion: the self-assessment manikin and the semantic differential," *Journal of behavior therapy and experimental psychiatry,* vol. 25, no. 1, pp. 49-59, 1994.