A BONE CONDUCTION BASED SPATIAL AUDITORY DISPLAY AS PART OF A WEARABLE HYBRID INTERFACE

Amit Barde, Gun Lee

HIT Lab NZ,
University of Canterbury,
Private Bag 4800,
Christchurch 8140, New Zealand.
**amit.barde@pg.canterbury.ac.nz,
gun.lee@canterbury.ac.nz**

Matt Ward, William S. Helton

Department of Psychology,
University of Canterbury,
Private Bag 4800,
Christchurch 8140, New Zealand.
**matt.ward@ pg.canterbury.ac.nz,
deak.helton@canterbury.ac.nz**

Mark Billinghurst

School of ITMS,
University of South Australia,
Mawson Lakes, SA 5095, Australia.
mark.billinghurst@unisa.edu.au

ABSTRACT

Attention redirection trials were carried out using a wearable interface incorporating auditory and visual cues. Visual cues were delivered via the screen on the Recon Jet – a wearable computer resembling a pair of glasses – while auditory cues were delivered over a bone conduction headset. Cueing conditions included the delivery of individual cues, both auditory and visual, and in combination with each other. Results indicate that the use of an auditory cue drastically decreases target acquisition times. This is true especially for targets that fall outside the visual field of view. While auditory cues showed no difference when paired with any of the visual cueing conditions for targets within the field of view of the user, for those outside the field of view a significant improvement in performance was observed. The static visual cue paired with the binaurally spatialised, dynamic auditory cue appeared to provide the best performance in comparison to any other cueing conditions. In the absence of a visual cue, the binaurally spatialised, dynamic auditory cue performed the best.

1. INTRODUCTION

One of the most common ways of interacting with mobile devices is through visual interfaces. Similarly, widely available wearable computers such as Google Glass [1] and the Recon Jet [2] also use visual interfaces as the primary means for information delivery. However, visual presentation can overwhelm a user due to an inordinate number of data streams vying for the same screen space, or the users' inability to divide attention between the presented information streams on the screen and the world around them. Either way, the information presented only via the visual faculty can overload the users' senses and cognitive ability [3].

Addressing this problem for wearable displays, with their severely limited screen space and constant demand on the user's attention, is particularly important. Attention critical tasks such as driving, search and rescue etc. maybe adversely

affected by the use of such devices [4] [5] [6]. This presents us with a set of unique challenges; (1) information presentation without overloading the user and (2) unobtrusive information delivery requiring minimum attention from a user perspective.

As part of our research, we are interested in exploring the use of binaurally spatialised auditory cues delivered over a bone conduction headset (BCH) for wearable interfaces. By doing this, information can be presented using auditory and visual cues and hopefully reduce the problem of information overload. In the remainder of this paper we first review related work and then describe an experiment exploring the effectiveness of using several combinations of auditory and visual cues for information presentation in a wearable computer interface. We then summarise results from the experiment and conclude, providing suggested directions for future research.

2. BACKGROUND

Wearable spatial auditory displays have been the subject of research for well over two decades now [7] [8] [9] [10] [3]. Almost all of these systems incorporate the use of Head Related Transfer Functions (HRTF) [11], either individualised or non-individualised, to deliver a three dimensional synthetic rendering of the acoustic environment. However, the drawback with these systems is the use of headphones to deliver the required auditory cues or create an auditory environment. This can isolate a user from his/her acoustic environment [12] [13]. One solution to this problem is the implementation of techniques that allow the capture and reproduction of the ambient acoustic environment. Harma et al. [14] and Tikander et al. [15] have demonstrated the use of such an 'audio augmented reality' device. Devices such as the Here Active Listening system [16] developed by Doppler Labs are gaining popularity in the consumer market.

All these systems need to be inserted in to the ear canal, and rely on external microphone inputs to reproduce the ambient environment around the user. Their sizable form factor also occludes the pinna, meaning signals reaching the

microphone are not necessarily those that have been filtered as a result of the reflections resulting from the shape of user's pinna. From exiting literature we know that the occlusion of the pinna can cause a greater number of front – back confusions and a diminished ability to localise in the vertical plane [17].

The problems described above are unacceptable in attention critical environments that demand high levels of awareness of auditory and visual cues. Leaving the ears open to the ambient acoustic environment while engaging in a visually demanding task is safer than having the ears plugged. To overcome this problem, we use a bone conduction headset (BCH) to deliver binaurally spatialised audio as part of a wearable interface. Previous studies exploring the use of the BCH as an auditory display device appear promising [18] [19] [20] [21] [22]. While the use of the BCH has been primarily restricted to its implementation as an auditory display for the visually challenged [18] [20], some studies have demonstrated its effectiveness even for sighted users [23] [24]. However, besides [23] [24] we are unaware of any studies that incorporate the use of BCH as part of a wearable interface. We hope to demonstrate the practical utility of a BCH, as part of wearable interface, to deliver binaurally spatialised cues.

In the following sections we describe our implementation of the BCH as a spatial auditory display device as part of a wearable hybrid interface. We then present a user study conducted to evaluate the use of audio-visual cues in a visual search task. The ability to reorient attention with the use of these cues is explored. Existing studies demonstrate that there is a significant increase in visual search task performance when auditory cues are used [25] [26] [27]. The amalgamation of the visual and auditory faculties and the ability to exploit their mechanisms of perception could make for a more efficient interface than one that relies completely on a single modality. Such an interface assumes great importance in attention critical fields such as driving and search and rescue. Being able to receive task specific information without having to divide attention between the primary task and information retrieval that may directly affect the outcome of the task is important. For example, receiving binaurally spatialised auditory beacon based navigation during a driving task is safer than having to constantly direct one's visual attention to a GPS device. While the navigation information is critical to the primary task of driving and directly affects its outcome, information delivery can take place in a manner that does not affect the primary task in manner that renders it unsafe. In our study we demonstrate the use of this and other auditory cues along with visual presentations made on wearable device.

3. PARTICIPANTS

30 participants (20 male, 10 female) between the ages of 18 and 34 (Mean: 24, Std. Dev: 4.3) volunteered to take part in the study. Participants reported normal hearing in both ears. No testing was carried out to verify their claims of having normal hearing since there appears to be no known relation between localisation performance and audiogram results unless the hearing loss is profound [25]. All participants were compensated with $20 shopping vouchers for their efforts.

4. METHOD

4.1. Apparatus

The apparatus used for the study can be divided in to three categories; 'real-world' analogue (for far domain stimuli presentation) and tracking equipment, handheld and worn tracked devices used by the participants and a bone conduction headset (BCH) used to deliver auditory cues.

The real-world analogue used for this experiment was a set of three screens connected to each other at 60°. The screens measured 2400mm x 1830mm (74.65° x 59.53° visual angle) and were mounted 600 mm above the floor. Images were projected on to these screens by three NEC LT265 projectors. The tracking system used for the study comprised of four ARTTRACK2 cameras mounted on top of the screens (see figure 1), paired with the DTrack software [28]. The cameras are capable of tracking objects up to a distance of 4.5m. For detailed specifications of the camera see [29].

Equipment used by participants consisted of a Recon Jet [2], a wearable 'smart glass' and the Steradian S-7X laser tag gun [30] (see figure 2). Both these devices had markers affixed to them to allow their positions to be tracked during the course of the experiment. The laser tag gun was modified such that depressing the trigger on the gun allowed the participant to 'shoot' targets that were displayed on the screens. This was achieved by connecting a pair of leads attached to the trigger inside the gun to the circuit board of a mouse. The Recon Jet has a widescreen 16:9 WQVGA display with images on it set to appear as they would on 30 inch HD display at 7 feet. For more detailed technical specification see [2]. The Recon Jet was connected wirelessly through a router. Tracking data was transferred to the PC using the VRPN software [31] [32].
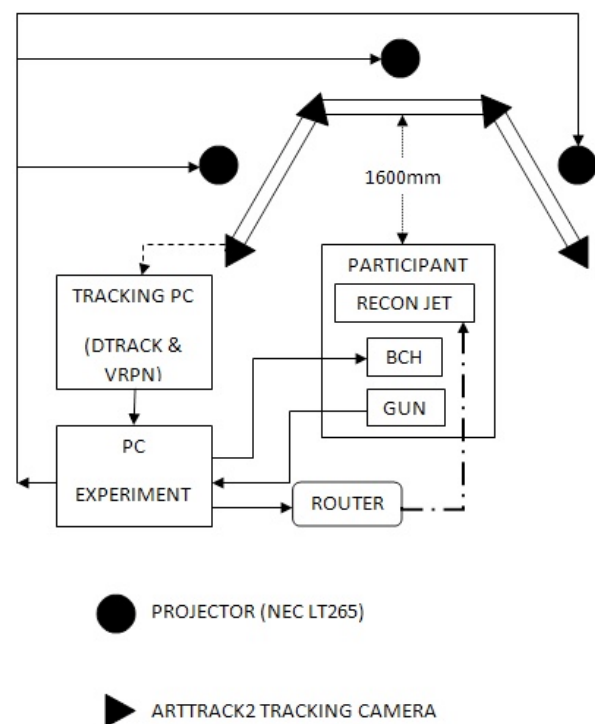

Figure 1: Block diagram of the experimental setup

Participants also wore a pair of bone conduction headsets (Aftershokz Sportz3) [33]. Auditory stimuli for the experiment were reproduced over the BCH. The auditory

stimuli were delivered to the BCH via the PC's on-board sound card.

Figure 2: Participant holding the laser tag gun with markers affixed on top to allow the position of the gun to be tracked. Also seen in the picture are the Recon Jet with markers for tracking, and the BCH.

4.2. Stimuli

Participants were presented auditory and visual stimuli for the experiment. A detailed explanation of the stimuli is given in the following sections.

4.2.1. Visual Stimuli

Visual stimuli were delivered on the screens representing the far field and the Recon Jet display. Stimuli displayed on the screens were targets that appeared at random intervals during the experiment, and a string of numbers at the bottom of the centre screen. Targets consisted of yellow discs of approximately 58mm radius that turned blue when shot and appeared at predefined positions of ±50° and ±100° (see figure 3). The targets were positioned at the centre of the screens. The numerical string used a black Arial typeface of 65mm size positioned in the horizontal centre and approximately 690mm below the vertical midpoint of the centre screen.
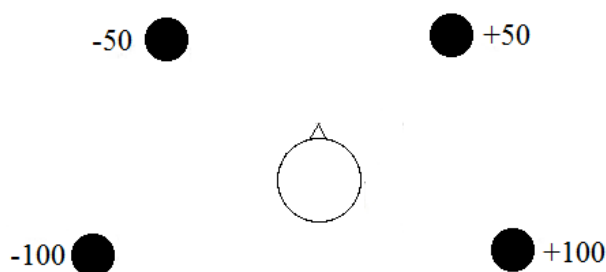

Figure 3: Target Positions

Text displayed on the Recon Jet used a white Arial type face and was positioned in the centre of the screen (see figure 4). Preceding messages were listed above in grey. Visual interrupt signals delivered on the Recon Jet consisted of static cues, pursuit visual cues and a blank screen. The static visual cues consisted of white arrows 1.3° in width and 6.5° in length (see figure 5). The arrows were angled at 40° for targets appearing at ±50° and 80° for targets at ±100°. Pursuit

visual stimuli caused all objects on the screen to move in the direction of the target at 16.2° per second.
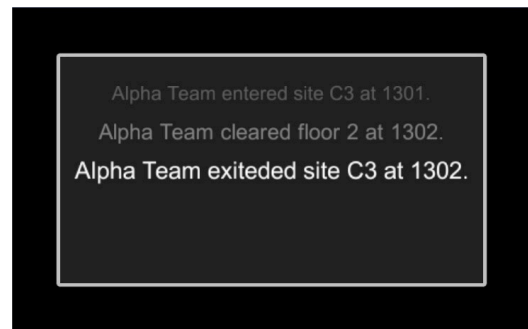

Figure 4: Messages displayed on the Recon Jet screen

In addition to the messages on the Recon Jet and targets projected on the screens, the participants also saw two smaller 'dots' on the screens. These dots represented the position of their head (yellow dot) and the position of the gun (blue dot). Both these moved around on the screen as the participant rotated along the horizontal arc on which the targets appeared.
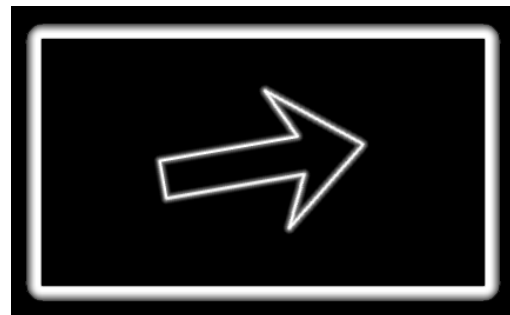

Figure 5: Visual Cue – White Arrow

4.2.2. Auditory Stimuli

Auditory stimuli consisted of a 1 second alarm sound or ping (25ms on set and offset rate). The same sound was used for two of the three types of auditory cues that were delivered. The alarm tone was presented either as a static sound or a binaurally spatialised dynamic audio cue moving in the direction of the target. The binaurally spatialised, dynamic cue simulated the motion of the alarm from the participant's position towards the target on the screen. The cue was designed in accordance with alarm design guidelines prescribed by Walker and Kramer in [34]. The duration and level of the auditory cue were chosen to represent those used by previous researchers [18] [35] [20] [36] [37]. Despite studies demonstrating that wideband noise is easier to localise [18] [20] [36] [38] [39] than most other forms of stimuli, we chose to use a ping for its aesthetic appeal [40]. The third auditory cue consisted only of silence. The static auditory cue was delivered at approximately 70dBA. The dynamic cue on initiation will have had approximately the same loudness level, but fell quickly as the cue moved towards the target. A logarithmic fall off with the addition of the Doppler Effect was modeled to replicate real world auditory percepts.

The visual and auditory cues used here are analogous in that they encompass similar perceptual characteristics, but in different domains (see table 1).

AUDITORY CUES		VISUAL CUES	
S0	No sound (Silence)	V0	Blank Screen
S1	Static alarm	V1	Static (Arrows pointing in the direction of the target)
S2	Binaurally spatialised, dynamic alarm	V2	Pursuit visual cue

Table 1: A list of the auditory and visual cues used in the experiment. A total of 9 cues encompassing a combination of all the cues above were presented to the participants.

The experiment was designed and built within the Unity3D [41] environment. Binaural spatialisation of the auditory cue was achieved using the 3Deception Binaural Engine plug-in for Unity developed by Two Big Ears [42]. We've chosen to adopt the use of a plug-in versus the traditional approach of using individualised HRTFs or HRTF libraries since we believe this lends a greater degree of ecological validity to the study. The plug-in was chosen after an extensive phase of testing and comparisons with existing binaural engines.

5. PROCEDURE

Participants were seated on a rotating chair 1600mm from the central screen (see figure 1). The position was situated approximately on the normal from the central screen such that targets could be presented anywhere on a 200° horizontal arc. Participants were allowed to adjust the height of the chair for comfort.

Participants were then handed the Recon Jet and BCH to put on. If required, they were helped positioning the screen of the Recon Jet so that the text displayed on the screen appeared clear. The BCH was put on such that the drivers of the headset sat in front of the ears on the mandibular condyle. Participants were also handed the laser tag gun. Following this, a short calibration process was run. This was to ensure that participants had a full range of motion that allowed them to reach targets at ±100°, check if tracking information was being gathered in the right manner and ensure that participants were able to read text appearing on the Recon Jet and the main screen. Following the calibration process, six practice trials were conducted. These trials allowed participants to see the different audio – visual cues that could be presented to them during the course of the experiment.

The main experiment was separated in three blocks. Each block was followed by a five minute break. During each block participants were instructed to read aloud number strings appearing on the central screen and messages appearing on the Recon Jet. Messages displayed on the Recon Jet were chosen at random from one of three structure types; [Alpha] team entered site [C][37] at time [1407], [Alpha] cleared floor [2] at time [1407], or [Alpha] team exited site [C][37] at time [1407] (see figure 5).

During a third of the trials for which messages were displayed on the Recon Jet, a cue would interrupt the participant one second after the message's onset. Simultaneously, a target would appear at ±50° or ±100°. Participants were required to 'shoot' or mark the target using the modified laser tag gun as quickly as possible (see figure 6). Once the target had been shot, participants returned their gaze to the central screen, and the alternating display of number stings on the central screen and messages on the Recon Jet resumed. Within each block there was one target event trial for each combination of visual and audio cues for each position for a total of 36 target events (9 event types x 4 target locations) and 72 non-event messages. The experiment took on average 65 minutes to complete.
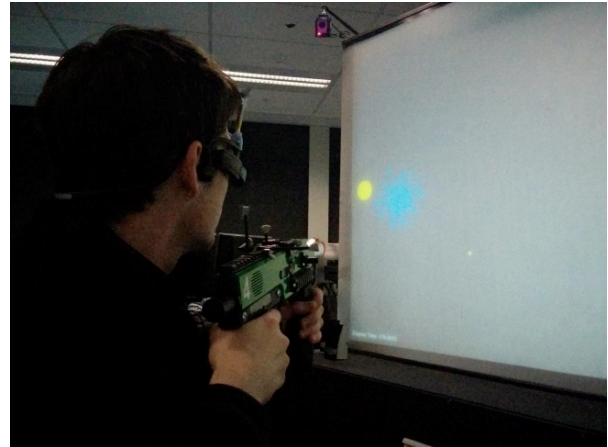

Figure 6: Participant attempting to shoot a target.

6. RESULTS

Two participants were excluded from the analysis – one for not following instructions, while the other had to be excluded due to failure of the on-board sound card to deliver audio signals. Additional technical difficulties with the Recon Jet, primarily associated with power management, meant data from the third block of trials for an additional six participants was recorded incompletely or lost. Data gathered from the third block was excluded from the analysis for all participants to maintain uniformity.

A 3x3x4 (visual cues: 3 auditory cues: 3 target positions: 4) repeated measures analysis of variance (ANOVA) was carried out to test for the main and interaction effects between the factors. Since the data violated Mauchly's test of sphericity, values as determined by the Greenhouse-Geisser correction were used. The results showed significant main effects of all three independent variables; audio cues ($F_{(1.677, 80.513)} = 104.671$, $p < 0.001$), visual cues ($F_{(1.767, 84.822)} = 60.736$, $p < 0.001$) and target positions ($F_{(2.244, 107.732)} = 54.592$). The results also showed significant two way interactions between all pairs (visual x audio: $F_{(3.038, 145.835)} = 16.041$, $p < 0.001$; visual x position: $F_{(3.432, 164.742)} = 13.754$, $p < 0.001$; audio x position: $F_{(3.928, 188.535)} = 8.248$, $p < 0.001$). In addition to this, a significant three way interaction between the three independent variables was also observed ($F_{(5.528, 265.30)} = 4.054$, $p = 0.01$).

To look into the details of the interaction effects, we tested simple two-way interaction effects by fixing the levels of the target position. For the target at -100°, statistically significant interactions were found between the auditory and visual cues ($F_{(2.942, 158.852)} = 11.414$, $p < 0.001$). Main effects of the visual ($F_{(1.806, 97.504)} = 44.114$, $p < 0.001$) and auditory cues ($F_{(1.784, 96.32)} = 47.764$), $p < 0.001$) also showed statistically significant results. Similar results are observed for the target at +100° with interaction effects between the two types of cues being statistically significant ($F_{(3.258, 169.395)} = 5.852$, $p = 0.001$). The main effects of the cues also show statistically significant results (Audio: F

(1.562, 81.214) = 20.299, p < 0.001; Visual: F (1.927, 100.179) = 30.847, p < 0.001).

No significant interaction between the audio and visual cues were seen for targets at -50° (F (2.209, 119.266) = 1.839, p = 0.159) and +50° (F (3.1, 161.204) = 0.971, p = 0.410). For -50° there were significant main effects for both the auditory (F (1.558, 84.145) = 37.285, p < 0.001) and visual cues (F (1.802, 97.331) = 5.982, p = 0.005), while +50° displayed similar effects only for the auditory cues (F (1.785, 92.81) = 30.743, p < 0.001) and not the visual cues (F (1.717, 89.272) = 1.859, p = 0.167). The lack of a significant effect for the visual cues suggests that the peripheral vision over rides any of the visual cues when targets appear in these regions. Participants appear to lock onto these targets as result of the auditory cues and peripheral vision, rendering the visual cues ineffective.

The preceding analysis was then followed up with an analysis of variance (ANOVA) for each target position to compare performance between the different cues and their combinations. All positions displayed a significant difference in performance between the different cue types and their combinations (-100°: F (4.494, 242.66) = 31.008, p < 0.001; +100°: F (3.656, 190.112) = 17.182, p < 0.001; -50°: F (3.851, 207.944) = 12.063, p < 0.001; +50°: F (5.018, 260.93) = 9.672, p < 0.001). Post hoc tests using the Bonferroni correction of pair wise comparisons between the cueing conditions give a detailed picture of the effectiveness of the cues for each of the four positions. For the auditory cueing conditions (V0S0, V0S1 and V0S2) only, the static (V0S1) and dynamic auditory (V0S2) cues outperform the no cue condition, V0S0, at all target positions (-100°: p < 0.001; +100° (V0S2): p < 0.001; -50° (V0S1): p = 0.006; -50° (V0S1): p < 0.001; +50 (V0S1): p < 0.001; +50 (V0S2): p = 0.04) except +100° V0S1 (p = 1). No significant difference was observed between the static (V0S1) and dynamic (V0S2) cueing conditions at -100° (p = 0.194), -50° (p = 1) and +50° (p = 1). A significant difference, though, was observed at +100° (p = 0.008). While further investigation is required, the binaurally spatialised auditory cue consistently demonstrates a faster onset of head motion time across all targets (see figure 7).
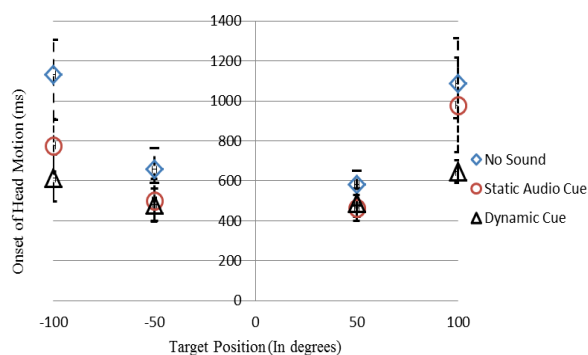

Figure 7: Average time for onset of head motion measured across all auditory cueing conditions.

For the cueing conditions using V1 paired with the auditory cues (V1S0, V1S1 and V1S2), a significant difference is seen between V1S0 and V1S2 at all positions (-100°: p = 0.02; -50°: p < 0.001; +50°: p < 0.001; +100°: p = 0.001). Significant differences were also seen between V1S0 and V1S1 at +100° (p = 0.012), -50° (p < 0.001) and +50° (p = 0.001), while -100° showed no significant difference between the cues (p = 0.06). This result is similar to the one

obtained with only auditory cues earlier. No significant differences were observed between V1S1 and V1S2 at any of the positions (p = 1). Both these conditions show closely matched onset times for head motion, with V1S2 displaying a marginally quicker onset (see figure 8).
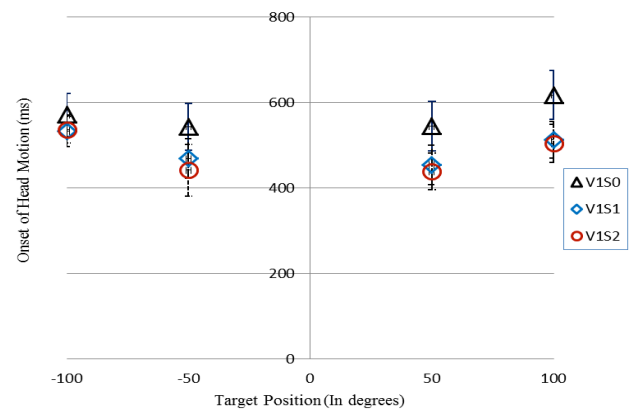

Figure 8: Comparisons between on-set of head motion times for the static visual cue (V1) paired with the auditory cues. A significant difference exists between on-set of head motion times for no auditory cue vs. auditory cueing conditions. No significant difference is seen between the static (S1) and dynamic (S2) auditory cueing conditions when paired with the static visual cue (V1).

For the cueing conditions using V2 paired with the auditory cues (V2S0, V2S1 and V2S2), a significant difference is observed between conditions V2S0 and V2S1 at -100° (p < 0.001), -50° (p = 0.01) and +50° (p = 0.003). No significant difference between the cueing conditions is seen at +100° (p = 0.261). Comparisons between V2S0 and V2S2 display significant differences at -100° (p = 0.001), -50° (p = 0.044), +100° (p = 0.017) and +50° (p < 0.001). Comparisons between the V2S1 and V2S2 pair does not show any significant difference at -100°, +100°, -50° and +50° (p = 1). These cueing conditions (V2S1 and V2S2) appear to have similar onset times across all targets (see figure 9).
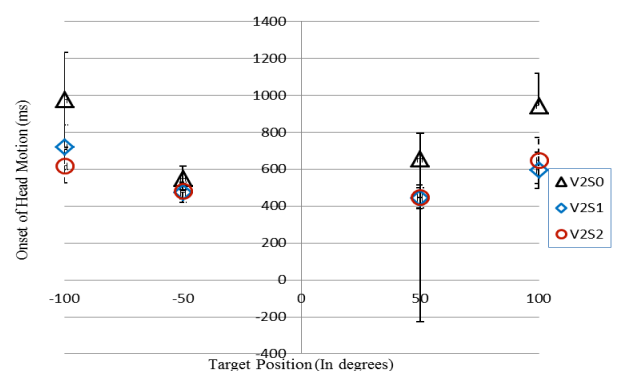

Figure 9: Comparisons between on-set of head motion times for the pursuit visual cue (V2) paired with the auditory cues.

From the analysis that has been carried out, it is clear that the presence of a visual or auditory cue definitely elicits a quicker onset of head motion from the time the target appears at any one of the positions. The lack of significant differences (-100°: p = 1, -50°: p = 0.753, +50°: p = 1 and +100°: p = 1) between the spatialised auditory cue (V0S2) and the static visual cue (V1S0), points to the fact that both these cues are

nearly equally good at redirecting attention. A combination of these two cues though, consistently registers the quickest time for the onset of head motion across all positions even though in some cases these differences do not appear statistically significant. This lack of statistical significance appears mainly when this cue (V1S2) is compared with other cues that include either a binaurally spatialised auditory cue (S2) or a static visual cue (V1). While V1S2 appears to be best suited for attention directions tasks, a comparison of the onset of head motion times between +100° and -100° for this cue shows that the cue performs better for the left side i.e. -100° (506.61 ms vs. 529.53 ms) (see figure 10).
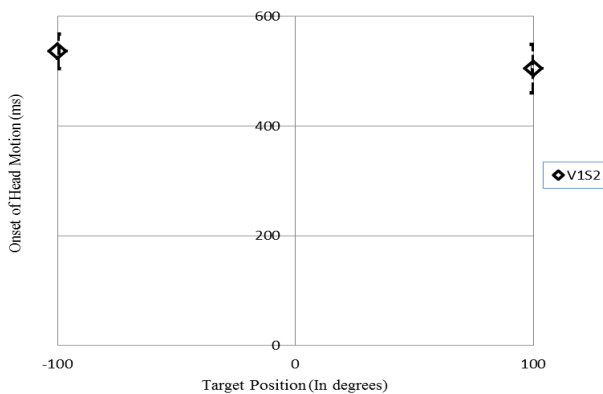

Figure 10: Onset of head motion times for the cueing condition V1S2 (static visual cue and dynamic auditory cue) for ±100°.

7.   DISCUSSION

We've been able to demonstrate the benefits of using auditory cues in an attention redirection task via this study. The results in this case can be categorized in to two distinct types: (1) auditory cues only and (2) audio-visual cues. The first part of the results section falls under the auditory cues category. The results obtained for these cueing conditions tend to suggest that the binaurally spatialised, dynamic auditory cues is effective for redirecting attention to targets that do not occupy the user's field of view i.e. ±100°. The absence of a significant difference between the static (S1) and dynamic (S2) cueing conditions for targets at ±50° is likely due to the fact the targets fall within the user's peripheral vision. The onset of the auditory cues could possibly be the precursor to the participants localizing the target using their peripheral vision. This could be responsible for over-riding both the visual cues, negating their effect. This effect extends across the two auditory cueing conditions S1 and S2 paired with the two visual conditions V1 and V2 for targets at ±50°. Conversely, when either of the auditory conditions paired with a visual cue was compared to the performance without an auditory cue, a clear difference in the onset of head motion times is observed. This is indicative of the fact that even in the presence of a visual cue delivered during a visually demanding task, an auditory cue is more likely to attract attention and help reorient user attention in the space around him/her. Another observation that points to the effectiveness of the binaurally spatialised dynamic cue is the absence of a considerable difference between targets on the same side i.e. ±100° and ±50°. This result effectively demonstrates that the binaurally spatialised auditory cue is as good at redirecting attention to targets outside the visual field as the visual percept is at acquiring targets at ±50° in this experiment.

In the case of combinations of the visual cues, V1 and V2, with auditory cues S1 and S2, the pairing of the static visual cue, V1, with the dynamic auditory cue, S2 appears to provide the best results. As we've demonstrated with the auditory cueing condition only, these results show a superior performance in comparison to other cueing condition pairings when compared with onset of head motion and target acquisition times for targets outside the visual range. This study clearly indicates that the use of auditory cues in conjunction with visual cues to reorient attention is possible. The results from our study compare favourably with those of Perrott et al. [25], Nelson et al. [26] and Rudmann & Strybel [27].

8.   CONCLUSION

We've demonstrated the use of a binaurally spatialised, dynamic auditory cue in conjunction with a visual cue to redirect user attention. These reorientation cues appear to be most effective for targets outside the visual field, but have also shown to be of use within the peripheral vision in comparison to having no auditory cue at all. The use of an auditory cue or alarm in a visually demanding task cannot be underestimated. The dynamic auditory cue appears to be able to redirect the user's attention without inducing a frantic search of the visual field, a behavior that was seen with the static auditory cues. Similar to a '3D' auditory cue delivering azimuth, elevation and distance information used by Nelson el al. [26], our dynamic auditory cue exhibits superior performance compared to the static cue. These results also demonstrate that the binaurally spatialised, dynamic auditory cue will be useful in the event that a user does not latch on to a visual cue that may be presented simultaneously. The outcomes from this study also appear to suggest that a 'dual delivery' of cues across two different modalities appears to ensure that the system is somewhat fail safe.

For the purpose of this experiment only four specific targets were used. In the future, it will be worthwhile exploring how both visual and auditory cues will perform in the presence of visual and auditory distractors. It also worth comparing the performance of binaurally spatialised, co-located cues with spatialised dynamic cues moving towards a target. The results from such experiments could provide further vindication for using the BCH as an auditory display incorporated in to a wearable computing interface and also give us an idea as to which of the cues provide better results for binaural spatialisation over a BCH.

9.   REFERENCES

[1]   Heinrich, M.J., et al., *Wearable computing device*, U.S.P.a.T. Office, Editor. 2013, Google Inc.: U.S.A. p. 11.

[2]   ReconInstruments. *Recon Jet*. 2015 [cited 2015 01 August 2015]; Available from: http://www.reconinstruments.com/products/jet/.

[3]   Walker, A. and S. Brewster, *Spatial Audio In Small Screen Device Displays*. Personal Technologies, 2000. **4**(2-3): p. 144-154.

[4]   WHO, *Mobile Phone Use: A growing Problem of Driver Distraction*, in *Deacde of Action for Road Safety*. 2011, World Health Organization.

[5]   CDC. *Injury Prevention & Control: Motor Vehicle Safety*. 2015 [cited 2016 3rd Feb]; Available from: http://www.cdc.gov/motorvehiclesafety/distracted_driving/.

[6]   DOT. *Distraction: Facts and Statictics*. 2015 [cited 2016 3rd Feb]; Available from: http://www.distraction.gov/stats-research-laws/facts-and-statistics.html.

[7]   Wightman, F.L. and D.J. Kistler, *Headphone Simulation Of Free Field Listening. II: Psychophysical Validation.* The Journal of the Acoustical Society of America, 1989. **85**(2): p. 868 - 878.

[8]   Burgess, D.A. *Techniques For Low Cost Spatial Audio*. in *5th Annual ACM Symposium On User Interface Software And Technology*. 1992. Monteray, California, United States: ACM.

[9]   Wenzel, E.M., et al., *Localization using nonindividualized head‐related transfer functions.* The Journal of the Acoustical Society of America, 1993. **94**(1): p. 111-123.

[10]   Sawhney, N. and C. Schmandt. *Nomadic Radio: Scaleable and Contextual Notification For Wearable Audio Messaging*. in *SIGCHI conference on Human Factors in Computing Systems*. 1999. ACM.

[11]   Cheng, C.I. and G.H. Wakefield, *Introduction to Head Related Transfer Functions: Representations of HRTFs in Time, Frequency and Space.* Journal of the AES, 2001. **49**(4): p. 231 - 249.

[12]   Wilkins, P.A. and W.I. Acton, *Noise and accidents—a review.* Annals of Occupational Hygiene, 1982. **25**(3): p. 249 - 260.

[13]   Lichenstein, R., et al., *Headphone Use and Pedestrian Injury and Death in the United States: 2004 - 2011.* Injury Prevention, 2012.

[14]   Harma, A., et al., *Augmented Reality Audio For Mobile And Wearable Appliances.* Journal of the Audio Engineering Society, 2004. **52**(6): p. 618-139.

[15]   Tikander, M., M. Karjalainen, and V. Riikonen. *An Augmented Reality Audio Headset*. in *11th Int. Conference on Digital Audio Effects (DAFx-08)*. 2008. Espoo, Finland.

[16]   Labs, D. *Here Active Listening*. 2015 [cited 2016 6 Feb]; Available from: https://www.hereplus.me/.

[17]   Oldfield, S.R. and S.P.A. Parker, *Acuity of Sound Localization: A Topography of Audiotory Space. II. Pinna Cues Absent.* Perception, 1984. **13**: p. 601 - 617.

[18]   Walker, B.N. and J. Lindsay, *Navigation Performance In A Virtual Environment With Bonephones*, in *International Conference On Auditory Displays*. 2005: Limerick, Ireland.

[19]   MacDonald, J.A., P.P. Henry, and T.R. Letowski, *Spatial Audio Through A Bone Conduction Interface.* International Journal Of Audiology, 2006. **45**(10): p. 595 - 599.

[20]   Walker, B.N. and J. Lindsay, *Navigation Performance With A Virtual Auditory Display: Effects Of Beacon Sound, Capture Radius and Practice.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2006. **48**(2): p. 265 - 278.

[21]   Lindeman, R.W., H. Noma, and P.G.d. Barros, *Hear-Through and Mic-Through Augmented Reality: Using Bone Conduction To Display Spatialized Audio.* 2007.

[22]   Stanley, R.M., *Measurement And Validation Of Bone Conduction Adjustment Functions In Virtual 3D Audio Displays*, in *School of Psychology*. 2009, Georgia Institute of Technology.

[23]   Valjamae, A., et al., *Binaural Bone Conducted Sound In Virtual Environments: Evaluation Of A Portable, Multimodal Motion Simulator Prototype.* Acosutical Science And Technology, 2008. **29**(2): p. 149 - 155.

[24]   Villegas, J. and M. Cohen, *GABRIEL: Geo-Aware Broadcasting For In-Vehicle Entertainment And Localizability*, in *AES 40th International Conference*. 2010, AES: Tokyo, Japan.

[25]   Perrott, D.R., et al., *Aurally Aided Visual Search in the Central Visual Field: Effects of Visual Load and Visual Enhancement of the Target.* Human Factors, 1991. **33**(4): p. 389 - 400.

[26]   Nelson, W.T., et al., *Effects of Localized Auditory Information on Visual Target Detection Performance Using a Helmet-Mounted Display.* Human Factors, 1998. **40**(3): p. 452 - 460.

[27]   Rudmann, D.S. and T.Z. Strybel, *Auditory Spatial Facilitation of Visual Search Performance: Effects of Cue Precision and Distractor Density.* Human Factors, 1999. **41**(1): p. 146 - 160.

[28]   ART, *DTrack*. 1999, Advanced Realtime Tracking: Munich, Germany.

[29]   ART. *ARTTRACK2*. 1999 [cited 2015 1st Dec]; Available from: http://www.ar-tracking.com/products/discontinued/arttrack2/.

[30]   Technologies, S. *Steradian S7-X Laser Tag Gun*. 2003 [cited 2015 12 Sept]; Available from: http://www.steradiantech.com/xseries/s7x/.

[31]   II, R.M.T., *Virtual Reality Peripheral Network (VRPN)*. 1998.

[32]   II, R.M.T., et al. *VRPN: A Device-Independent, Network-Transparent VR Peripheral System*. in

Proceedings of the ACM Symposium on Virtual Reality Software and Technology. 2001. ACM.

[33]   Aftershokz. *Sportz3*. [cited 2015 2nd June]; Available from: http://aftershokz.com/collections/wired/products/sportz-3.

[34]   Walker, B.N. and G. Kramer, *Auditory Displays, Alarms and Auditory Interfaces*, in *International Encyclopedia of Ergonomics and Human Factors*, W.K. Informa Healthcare, Editor. 2006, CRC Press. p. 1021 - 1025.

[35]   Walker, B.N. and R.M. Stanley, *Thresholds Of Audibility For Bone Conduction Headsets*, in *International Conference On Auditory Display*. 2005, ICAD: Limerick, Ireland.

[36]   Gardner, W.G., *3D Audio Using Loudspeakers*, in *School of Architecture and Planning*. 1997, Massachusetts Institute of Technology.

[37]   Wersenyi, G., *Localization In A HRTF Based Minimum Audible Angle Listening Test On a 2D Sound Screen For GUIB Applications*, in *115th Convention of the Audio Engineering Society* 2003.

[38]   Stevens, S.S. and E.B. Newman, *The Localization of Actual Sources of Sound.* The American Journal of Psychology, 1936. **48**(2): p. 297 - 306.

[39]   Weinrich, S.G., *Horizontal Plane Localization Ability and Response Time as a Function of Signal Bandwidth*, in *98th Convention of the AES*. 1995: Paris, France.

[40]   Walker, B.N., R.M. Stanley, and J. Lindsay. *Task, User Characteristics, and Environment Interact to Affect Mobile Audio Design*. in *PERVASIVE*. 2005.

[41]   Helgason, D., N. Francis, and J. Ante, *Unity3d*. 2005, Unity Technologies.

[42]   Thakur, A. and V. Nair, *3Deception*. 2015, Two Big Ears: Edinburg, Scotland.