Whiskers: Exploring the Use of Ultrasonic Haptic Cues on the Face

Hyunjae Gil[1], Hyungki Son[2], Jin Ryong Kim[2], Ian Oakley[1]

[1]Dept. of Human Factors Engineering
UNIST, Ulsan, Korea

[2]Electronics and Telecommunications Research Institute, Daejeon, Korea

hyunjaegil@gmail.com, {hagonhk, jessekim}@etri.re.kr, ian.r.oakley@gmail.com

ABSTRACT
Haptic cues are a valuable feedback mechanism for smart glasses. Prior work has shown how they can support navigation, deliver notifications and cue targets. However, a focus on actuation technologies such as mechanical tactors or fans has restricted the scope of research to a small number of cues presented at fixed locations. To move beyond this limitation, we explore perception of in-air ultrasonic haptic cues on the face. We present two studies examining the fundamental properties of localization, duration and movement perception on three facial sites suitable for use with glasses: the cheek, the center of the forehead, and above the eyebrow. The center of the forehead led to optimal performance with a localization error of 3.77mm and accurate duration (80%) and movement perception (87%). We apply these findings in a study delivering eight different ultrasonic notifications and report mean recognition rates of up to 92.4% (peak: 98.6%). We close with design recommendations for ultrasonic haptic cues on the face.

Author Keywords
Tactile; Haptic; Face; SmartGlasses; Notifications.

ACM Classification Keywords
H.5.2. User Interfaces – Haptic IO.

INTRODUCTION
Augmented Reality (AR) smart-glasses are an emerging consumer technology with the potential to provide rich, interactive graphical contents superimposed over the real world. While the most striking uses of the technology involve impressively high fidelity mixed reality simulations [1], the most fundamental are likely to be more mundane. For example, Google Glass, a headset commonly touted as a privacy violating consumer failure [10], has been reinvented as a tool for business productivity [24] – helping human workers navigate spaces, follow instructions or get in touch with colleagues while on the job. Indeed, researchers have long recognized that the ultimate value of AR will rest less on how well it entertains us and more on how well it supports diverse everyday activities: travelling to new places [20], learning new skills [22] or providing targeted knowledge to solve specific problems [9].

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-5620-6/18/04...$15.00
https://doi.org/10.1145/3173574.3174232

Visual displays are a core technology for conveying content relating to these tasks: they can render dense, rich and precise information. In equal measure, they can disturb and distract by presenting irrelevant, unnecessary or occluding content [7]. Recognizing the importance of these problems, researchers have begun to explore how AR systems could be enhanced to present content without detracting from a user's main task by, for example, targeting peripheral vision [25] or using alterative sensory modalities such as audio [38] or haptic cues delivered to the head [5]. We argue that haptics has particular promise in this area and, indeed, prior authors' activities support this. Operating generally in VR settings, prior work has deployed actuators as diverse as tactors [41], fans [23], flywheels [8] and Peltier elements [29] in Head Mounted Displays (HMDs) to support tasks ranging from navigation or guidance [17] through spatial awareness [28] to increasing immersion or presence [32].

This paper builds on this work by investigating what we term *whiskers*: haptic cues delivered to the face via a commercial ultrasonic actuator [4]. This is valuable because, in contrast to the physical tactors [17] or fan based air jets [23] used in prior work, ultrasonic displays may improve comfort, ease fit and deliver a much more diverse set of cues. Specifically, although they are non-contact, they have a high resolution and range – one actuator can accurately position spatiotemporally varying cues over a large area, potentially rendering a rich variety of information. To investigate the value of this idea, this paper presents three lab studies exploring perception and comprehension of ultrasonic haptic cues on the face. Two studies tackle basic perception while the third is embedded in a more realistic notification scenario in which application specific meanings are assigned to each cue. The results cast light on perceptual properties such as spatial and temporal resolution that we distill into practical recommendations for how to create effective cues in this modality.

We believe that in-air haptics are a viable and valuable feedback modality for smart-glasses. The data and recommendations in this article provide a detailed characterization of performance that future designers and researchers can apply to develop distinguishable and effective feedback systems using ultrasonic in-air haptics. In sum, the contributions of this paper are: 1) the first empirical data describing recognition performance for a range of ultrasonic haptic cues on the face; 2) data from an applied study that integrates these cues into a notification delivery scenario and 3) recommendations for designing ultrasonic haptic cues for display on the face based on these outcomes.

RELATED WORK

The recent commercial success of HMDs has attracted considerable research interest in the idea of presenting tactile cues to the head. A common and elegantly simple approach is simply to integrate standard mechanical vibrotactile actuators into a headband [28] or helmet [17]. Researchers have demonstrated such systems can be used to signify directional or spatial cues, including in complex situations such as 3D VR scenarios [17, 28]. Such systems are also useful for presenting tactons [2], multi-dimensional cues in which additional stimuli parameters are encoded via properties such as frequency or amplitude. Other forms of physical actuators have also been used to enhance HMDs. Temperature displays [29, 32] show promise for increasing the levels of immersion that can be achieved with VR headsets, or even for directional cueing [30]. Researchers have also considered non-tactile force based designs that use flywheels [33] to convey more realistic sensations of impact or directional pulls.

While these approaches show the potential and value of head based tactile feedback, other authors have observed that the requirement for physical contact inherent in these designs is problematic: it complicates wearable fit [23] and limits the number of locations that can be addressed. Accordingly, a variety of authors have explored delivering non-contact tactile stimuli to the head. The dominant approach is to use air-flow delivered directly from fans [23, 33] or via vortexes [35]. Authors have characterized the subjective experience of these types of cue [33, 35] and reported success in conveying notifications by targeting different locations such as the neck or cheek [23]. Airflow haptics is also reported to improve immersion in VR environments – contributing to more accurate simulations of real world conditions such as weather [15]. While this literature shows the potential of in-air haptic cues to the head, we note a discrepancy between the relatively crude actuators that have been studied and the current state of the art in-air haptic actuators that can render a much greater diversity of stimuli.

Arguably the most prominent technology for sophisticated in-air haptics is ultrasonic actuation [4, 12]. Such systems operate by focusing ultrasonic sound waves emitted from a grid of actuators on tightly targeted points in space. They are capable of accurately positioning cues over relatively large spatial regions and with a range of frequencies and amplitudes [4]. While these actuators are currently physically large, we note future devices may shrink considerably if the addressable spatial region is reduced; accordingly, researchers have begun to explore their use in wearable scenarios [18, 34]. Prior work on these devices has characterized perceptual performance on the hand, suggesting cues can be localized with an error of 8.5mm, that directional movement can be readily recognized [40] and that two points are distinguished between 31% (same frequency) and 86% (different frequencies) of the time when they are separated by 30mm [4]. Reflecting the promising potential of this technique, researchers have used it to enhance a wide range of mid-air interactions in applications such as holography [11], non-contact tactile display [13, 14, 31], musical instruments [16], communicating emotions [27] and for increasing agency in input [26].

This paper targets the discrepancy between the current interest in head based haptic stimulation as a feedback channel and the relative simplicity of the actuators used to display cues. We identify a research opportunity to characterize human recognition performance with state-of-the-art ultrasonic in-air haptic cues delivered to the head; we know of no prior studies exploring this context and actuator technology. Furthermore, we argue ultrasonic technology is a particularly good fit for delivering cues to regions of the face around the eyes – areas suitable for an HMD mounted actuator. This paper seeks to address the lack of knowledge on this topic with initial lab studies of ultrasonic haptic cues delivered to the face. The outcomes of these studies characterize basic perceptual and recognition performance and can serve to inform the future design of haptic cues, effective interaction techniques and miniaturized wearable ultrasonic actuators intended for integration into HMDs.

TACTILE PERCEPTION ON THE FACE

The face provides a unique context for displaying tactile sensations. This is partly because of its complex and diverse geometry and abundance of bodily landmarks [6] in the form of facial features. Landmarks, in this sense, represent areas of the skin that are readily identifiable due to their proximity to a site of high saliency: the tip of the nose, for example, can likely be identified more easily and accurately than a similarly sized site situated on the cheek. Another factor that makes the face unique is that it is populated by a reduced set of touch sensing mechanoreceptors. Four standard types of receptor are present: Ruffini corpuscles, Meissner corpuscles, Merkel cell disks, and hair receptors [37]. The face lacks Pacinian corpuscles, the receptors responsible for detecting the relatively high frequency vibrations (~250Hz) commonly used in research on tactile feedback. Such cues are, in practice, undetectable on the face. Accordingly, this work targeted Meissner corpuscles which respond to stroking and fluttering sensations (touch and vibration). They are most sensitive to vibrations between 10 and 50 Hz. Meissner corpuscles are also rapidly adapting receptors [37], meaning that they respond strongly on stimulus onset, but this response decays rapidly over time, ultimately providing little information as to the duration of a static prolonged stimulus.

Beyond this basic physiology, psychophysical literature has also characterized tactual responses on the face. Prior studies have reported spatial resolution in the form of two-point discrimination thresholds derived from point stimuli. Due to variations in cell types and distributions, responses differ: the cheek is 9.0mm to 13.1mm; the lateral forehead 13.4mm to 15.0mm and; the medial forehead is 12.7mm to 13.0mm [37]. However, reports also vary. In recent work exploring air based haptics, esthesiometer derived two-point thresholds were reported to be lower: 5.8mm, 8mm and 12mm for the cheek, medial forehead and anterior forehead [35]. It is worth noting that these figures all substantially exceed thresholds on the finger (2-4mm) but match reported values for the palm (8-12mm). As such, cues delivered to the face will likely not be detected as accurately as those to the finger, but may attain performance similar to other hand regions. We note that while this literature provides a useful guide to the perceptual experience of touch on the face, it generally relies on traditional stimulation techniques, such as pinpricks, and may not directly apply to other forms of stimuli, such as the in-air haptics cues studied in this paper. As such, we argue there is value in determining perceptual performance with ultrasonic in-air haptic cues in order to complement existing knowledge about the tactile sensitivity of the face.

SYSTEM

We delivered in-air tactile cues to the face using ultra-sound [4, 12]. Specifically, we used the UltraHaptics evaluation kit (www.ultrahaptics.com/products/evaluation-kit/), a system that features a 16×16 array of ultrasonic transducers that can render cues across a 60° field of view and at distances of up to 800mm from the top surface of the transducer. Focal points are reported to be 8.6mm in diameter, as measured by a high-performance microphone [4]. We note that cue diameter, as perceived by a user, is likely to be lower as the intensity of stimulation is not uniform – it drops towards the periphery of the region and will also vary according to the physiology of, and mechanoreceptors present in, the target skin region. To deliver cues to the face, we selected a modulation frequency of 40Hz to maximize the response of Meissner corpuscles and a magnitude (in the UltraHaptics API) of 2.0. Subjective testing during development with pilot participants indicated there was a relatively narrow band of viable frequency/amplitude choices and that these settings led to optimal cue perception. They were fixed for all empirical work in this paper. The sensation they evoke on the skin is best described as light pressure or fluttering.

Ultrasonic haptics requires accurate 3D body tracking to deliver cues to the skin: each is delivered to a highly specific location in space. Even relatively small tracking errors can diminish or alter perception of the cues. This is a particularly important issue when considering delivery of cues to the face due to its complex and various geometry: it features non-trivial curvature that differs substantially from individual to individual. To ensure accurate cue delivery, we used an Intel RealSense SR300 depth camera to capture the shape of participant's faces in real time. This device has a 640 by 480 resolution, a frame rate of 60 Hz and is capable of tracking depth at short ranges: the minimum distance is 200mm. This device is capable of generating an accurate model of a user's face in a short period of time.

The lab setup used in all work described in this paper combined these devices (see Figure 1, left). They were mounted on a robust support structure in a vertical arrangement with the haptic display directly above the tracker. The assembly was positioned 250mm in front of a participant's face, an appropriate range for both tracking and actuation. To increase the reliability of the tracking, we also stabilized participant's faces by ensuring they were always comfortably seated with their heads on a chin-rest. The chair and chin rest were adjusted for each participant to align his or her nose with the bottom of haptic display and the top of the tracker. This provided optimal fields of view for both devices. To alleviate the impact of audio cues while using the system, participants always wore in-ear headphones (to deliver instructions) underneath a 3M audio isolation headset. While we know of no evidence suggesting ultrasonic haptic cues are harmful, participants also wore small protective plastic glasses during all studies to physically isolate their eyes from the haptic stimuli.

STUDY 1: LOCALIZATION & DURATION

We conducted an initial study of perception of ultrasonic haptic cues on the face. It involved three variables: three face *sites,* a set of five horizontally aligned equidistant *locations* at each site and cue lengths of three different *durations*. These variables encompass basic properties relating to determining appropriate face sites to target, the spatial resolution of the skin and the sensitivity to shorter or longer cues. We describe and justify our choices for these variables in more detail below.

Sites: The three face sites were: the left *cheek*; the left side of the forehead, directly above the *brow* and; the center of the forehead, directly above the *bridge* of the nose. We subsequently refer to these as the *cheek*, *brow* and *bridge*. They are shown in Figure 2. The face sites were chosen following a process of informal experimentation and to reflect four key constraints. First, they are viable locations for feedback from our motivating scenario of a smart-glasses system: they are regions adjacent to the eyes and thus, the rims and bridge of a pair of glasses. Secondly, they are at 90° to the ears, making it effectively impossible for the ultrasonic cues to directly address the ears, a potential hazard [21]. Thirdly, they are experimentally expedient: faces are most readily tracked when observed frontally and all three regions can be addressed with minimal movements of the head with respect to the haptic actuator in our study setup. Fourth, they are (relatively) flat, further simplifying the process of applying haptic cues. While additional sites, such as the temple or nose, may be worthy of study, we argue that study of these three sites will yield a representative picture of performance appropriate for an initial study.

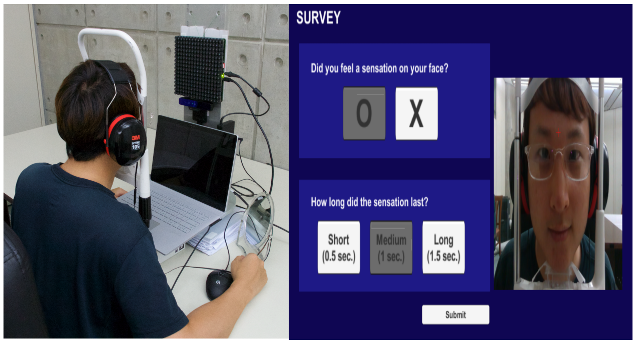

Figure 1. Left shows experimental setup including the UltraHaptics display and SR300 (depth camera) located 250mm in front of participants' faces. It also shows the laptop used enter data and answer questions. Right shows the experimental answer sheet for study 1. If participants answer 'O' to the first question, the second question and captured face figure appear. Cursor keys are used to move the red crosshair to the perceived stimulus location.

Locations: For each body site, we delivered cues to five locations arranged in a 14mm horizontal line; they were spaced at 3.5mm. These choices differ from prior studies of ultrasonic cues to the hand: Wilson *et al.* [40], for example, look at a 2D grid of points spaced at 10mm. Our choices do not reflect a belief that tactile localization performance will be higher on the face than the hand. Rather, they sought to fit within our selected body sites (e.g., the bridge is relatively small) and to span relatively flat face surfaces. The small inter-cue distances and avoidance of vertical target variations simplified the process of meeting these constraints.

Durations: We investigated the impact of cue duration on perception. We selected levels of 500ms, 1000ms, and 1500ms. These values were chosen based on previous work on ultrasonic haptics to the hand that used 100ms and 1000ms cue durations [40]. Using a range of values for this basic display parameter is particularly important due to the fact that cues in this work target the less frequently studied Meissner corpuscles. Durations that are effective may differ from those that are viable in systems for other skin sites.

Methods

Twelve participants (mean age 24, six female) were recruited via online advertisements on university/research institute forums and groups. They were compensated for their participation with ~10 USD. They were screened to exclude individuals with neuropathy or specific injuries to the face with a questionnaire. Such conditions are rate; no prospective participants were excluded.

Participants completed the study in a quite office. In addition to the haptic display, depth camera and chin rest, a laptop computer was placed directly in front of the participants, in easy sight and reach. All data entry was using this device. The study began with a training session to familiarize participants with the experimental equipment and ultrasonic cues. It involved resting in the chin support and exposure to the full range of cues used in the experiment. They also completed up to five randomly selected trials for each face site to gain familiarity with the study procedure and interface, until they were satisfied they understood the process. No data was logged during this period. Including setup and explanations of the study procedures, this session typically took 30 minutes. The main study then started. Participants completed trials from each *site* in a separate block. The order of these blocks was balanced among participants, with two completing each of the six possible orders. Within each *site* block, participants completed two repetitions of each combination of five *locations* and three *durations*. Cues within each block were randomly ordered. In total, 1080 trials were recorded: 12 participants by 3 *sites* by 5 *locations* by 3 *durations* by 2 repetitions. The main part of the study took approximately one hour to complete.

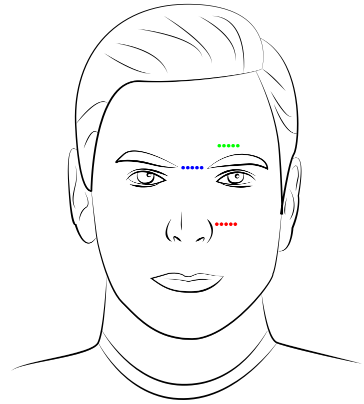

Figure 2. Five points in the localization study for each face site. Cheek is red; bridge is blue; brow is green.

At the start of each *site* block participants were asked to adopt a posture best suited to accurate face tracking. This was parallel to the actuator for the *bridge* condition and yaw rotated by 10° rightwards for *check* and *brow*. Each trial then followed a similar process. First participants clicked the laptop mouse to begin and a calibration process was performed to capture the exact face geometry and position in order to support precise cue delivery. This process involved capturing 50 samples and filtering the depth data with a low pass filter in order to remove noise. If the filtered depth data showed high variability (for example if the participant moved), calibration was restarted. The calibration process typically took three to five seconds. From calibration onwards, participants were required to stay still. One second after calibration was complete, participants received an audio cue (a beep on the headphones), there was a one second pause and the ultrasonic haptic stimuli was presented. After it was complete, participants were allowed to move and, in fact, were asked to mark the location they felt the cue (if any) on their face with the index finger of their non-dominant hand. This was intended to help separate performance in the perceptual task of detecting the cue on the skin from that of the memory task of recalling where they cue was while noting down this data.

Face site	500ms	1000ms	1500ms
Cheek	95.4%	97.8%	99.4%
Middle	95.8%	98.9%	100.0%
Forehead	94.1%	93.3%	96.9%

Table 1. Cue detection rate in study 1.

Duration	500ms (Response)	1000ms (Response)	1500ms (Response)
500ms (actual)	83.4%	15.4%	1.1%
1000ms (actual)	10.2%	76.0%	13.7%
1500ms (actual)	0.3%	21.7%	78.0%

Table 2. Confusion matrix for duration perception rate in study 1.

Participants then logged responses by answering a series of questions on the laptop. They first recorded a yes/no answer as to whether they felt a cue. If they answered negatively, this was recorded and the trial terminated, but participants were required to repeat it again later. If they answered affirmatively, a second question inquired about the duration, with three options: short (0.5 seconds), medium (1 second) and long (1.5 seconds). Finally, a screen capture of their face from the start of the trial was presented with a red cursor marking a neutral central location. They used the arrow keys on the keyboard to move this point to the perceived stimulus location. They were allowed to examine a mirror to help with this process. This point was then translated into real world coordinates using measurements from the depth camera. The response system for this study is shown in Figure 1 (right).

Results

Stimuli detection and duration recognition rates are shown in Tables 1 and 2. Detection rates were highest on the bridge with 1500ms duration (100%) and lowest on the brow with 1000ms cues (93.3%). Accuracy of duration classifications ranged from 83% (500ms) to 76% (1000ms), which the confusion matrix shows is almost entirely due to overlap between adjacent stimuli levels (i.e. between 500ms and 1000ms and between 1000ms and 1500ms). Localization accuracy data is shown in the heat maps in Figure 3. While these visualize the data in 2D, we numerically calculated accuracy for statistical assessment using the horizontal dimension only, as cues did not vary vertically. We calculated the absolute horizontal distance between the intended focal point and perceived point for each trial. The mean localization error was 4.82mm (SD = 2.98mm) for the cheek, 3.77mm (SD = 2.29mm) for the bridge and 9.04mm (SD = 5.43mm) for the brow.

Data from this study were analyzed with three-way repeated measures ANOVAs on site, location and duration for the measures of detection rate, accuracy of duration perception and mean absolute error. Greenhouse-Geisser sphericity corrections were applied if required and pairwise comparisons were conducted with *post-hoc* Bonferroni

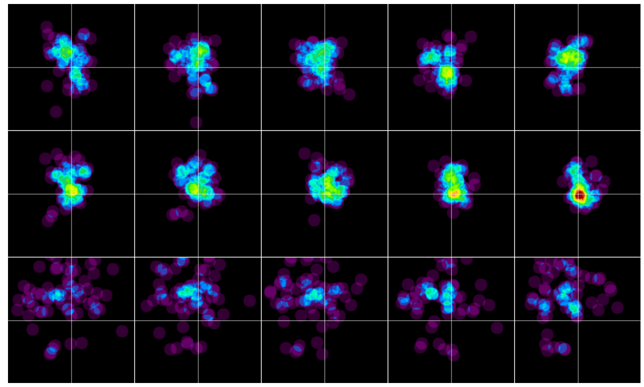

Figure 3. Heat map images illustrating distribution of perceived points. The size of each cell is 3cm x 3cm and white lines represent the target stimuli locations. The top row shows cheek data, the center bridge data and the bottom brow data.

adjusted t-tests. In the interests of brevity only significant results (at $p<0.05$) are reported. Neither detection rate nor duration accuracy showed significant differences with any variable. Data from mean accuracy showed significant differences only for site ($F(2, 22) = 17.22$, $p < 0.001$, $\eta_p^2 = 0.61$), that *post-hoc* testing revealed was due to the brow yielding lower accuracy than both other locations ($p<0.005$). This indicates that cues were perceived generally uniformly across all the stimuli applied in the study with the exception of low localization accuracy, clearly visible in Figure 3's heat maps, on the brow.

Discussion

Recognition performance across the range of parameters studied in this experiment was good (97%) and generally flat – few significant differences emerged. It is worth discussing some differences with respect to prior work. The mean detection rate we report is 97%, lower than Wilson *et al.*'s [40] 99%. This is no doubt in part attributable to greater tactile acuity on the hand, but we also note the role of individual differences in the response patterns of participants. Specifically, two participants in the current study reported a low cue detection rate (mean: 88%), while the remaining set reported 99%. Examining the performance of this pair in assessing other stimuli parameters (above the mean), we suggest these individuals may have ignored experimental instructions and opted to not respond to trials when they felt uncertain about all details. This behavior, while an inherent aspect of this kind of study, may have acted as a confound: pushed down detection rate while pushing up recognition rate for the other stimuli parameters. One of this pair also showed evidence of a strong practice effect, correctly detecting every cue in his or her final condition in the study. This suggests that additional experience or training may level out such variations and improve performance.

The study highlights interesting aspects of haptic cue design for the face. Specifically, we note that the reduced resolution of the brow may be due to a lack of meaningful bodily or facial landmarks [6] that can provide reference points to mark the location of a perceived cue. In contrast, the bridge spans the area between the eyes and one end of the range used on the cheek is proximate to the nose. These landmarks may be instrumental in the better localization performance in these conditions. Figure 4 explores this idea: the heat map shows data from the extremities of the stimuli ranges for each face site. Two distinct points are most clearly observed in the bridge condition, where the stimuli range connects two clear bodily landmarks. We also note that while we found no significant difference in duration perception rates, the confusion matrix clearly suggests that cue durations of 500ms and 1500ms are readily distinguishable; pairs separated by only 500ms are less distinct. These observations form practical guidance that can be used to design effective ultrasonic haptic cues for delivery to the face.

STUDY 2: DYNAMIC CUES

We conducted a follow up study to broaden the scope of the cues in this work from the static presentations in the first study to dynamic presentations of content that stimulates multiple skin locations. This idea relates to Tactile Apparent Motion (TAM), or the creation of what appears to be a smoothly moving tactile stimuli through sequential stimulation of fixed, discrete skin sites. TAM has been demonstrated with both vibrotactile and electro-cutaneous stimuli [19, 36] and has been shown to be affected by a range of parameters including stimulus duration, inter-stimulus distance, inter-stimulus onset interval (ISOI) and skin site [19]. However, as previous work has acknowledged [40], there are challenges to achieving TAM using ultrasonic actuators. Specifically, rendering multiple simultaneous cues (as in ISOI) is difficult – presenting two cues reduces their magnitude, potentially confounding these variations. We also conducted informal tests with short cue durations, such as the 100-300ms typical in TAM studies, that suggested they were infective on the face.

Accordingly, we opted to study a very limited subset of cues variations in TAM. Specifically, we fixed duration to 1000ms, the mid-point of the range successfully tested in study 1, and varied inter-stimulus distance. Although we were unable to vary ISOI, we did vary the simple *interval*, or temporal spacing between the cues. While not a traditional part of TAM, subjective testing indicated different intervals created distinct perceptual experiences that might impact recognition performance. We also opted to vary face sites and the horizontal direction of cue motion.

Ultimately, the study examined four variables: three facial *sites* (*cheek*, *brow* and *bridge*, as in study 1); the horizontal *direction* of movement (left/right); the horizontal inter cue *distance* between successive stimuli (2mm, 4mm, 6mm) and; the inter-cue *interval* between the consecutive cue presentations (100ms, 300ms). While this setup leaves many aspects of TAM unexamined, we believe it represents a practical set of parameters that is well suited to this article's actuator setup and core focus on haptic information delivery to the face.

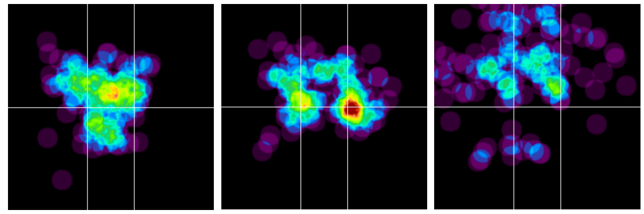

Figure 4. Heat maps illustrating data from the extremities of the range used in study 1 (14mm separation). White lines show the two stimuli locations. Left figure shows cheek data, center shows bridge data and right shows brow data.

Methods

Twelve new participants (mean age 23, six female) were recruited for this study using identical recruitment methods and screening procedures to study 1. They were exposed to a similar training session. They were compensated with approximately 10 USD and the entire study took approximately 90 minutes to complete.

As with study 1, participants completed all trials from each site in a separate block, and the order of these blocks was fully balanced among participants. Once again, trials in each block were randomly ordered. Each involved four repetitions of each level of all other variables. This led to a total of 1728 trials: 12 participants x 3 sites x 3 distances x 2 directions x 2 intervals x 4 repetitions. Trials followed the same trial initialization and calibration process used in study 1. Each trial involved a pair of stimuli, each of which was displayed for one second. The first stimuli in each trial was to the middle point of the range used for each site in study 1. This was followed by a short *interval* (100ms/300ms) before delivery of a second cue at a specific distance (2, 4, 6mm) in a specific direction (left/right). After stimuli were presented, participants answered three questions on the laptop: 1) whether they perceived a sensation; 2) if the second stimulus moved to left or right and; 3) if the second stimulus was perceived as near to or far from the first one.

Results

In this study, detection rates (not charted) were uniformly high (97.5%) across all stimuli parameters. Direction perception and distance perception results are shown in Figures 5 and 6 for all four study variables: site, direction, distance and interval. Distance perception scores were calculated by assigning a score of 50% to a stimulus marked as near and 100% to one marked as far.

All study data was analyzed using similar statistical techniques and procedures to the first study. In terms of direction perception, there were significant main effects of site ($F(2, 22) = 8.45$, $p = 0.002$, $\eta_p^2 = 0.44$) and distance ($F(2, 22) = 21.71$, $p < 0.001$, $\eta_p^2 = 0.66$). Post-hoc testing indicated that the cheek exhibited inferior direction perception to the bridge (p=0.08, 62% vs 79%) and that the large 6mm distances led to greater accuracy than both shorter

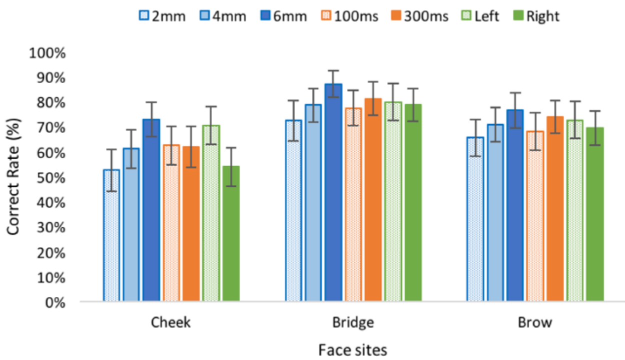

Figure 5. Direction perception rate for the distance, inter-cue intervals, and direction on the cheek, bridge and brow in study 2. Bars show standard error.

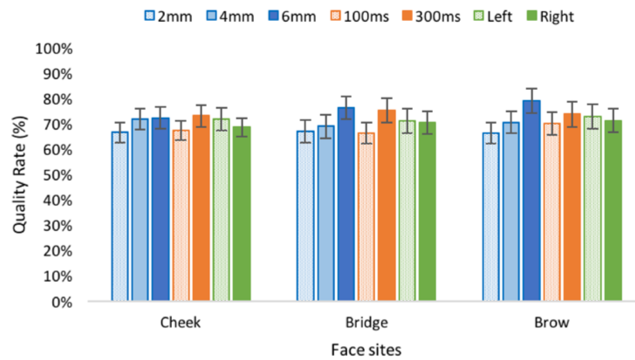

Figure 6. Distance perception for the distance, inter-cue intervals, and direction on the cheek, bridge and brow in study 2. Bars show standard error.

options (both $p<=0.001$, 79% vs 70% and 64%). For distance perception, main effects of distance ($F(2, 22) = 19.63$, $p < 0.001$, $\eta_p^2 = 0.64$) and time interval ($F(1, 1) = 41.59$, $p < 0.001$, $\eta_p^2 = 0.79$) were strong and there was also a comparatively weak interaction between interval and direction ($F(1, 11) = 5.35$, $p = 0.04$, $\eta_p^2 = 0.33$). Post-hoc testing on the distance variable indicated that the 6mm distance led to a greater proportion of "far" ratings ($p<=0.002$, 76% compared to 66% and 70%). The result for the interval variable indicates that the longer 300ms pause also led to a greater proportion of "far" ratings (74% vs 68%), while the interaction suggests the strength of this effect may be modulated by cue direction: it is more pronounced in cues moving to the left.

Discussion

This study reinforced the viability of the bridge as an appropriate site for the delivery of ultrasonic haptic stimuli: it performed better than the cheek in terms of direction perception. Possible explanations for the reduced performance in the cheek could be that, in comparison to the other two sites, it may exhibit more curvature, potentially hampering cue delivery. Furthermore, the mapping of left and right movements to a part of the body that is clearly on the left may be confusing. There is some data to support this idea. Although the interaction was not significant, rightward movements on the cheek were detected at rates close to chance (54%), while leftward ones considerably beat guessing (70%). Further studies could cast light on this issue. The results also indicate that movement perception increased substantially with the larger distances studied. Indeed, distances greater than the 6mm maximum used in this study would be preferable in a realistic system as they could be expected to further boost performance. Longer 300ms time intervals also modestly boosted perception of movement. A possible explanation for this is that the longer temporal break between the cues led to more distinct sensations, which were more salient to the rapidly adapting Meissner corpuscles we targeted [37].

STUDY 3: NOTIFICATIONS

Building on the results of the first two studies, we conducted a study to investigate how people can recognize notifications shown via ultrasonic haptics cues. The aim of this study is to extend the more abstract work in the first two studies to an application focused scenario where participants are required to associate cues with domain specific meanings. This approach follows the substantial literature on tacton design [2, 3] and reflects the idea that purely perceptual performance is an inadequate stand-in for performance in a real-world scenario. Users of face-based ultrasonic haptic notifications will need to not only reliably perceive displayed cues, but also rapidly and reliably map them to application level concepts such as directions, or specific warnings or other meanings. This lab study sought to assess performance in this more applied task.

When designing this study, we opted to explore only two face sites, *cheek* and *bridge*, due to better performance in measures such as absolute position error and stimuli detection rate in study 1. We were also interested in further studying direction perception on the cheek (revealed in the second study to be poor) to better understand this aspect of performance. We designed a set of eight notifications to deliver to these two sites, encoded by three binary parameters. To provide context to the task, we selected driving and navigational notifications. Cues to signify these messages were based on the outcomes of the first two studies. They were:

Type: Notifications could relate to either a navigational turn *instruction* or a parking proximity *warning*. Turn instructions were shown by *moving* cues; warning instructions were shown by *stationary* cues. Both types of cue were represented by four sequential ultrasonic stimuli. For stationary warnings, each cue was targeted at the same location. For moving directions, each cue was separated by 4mm in the same horizontal direction for a total of 12mm of linear movement. While recognition of movement for the 4mm distance was somewhat challenging (cheek: 62%, bridge: 79%), we opted for this configuration to explore if repeat aggregate presentations improved performance. Intervals between cue presentations was set to 300ms.

Direction: Notifications could relate to content to the *left* or the *right* (e.g. a left turn or a right proximity warning). Directions were simply presented as either the left or right extremities of the 14mm range used during the first two studies. These values exceed three standard deviations from the absolute error for the cheek (4.83 + 3*2.98 = 13.77mm), suggesting they will result in accurate performance. For stationary cues, this design was simple. For dynamic cues, it involves an inherent conflict between the location of the start point (left/right) and the direction of the subsequent movement (right/left). Based on subjective experimentation, we opted to encode direction via movement: a left message was signified by a cue that began on the right and moved to the left. A right cue used the opposite arrangement.

Urgency: Notifications came in two levels of importance: *near* (high urgency, e.g., an immediate turn) and *far* (low urgency, e.g., a distant object). Cue *duration* was used to represent this parameter. We selected the more rapid 500ms for the high urgency *near* messages and the more prolonged 1500ms for the low urgency *far* notifications. Study 1 suggests participants will be able to readily distinguish between these durations and we selected these values to explore whether they impact other aspects of performance, such as movement perception.

Cheek			Turning (Respose)				Warning (Response)			
			Left		Right		Left		Right	
			500ms	1500ms	500ms	1500ms	500ms	1500ms	500ms	1500ms
Turning (Actual)	Left	500ms	90.3%	0.0%	0.0%	0.0%	1.4%	0.0%	8.3%	0.0%
		1500ms	0.0%	84.7%	0.0%	0.0%	0.0%	1.4%	0.0%	13.9%
	Right	500ms	4.2%	0.0%	52.8%	1.4%	31.9%	0.0%	8.3%	1.4%
		1500ms	0.0%	2.8%	0.0%	68.1%	1.4%	23.6%	0.0%	4.2%
Warning (Actual)	Left	500ms	8.3%	1.4%	5.6%	0.0%	81.9%	2.8%	0.0%	0.0%
		1500ms	0.0%	4.2%	0.0%	6.9%	1.4%	86.1%	0.0%	1.4%
	Right	500ms	8.3%	0.0%	1.4%	0.0%	1.4%	0.0%	88.9%	0.0%
		1500ms	0.0%	12.5%	0.0%	1.4%	0.0%	1.4%	2.8%	81.9%

Bridge			Turning (Respose)				Warning (Response)			
			Left		Right		Left		Right	
			500ms	1500ms	500ms	1500ms	500ms	1500ms	500ms	1500ms
Turning (Actual)	Left	500ms	93.1%	0.0%	0.0%	0.0%	1.4%	0.0%	5.6%	0.0%
		1500ms	0.0%	94.4%	0.0%	0.0%	0.0%	0.0%	0.0%	5.6%
	Right	500ms	0.0%	0.0%	80.6%	0.0%	18.1%	1.4%	0.0%	0.0%
		1500ms	0.0%	0.0%	0.0%	97.2%	0.0%	2.8%	0.0%	0.0%
Warning (Actual)	Left	500ms	0.0%	0.0%	4.2%	0.0%	94.4%	0.0%	1.4%	0.0%
		1500ms	0.0%	1.4%	1.4%	4.2%	0.0%	91.7%	0.0%	1.4%
	Right	500ms	6.9%	0.0%	0.0%	0.0%	0.0%	0.0%	88.9%	4.2%
		1500ms	0.0%	0.0%	0.0%	0.0%	0.0%	1.4%	0.0%	98.6%

Table 3. Confusion matrix for notification recognition performance on the cheek and bridge in study 3.

Comparison	Outcome
Site	$F(1, 11) = 23.41, p = 0.001, \eta_p^2 = 0.68$
Type	$F(1, 11) = 4.85, p = 0.05, \eta_p^2 = 0.31$
Direction	$F(1, 11) = 5.82, p = 0.034, \eta_p^2 = 0.37$
Urgency	$F(1, 11) = 7.47, p = 0.019, \eta_p^2 = 0.4$
Site x Direction	$F(1, 11) = 5.01, p = 0.045, \eta_p^2 = 0.32$
Type x Direction	$F(1, 11) = 17.1, p = 0.002, \eta_p^2 = 0.61$
Site x Type x Direction	$F(1, 11) = 8.44, p < 0.014, \eta_p^2 = 0.43$

Table 4. Results from a four-way repeated measures ANOVA on the recognition rate of notifications in study 3 showing three interactions and four main effects.

Methods

Twelve new participants (mean age 22.5, five female) completed this study. Recruitment, screening, compensation and setup procedures were identical to the first two studies.

Before starting the main experiment, participants completed a 20-minute training session. It started with an experimenter explaining the notification scenario and how the different notifications mapped to the stimuli dimensions. To check if this information was correctly comprehended, participants were then asked to explain the meaning of four randomly selected notifications; all participants were able to do this. This session then exposed participants to all stimuli used in the study, and their meanings, on both face sites. A total 16 trials were conducted in the practice session (representing all cues in the study). If requested by participants, specific cues were presented twice during training. Immediately after the training was completed, participants began the main study. Once again face site was treated as an independent variable and presented in a fully balanced repeated measures design. For each site, participants completed 48 randomly ordered trials in which each notification was presented six times. The entire study took approximately 90 minutes per participant and logged a total of 1152 trials: 12 participants by 2 sites by 8 notifications by 6 repetitions.

Each trial followed a similar process to previous studies: participants clicked to start, stayed stationary during calibration and heard a beep to mark the start of a trial. In line with prior notification studies [3], cues were repeated. Specifically, each was presented twice in succession and delimited by a one second pause and another beep. After both exposures, participants answered a series of questions on the laptop: if they felt a cue; what type of information (warning/instruction) it represented; its direction (left/right) and; its importance (near/far).

Results and Discussion

Participants reported detecting a mean of 93.7% of the presented cues. The confusion matrices in Table 3 show the cue recognition data. It spans a range of performance: a grand mean of 85.9% and peak of 98.6% for a low urgency right warning message delivered to the bridge site – a 1500ms stationary cue on the right of the bridge. In contrast, the lowest performance, 52.8%, was observed for an urgent right turn message on the cheek – a 500ms cue that moves from left to right. On average, notifications were recognized correctly on the bridge 92.4% of the time and on the cheek 79.3%. Stationary cues (warnings) achieved a recognition rate of 91.5% and moving cues (directions) made 83.7%. Left cues were recognized correctly 92.5% of the time and right cues 85.2%. Finally, short cues (near, high urgency) were recognized with 98.4% accuracy and long cues (far, low urgency) achieved 99.1%.

We analyzed these results in terms of final cue recognition accuracy in line with the approaches and methods employed in the first two studies. A four-way repeated measures ANOVA revealed three interactions and four main effects, all shown in Table 4. This analysis indicates that the main effects summarized in the previous paragraph all relate to significant variations in performance: long (1500ms) moving cues on the bridge site are most readily recognized. We also note that the interaction effects all involve the direction variable. These can be interpreted as the superiority of left over right only impacting the cheek site and only for moving cues. This suggests that movement cues on the cheek need be aligned with its specific egocentric location: the left cheek is suitable for left moving sensations, but right moving sensations, especially in this study's scenario of indicating a real-world direction, make little sense to users. The right cheek, we speculate, would exhibit opposite behavior. We believe this issue was not present on the bridge due to its location at the front center of the face – from an egocentric perspective, both left and right directions are equally valid. Future work needs further explore this issue.

These results reinforce the viability of the bridge as a delivery site for ultrasonic haptic cues – the site variable showed the strongest effect in the study. Performance considerably exceeded that of the cheek and we also note its front central location avoided confusion when mapping movement based direction cues. We provide additional perspectives on this data in the following section that integrates outcomes from all three studies described in this article.

GENERAL DISCUSSION

It is important to contextualize the *whisker* cues studied in this work within the wider literature; multidimensional tactile notifications have been extensively studied. Brown and Brewster [3], for example, explored the design space of vibrotactile tactons via parameters including rhythm, roughness, frequency, intensity and spatial location, reporting recognition performance of 47.8% over their whole set of 27 cues, 80.1% for a subset of 18 and 96.7% for just the rhythm parameter. The results of this study suggest that *whiskers* match up reasonably well to this performance: in the final study, the eight tactons on the bridge were recognized with 92.4% accuracy. This suggests that ultrasonic haptic feedback to the face is an effective modality for haptic information display, providing performance in recognition tasks is broadly comparable to that achieved with conventional vibrotactile cues.

It is also worth contrasting the work in this paper with prior studies of ultrasonic cues, such as Wilson *et al.*'s [40] characterization of cue perception on the hand or Carter *et al.*'s [4] study of properties such as the two-point threshold. Compared to this literature, the current study differs substantially: both in the body site and in the type of mechanoreceptors it targets. Perhaps unsurprisingly, basic performance also differs. Wilson *et al.* [40] report detection rates of 98.9%, superior to the 93.7% reported in the final study in this paper. This indicates the cues we presented were closer to perceptual thresholds than Wilson's. While the hand's higher tactile acuity is a likely explanation for this difference, we also attribute this, in part, to issues with the calibration process required prior to each trial in our studies. The cues we delivered were highly sensitive to variations in the distance to the face and even small movements of participants during a trial could disrupt perception. Improving the face tracking system would likely boost performance.

We note that other aspects of performance may be superior on the face than the hand. Carter's two-point threshold data suggest 2cm separation for discrimination of two points with 50% accuracy while Wilson reports a 2D localization error of 8.5mm and recommends separating points by between 1.5 and 2cm (mean plus 1SD, a calculation that implies some overlap between locations). These figures are broadly coherent with each other. In our first study, we report 1D localization accuracy for whiskers to be as low as 3.77mm (on the bridge) and the notification study shows high recognition rates (up to 94.8%) for a pair of cues spaced 14mm apart. This strong performance may be due to the lower frequency stimuli we applied and the lack of large-receptive field Pacinian corpuscles in the face. We also suggest that the prevalence of bodily landmarks on the face may also aid localization performance.

Whiskers ultrasonic cues can also be contrasted to prior work that delivers in-air cues via technologies such as fans. The extensive study of this approach by Lee *et al.* [23], which also considers a wearable scenario and, in some experiments, sites on the face is most relevant. The focus of this work is also primarily on deriving optimal stimulus parameters in terms of properties such as cue intensity, duration and body site. While it convincingly illustrates that airflow-based haptic displays can yield good performance in recognition tasks (>80% in most cases), we note that information is generally conveyed by simply applying cues to separate body sites with different actuators – for example, four actuators targeting 90 degree offset locations around the wrist or neck each trigger a binary message. This results in systems that occupy large body sites and require multiple actuators. The ultrasonic approach in this paper differs in its use of a single high fidelity actuator capable of providing a range of cues to small body sites – we argue this may ultimately be a more practical approach.

Finally, we also note whiskers maintain the key advantage advocated by prior designers of head based haptic systems: a natural mapping between actuated points and real world orientations make cues ideal for various spatial awareness [28] and directional [17] tasks. Ultrasonic actuators also offer the potential to stimulate multiple face sites and, within each site, our studies show movement perception performance was effective (up to 87% recognition) even over the relatively small distances of 6mm used in the second study. With the larger multi-cue movements in the notification study, movement perception performance rose to 91.3% on the bridge. An interesting qualifier to these comments is the poor performance with rightward movements on the left cheek in the final study – stimuli designs need be coherent with egocentric bodily perspectives to be effective.

LIMITATIONS

This work suffers from a number of limitations, ranging from the fundamental, through the specific, to the practical. For the former, we note that our (typical) young adult sample (N=36, mean age 23 over all studies) may be poorly representative of general human tactile performance. The number of Meissner corpuscles in the skin decreases with age [39]; testing cues on older adults is an important next step for this work. Further limitations relate to the specific study setup. Although the depth camera accurately measures face shape, even small variations due to participant movements could disrupt cue delivery; a faster calibration free tracking system would likely yield improved performance. Ultrasonic haptic cue strength also varies with distance from the center of the actuator array; this may have contributed to lower performance in the cheek, the most distant site. In the future, more rapid and accurate mechanisms need to be used to target cues on the face and we also need ensure cue strengths are fully controlled.

In terms of practical issues, this work is motivated by a wearable scenario, but relies on actuators that are both large and fixed; consequently, we required that participants' heads were also held still. We note that prior work has shown that wearable ultrasonic haptic systems can be developed for the hands [34] and we highlight that creating fully wearable ultrasonic haptic actuator designs for the face is a clear next step for this work – we hope that the data in this paper, particularly about the viability of cues delivered to small face regions, can help guide future efforts to achieve this. Candidate designs include taking advantage of the fact that current HMDs (e.g., MS HoloLens and Samsung Gear VR) protrude from the face by 6-7cm and thus provide a viable platform for mounting actuators and distance sensors. We speculate that parabolic actuator arrangements on top of HMDs will enable systems capable of focusing perceivable cues on (or just above) the bridge site studied in this paper with relatively limited numbers of actuators (e.g., 1-3 rows of 8-16 actuators). We also note that some of the findings from this work may be applicable to non-wearable scenarios, such as driving or during computer workstation use [13], where the face is constrained in a fairly stable position – a large number of more powerful actuators could be embedded in the dash of a vehicle or the frame of a monitor and arranged (e.g., in parabolas) to address the face. Finally, we also note the current studies report lab data and a next step for this work would be to explore cue perception in more realistic real world settings involving distractions in terms of additional environmental cues (e.g., wind), cognitive activity (e.g., due to performance of actual navigational or driving tasks) and other forms of noise and interference common in mobile and wearable settings. Only by conducting such studies can we determine the true viability and usefulness of the ultrasonic haptic whiskers proposed in this paper.

RECOMMENDATIONS AND DESIGNS

This work examined the feasibility of whiskers: ultrasonic haptic stimuli on the face. In this section, we distill this work into practical recommendations for cue design.

Spatial resolution: The error distance for localization varied across face sites: 4.82mm (SD = 2.98mm) for the cheek, 3.77mm (SD = 2.29mm) for the bridge and 9.04mm (SD = 5.43mm) for the brow. Based on a threshold of 2SDs, we suggest spacing cues by 11mm on the cheek, 9mm on the bridge, and 20mm on the brow.

Duration: In study 1, the extreme durations were most accurately classified. We recommend separating stimuli durations into two levels separated by one second.

Interval: Longer time intervals between successive cues enhanced the perception of changes in position. Cues should be separated by intervals of 300ms.

Distance: Each face site led to different distance perception rates. The larger 6mm distances we studied conveyed changes in position as follows: 73% distance perception for the cheek, 77% for the brow, and 87% for the bridge. Smaller 4mm distances may be viable on the bridge (79%), but larger distances (8-12mm) may be more suitable for other sites.

Movement: While movements can be created by taking account of interval and distance recommendations, they must also match egocentric expectations. Central body sites (e.g., the bridge) can support left and right, while others (e.g., the cheek) may not. Effective designs may need to reflect egocentric bodily perspectives.

Site: When selecting face sites to target, consider proximity to landmarks in order to boost performance. The bridge site studied in this paper performed optimally in part due to its position between the brows: two clear landmarks.

CONCLUSION

In summary, this paper proposed whiskers – high fidelity ultrasonic haptic cues to the face. It explores their feasibility across a wide range of cue parameters and in a practical notification task. The data demonstrates the fundamental viability of ultrasonic haptic cues on the face. This can both inform designers about how to best create such cues and also provide developers with practical targets for next generation wearable hardware that can produce them. In this way, this paper hopes to facilitate the integration of high fidelity ultrasonic non-contact haptic displays into next generation HMDs.

ACKNOWLEDGMENTS

This research was partly supported by 'The Cross-Ministry Giga KOREA Project' grant funded by the Korea government (MSIT) (No.GK17C0100, Development of Interactive and Realistic Massive Giga Content Technology) and by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2017R1D1A1B03031364).

REFERENCES

1. Mark Billinghurst, Adrian Clark, and Gun Lee. 2015. A Survey of Augmented Reality. *Found. Trends Hum.-Comput. Interact.* 8, 2-3 (March 2015), 73-272. http://dx.doi.org/10.1561/1100000049
2. Stephen Brewster and Lorna M. Brown. 2004. Tactons: structured tactile messages for non-visual information display. In *Proceedings of the fifth conference on Australasian user interface - Volume 28* (AUIC '04), A. Cockburn (Ed.), Vol. 28. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 15-23.
3. Lorna M. Brown, Stephen A. Brewster, and Helen C. Purchase. 2006. Multidimensional tactons for non-visual information presentation in mobile devices. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (MobileHCI '06). ACM, New York, NY, USA, 231-238. http://dx.doi.org/10.1145/1152215.1152265
4. Tom Carter, Sue Ann Seah, Benjamin Long, Bruce Drinkwater, and Sriram Subramanian. 2013. UltraHaptics: multi-point mid-air haptic feedback for touch surfaces. In *Proceedings of the 26th annual ACM symposium on User interface software and technology* (UIST '13). ACM, New York, NY, USA, 505-514. http://dx.doi.org/10.1145/2501988.2502018
5. Alvaro Cassinelli, Carson Reynolds, and Masatoshi Ishikawa. 2006. Augmenting spatial awareness with haptic radar. In *10th IEEE International Symposium on Wearable Computers*. IEEE, 61-64. http://doi.org/10.1109/ISWC.2006.286344
6. Roger W. Cholewiak and Amy A. Collins. 2003. Vibrotactile localization on the arm: Effects of place, space, and age. *Attention, Perception, & Psychophysics*, *65*(7), 1058-1077. https://doi.org/10.3758/BF03194834
7. Joseph L. Gabbard, Gregory M. Fitch, and Hyungil Kim. Behind the Glass: Driver challenges and opportunities for AR automotive applications. *Proceedings of the IEEE 102*(2), 124-136. https://doi.org/10.1109/JPROC.2013.2294642
8. Jan Gugenheimer, Dennis Wolf, Eythor R. Eiriksson, Pattie Maes, and Enrico Rukzio. 2016. GyroVR: Simulating Inertia in Virtual Reality using Head Worn Flywheels. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (UIST '16). ACM, New York, NY, USA, 227-232. https://doi.org/10.1145/2984511.2984535
9. Steven Henderson and Steven Feiner. 2011. Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (October 2011), 1355-1368. http://dx.doi.org/10.1109/TVCG.2010.245
10. Jason Hong. 2013. Considering privacy issues in the context of Google glass. *Commun. ACM* 56, 11 (November 2013), 10-11. https://doi.org/10.1145/2524713.2524717
11. Takayuki Hoshi, Masafumi Takahashi, Kei Nakatsuma, and Hiroyuki Shinoda. 2009. Touchable holography. In *ACM SIGGRAPH 2009 Emerging Technologies* (SIGGRAPH '09). ACM, New York, NY, USA, Article 23, 1 page. http://dx.doi.org/10.1145/1597956.1597979
12. Takayuki Hoshi, Masafumi Takahashi, Takayuki Iwamoto, and Hiroyuki Shinoda. 2010. Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. *IEEE Trans. Haptics* 3, 3 (July 2010), 155-165. http://dx.doi.org/10.1109/TOH.2010.4
13. Takayuki Hoshi. 2011. Development of aerial-input and aerial-tactile-feedback system. In *IEEE World Haptics Conference (WHC),* IEEE, 569-573. https://doi.org/10.1109/WHC.2011.5945548
14. Takayuki Hoshi. 2012. Handwriting transmission system using noncontact tactile display. In *IEEE Haptics Symposium (HAPTICS),* IEEE, 399-401. https://doi.org/10.1109/HAPTIC.2012.6183821
15. Felix Hülsman, Nikita Mattar, Julia Fröhlich, and Ipke Wachsmuth. 2014. Simulating wind and warmth in virtual reality: conception, realization and evaluation for a cave environment. *JVRB-Journal of Virtual Reality and Broadcasting*, *11*(10). http://dx.doi.org/10.20385/1860-2037/11.2014.10
16. Inwook Hwang, Hyungki Son, and Jin Ryong Kim. 2017. AirPiano: Enhancing music playing experience in virtual reality with mid-air haptic feedback. In *IEEE World Haptics Conference (WHC),* IEEE*,* 213-218. https://doi.org/10.1109/WHC.2017.7989903
17. Oliver Beren Kaul and Michael Rohs. 2017. HapticHead: A Spherical Vibrotactile Grid around the Head for 3D Guidance in Virtual and Augmented Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 3729-3740. https://doi.org/10.1145/3025453.3025684
18. Cédric Kervegant, Félix Raymond, Delphine Graeff, and Julien Castet. 2017. Touch hologram in mid-air. In *ACM SIGGRAPH 2017 Emerging Technologies* (SIGGRAPH '17). ACM, New York, NY, USA, Article 23, 2 pages. https://doi.org/10.1145/3084822.3084824
19. Jacob H. Kirman. 1974. Tactile apparent movement: The effects of interstimulus onset interval and stimulus duration. *Perception & Psychophysics*, *15*(1), 1-6. https://doi.org/10.3758/BF03205819
20. Panos E. Kourouthanassis, Costas Boletsis, and George Lekakos. 2015. Demystifying the design of mobile

augmented reality applications. *Multimedia Tools and Applications*, *74*(3), 1045-1066. https://doi.org/10.1007/s11042-013-1710-7

21. B. W. Lawton. 2001. Damage to human hearing by airborne sound of very high frequency or ultrasonic frequency or ultrasonic frequency. *Health & Safety Executive*, 77pp.
22. Kangdon Lee. 2012. Augmented reality in education and training. *TechTrends*, *56*(2), 13-21. https://doi.org/10.1007/s11528-012-0559-3
23. Jaeyeon Lee and Geehyuk Lee. 2016. Designing a Non-contact Wearable Tactile Display Using Airflows. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (UIST '16). ACM, New York, NY, USA, 183-194. https://doi.org/10.1145/2984511.2984583
24. Stephen Levy. 2017. Google Glass 2.0 is a Startling Second Act. In Wired Backchannel. Retrieved September, 2017 from https://www.wired.com/story/google-glass-2-is-here/.
25. Kris Luyten, Donald Degraen, Gustavo Rovelo Ruiz, Sven Coppers, and Davy Vanacken. 2016. Hidden in Plain Sight: an Exploration of a Visual Language for Near-Eye Out-of-Focus Displays in the Peripheral View. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16). ACM, New York, NY, USA, 487-497. https://doi.org/10.1145/2858036.2858339
26. Patricia Ivette Cornelio Martinez, Silvana De Pirro, Chi Thanh Vi, and Sriram Subramanian. 2017. Agency in Mid-air Interfaces. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 2426-2439. https://doi.org/10.1145/3025453.3025457
27. Marianna Obrist, Sriram Subramanian, Elia Gatti, Benjamin Long, and Thomas Carter. 2015. Emotions Mediated Through Mid-Air Haptics. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15). ACM, New York, NY, USA, 2053-2062. https://doi.org/10.1145/2702123.2702361
28. Victor Adriel de Jesus Oliveira, Luca Brayda, and Luciana Nedel. 2017. Designing a vibrotactile head-mounted display for spatial awareness in 3d spaces. *IEEE Transactions on Visualization and Computer Graphics*, *23*(4), 1409-1417. https://doi.org/10.1109/TVCG.2017.2657238
29. Roshan Lalintha Peiris, Wei Peng, Zikun Chen, Liwei Chan, and Kouta Minamizawa. 2017. ThermoVR: Exploring Integrated Thermal Haptic Feedback with Head Mounted Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 5452-5456. DOI: https://doi.org/10.1145/3025453.3025824
30. Roshan Lalintha Peiris, Wei Peng, Zikun Chen, and Kouta Minamizawa. 2017. Exploration of cuing methods for localization of spatial cues using thermal haptic feedback on the forehead. In *IEEE World Haptics Conference (WHC),* IEEE, 400-405. https://doi.org/10.1109/WHC.2017.7989935
31. Yasuaki Monnai, Keisuke Hasegawa, Masahiro Fujiwara, Kazuma Yoshino, Seki Inoue, and Hiroyuki Shinoda. 2014. HaptoMime: mid-air haptic interaction with a floating virtual screen. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (UIST '14). ACM, New York, NY, USA, 663-667. https://doi.org/10.1145/2642918.2647407
32. Nimesha Ranasinghe, Pravar Jain, Shienny Karwita, David Tolley, and Ellen Yi-Luen Do. 2017. Ambiotherm: Enhancing Sense of Presence in Virtual Reality by Simulating Real-World Environmental Conditions. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 1731-1742. https://doi.org/10.1145/3025453.3025723
33. Michael Rietzler, Katrin Plaumann, Taras Kränzle, Marcel Erath, Alexander Stahl, and Enrico Rukzio. 2017. VaiR: Simulating 3D Airflows in Virtual Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 5669-5677. https://doi.org/10.1145/3025453.3026009
34. Antti Sand, Ismo Rakkolainen, Poika Isokoski, Jari Kangas, Roope Raisamo, and Karri Palovuori. 2015. Head-mounted display with mid-air tactile feedback. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology* (VRST '15), Stephen N. Spencer (Ed.). ACM, New York, NY, USA, 51-58. https://doi.org/10.1145/2821592.2821593
35. Yuka Sato and Ryoko Ueoka. 2017. Investigating Haptic Perception of and Physiological Responses to Air Vortex Rings on a User's Cheek. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 3083-3094. https://doi.org/10.1145/3025453.3025501
36. Carl E. Sherrick and Ronald Rogers. 1966. Apparent haptic movement. *Perception & Psychophysics*, *1*(3), 175-180. https://doi.org/10.3758/BF03215780
37. Maria Z. Siemionow, Bahar Bassiri Gharb, and Antonio Rampazzo. 2011. The face as a sensory organ. In *The Know-How of Face Transplantation*. Springer London, 11-23. https://doi.org/10.1007/978-0-85729-253-7
38. Yolanda Vazquez-Alvarez, Ian Oakley, and Stephen A. Brewster. 2012. Auditory display design for exploration in mobile audio-augmented

reality. *Personal and Ubiquitous computing*, *16*(8), 987-999. http://dx.doi.org/10.1007/s00779-011-0459-0

39. Ronald T. Verrillo. 1980. Age related changes in the sensitivity to vibration. *Journal of gerontology*, *35*(2), 185-193. https://doi.org/10.1093/geronj/35.2.185

40. Graham Wilson, Thomas Carter, Sriram Subramanian, and Stephen A. Brewster. 2014. Perception of ultrasonic haptic feedback on the hand: localisation and apparent motion. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14). ACM, New York, NY, USA, 1133-1142. http://dx.doi.org/10.1145/2556288.2557033

41. Flynn Wolf and Ravi Kuber. 2018. Developing a Head-Mounted Tactile Prototype to Support Situational Awareness. *International Journal of Human-Computer Studies*, *109*, 54-67. https://doi.org/10.1016/j.ijhcs.2017.08.002