Investigating Different Modalities of Directional Cues for Multi-task Visual-Searching Scenario in Virtual Reality

Taizhou Chen
School of Creative Media
City University of Hong Kong
taizhou.chen@my.cityu.edu.hk

Yi-Shiun Wu
School of Creative Media
City University of Hong Kong
yiswu2@cityu.edu.hk

Kening Zhu
School of Creative Media
City University of Hong Kong
kening.zhu@cityu.edu.hk

Abstract

In this study, we investigated and compared the effectiveness of visual, auditory, and vibrotactile directional cues on multiple simultaneous visual-searching tasks in an immersive virtual environment. Effectiveness was determined by the task-completion time, the range of head movement, the accuracy of the identification task, and the perceived workload. Our experiment showed that the on-head vibrotactile display can effectively guide users towards virtual visual targets, without affecting their performance on the other simultaneous tasks, in the immersive VR environment. These results can be applied to numerous applications (e.g. gaming, driving, and piloting) in which there are usually multiple simultaneous tasks, and the user experience and performance could be vulnerable.

CCS Concepts

• **Human-centered computing → Virtual reality**; *Haptic devices*;

Keywords

Directional cue, Vibration, Visual, Auditory, Multi-task, Virtual reality

ACM Reference Format:
Taizhou Chen, Yi-Shiun Wu, and Kening Zhu. 2018. Investigating Different Modalities of Directional Cues for Multi-task Visual-Searching Scenario in Virtual Reality. In *VRST 2018: 24th ACM Symposium on Virtual Reality Software and Technology (VRST '18), November 28-December 1, 2018, Tokyo, Japan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3281505.3281516

1 Introduction

While searching objects in real-life environments, humans perceive rich, coherent multimodal feedback comprised of visions, sounds, smells, tastes, and haptic sensations. Many virtual-reality (VR) applications seek to emulate this richness in object searching, and incorporate multisensory feedback, typically combining visual, auditory, and haptic feedbacks. The visual modality is usually preferred for cueing the target-searching tasks in VR, as it has a higher information bandwidth than the other modalities [21]. Besides the visual rendering, common VR applications also use spatial sound for immersive user experience. Spatial sound effect could also serve an indication for object location, to improve the user performance in various applications (e.g., driving/piloting simulator [1], language education [6], etc.). However, directional sound could be easily affected by any background sound in the application. Research showed that tactile directional feedback could outperform the audio-based cueing [2, 11], and it can potentially help offloading visual perception [10]. Specifically, vibrotactile devices integrated with VR head-mounted displays (e.g., vibrational headbands and helmets) have been used for increasing perceived presence in virtual environments [13], for obstacle detection [4, 1], and to indicate elements that are placed outside the visual field [9, 18, 19]. The information conveyed by the vibration is explicit, as the sensation directly evokes the behavior [1]. Existing researches [7, 16, 17] has distilled guidelines and insights for localization with on-head vibrotactile feedback, suggesting that vibrotactile HMDs could be optimized with few (4 to 8) motors placed on the forehead, occipital, and temple regions of the head.

While existing research proved the effectiveness of on-head vibrotactile cue over audio for visual target searching [12, 14, 15], these studies focused on simple single-task scenario (i.e. there was only one task of searching one visual target at a time). Target searching or navigation in real life often involves multiple simultaneous visual tasks or distraction. For example, drivers often need to look at points of interests outside their car, (road signs, traffic lights) as well as inside the car (GPS system, radio) while keeping into consideration external factors (traffic, pedestrians, etc.). Multiple-resource theory [22] argues that the distribution of tasks and information across various modalities of sensation could reduce cognitive workload. Cue effectiveness may likely vary in the distractive and multitasking situation (e.g., responding to an interruption while looking for targets), which requires higher workload. This situation could also happen in VR which have been commonly used for simulating the real-life experience. Furthermore, the HMD-based VR, which is usually visually intensive, may affect the user performance as well. While it's arguable that tactile feedback may outperform audio, and offload visual processing for the users in complex target-searching tasks, there is lack of quantitative and qualitative evidence showing how they would perform differently in a complex multi-task VR scenario.

To fill this gap, we investigated how the three different directional cues (visual arrow icon, spatial sound effect, and on-head directional vibration) will affect visual target searching in a multi-task VR scenario. In our experiment, participants searched two visual targets

VRST '18, November 28-December 1, 2018, Tokyo, Japan

ACM ISBN 978-1-4503-6086-9/18/11.
https://doi.org/10.1145/3281505.3281516

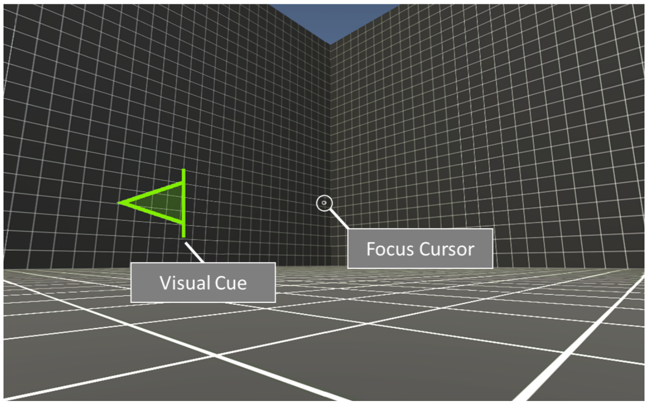

Figure 1: Visual cue: arrow icon

simultaneously among many distractions. Our results showed that the vibrotactile cue was significantly faster and more accurate than the auditory cue. Although the visual cue outperformed the other two for the task-completion time and the range of head movement, the vibrotactile cue was significantly more accurate than the visual cue in the hard mode of the identification task.

2 Design of Directional Cues

For the purpose of our study, we developed an experimental virtual environment in Unity 3D. As shown in Fig. 4, The virtual environment contained multiple spherical objects. Along with the rendering of the targets, we implemented three types of directional cues: visual, audio, and vibration.

2.1 Visual Cue

Ward et al.'s experiments [20] showed that the arrow icon is one of the more efficient visual cueing option for indicating directions. Arrow widely used in games [23]. Compared to other visual cues, such as attention funnels [1], arrow occupies a smaller space in the display area. In our study, when a target is spawned outside the current view, an arrow icon (left/right) will be displayed on the left or right side of the display (Fig. 1), indicating the target's position relative to the current viewing direction.

2.2 Auditory Cue

For the audio cue, we followed the setup in HapticHead [12], and used the built-in sound effect in combination with Unity3D 5.3's included spatial sound system with "spatial blend" set to 1 (full 3D) and Mighty Rock HE8G noise cancellation earplugs. Fig. 2 illustrates the frequency spectrum of the auditory cue. Please refer to our supplementary video to listen to the auditory cue.

2.3 Vibrotactile Cue

We built a vibrotactile headband (Fig. 3(a)) with eight electromechanical vibration motors controlled by the PWM signals from one Arduino Uno board. All motors, 10mm flat vibration motor (Model No.: 1027, DC3.0V, 90mA, 9000 rpm), were attached to elastic band with wires, to be easily worn around the head, and allow easy adjustment to different head sizes. As shown in Fig. 3(b), all motors were placed at equal distance from the center of the forehead over the Cardinal (North, East, South, and West) and Collateral points (NE, SE, SW, and NW).

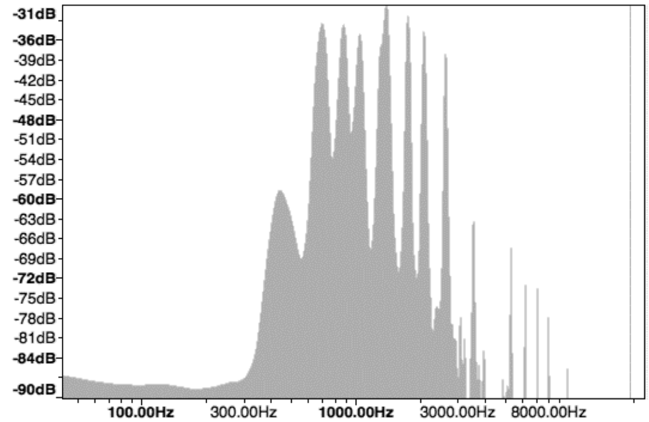

Figure 2: Frequency spectrum of the auditory cue

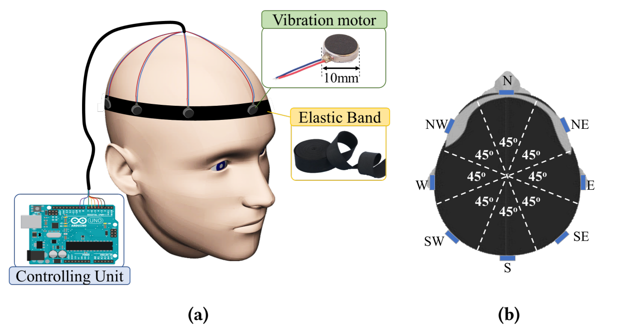

Figure 3: Vibrotactile headband design: (a) system diagram (b) motor arrangement.

The VR scene calculated the angular distance between the camera and the target, and signaled the Arduino to activate the corresponding motor at the highest intensity. The same actuator was actuated constantly if the user stayed still. In addition, actuator actuation was adjusted with head rotation. As the user turns the head towards the target, the signal travels along the trajectory towards the front of the head.

3 Experiment: Multiple-Target Searching

We conducted a within-subject experiment to compare the effectiveness of these directional cues for multi-task target searching in VR.

3.1 Participants

We recruited twenty-four participants (twelve females). Their ages ranged from 23 to 32 years (M = 24.8, SD = 2.44). All have tried VR HMDs for only once or twice.

3.2 Tasks

In this experiment, the participants needed to accomplish a primary task, finding the visual target (a sphere) among distraction as fast as possible, as the primary task, by following one particular type of the three directional cues. We further increased the complexity of the visual-searching tasks by adding distractive objects and parallel tasks. The VR experiment contained 100 spheres as the distractive objects (Fig. 4). There was only one real target sphere which was always spawned at the eye-level of the user. Thus, the directional

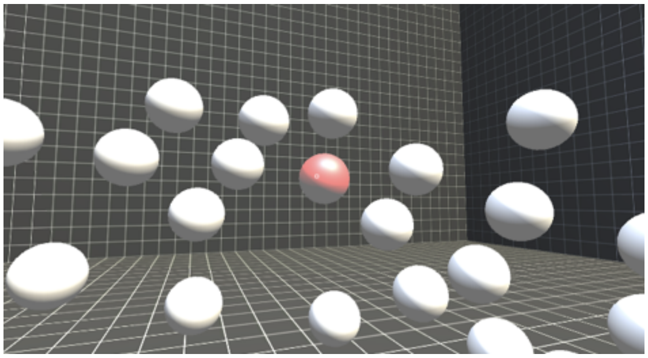

Figure 4: Visual indication of finding the primary target.

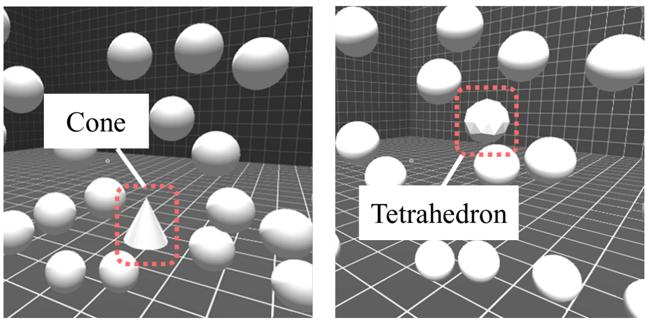

Figure 5: Target objects for the secondary task: left - easy mode, right- hard mode.

cues, especially the auditory and the vibrotactile cues, were displayed in 2D. When the real target fell into the participant's view, the directional cue would disappear/pause. When the participant focused/hovered the cursor on the real target, it will be "highlighted" based on the type of directional cue used in this session. That is, the target sphere turned red in the Visual session (Fig. 4); a special sound effect which was different from the directional auditory cue in the Auditory session; all motors vibrated at the highest intensity in the Vibrotactile session. The participants confirmed by hovering the cursor on the object for three seconds.

Besides performing the primary task, the participants needed to simultaneously perform a secondary visual-searching task in which they need to identify a visually distinctive object among the distractive objects. To ensure the possibility of accomplishing the secondary task, the visually distinctive object was spawned at a random location within the shortest angular path between the current camera direction and the primary-task visual target, yet didn't fall into the same field of view as the primary target. In addition, we implemented two modes, easy and hard, of difficulties for this secondly visual-searching task. As shown in Fig. 5, the visually distinctive object is a cone in the easy mode, while it's a polyhedron which is visually more closed to a sphere in the hard mode. Note that the participant didn't need to select the visually distinctive object with his/her gaze. Once the participant was sure of finding the visually distinctive object in his/her view, he/she placed a single click on the Bluetooth mouse in his/her hand (Fig. 6). were six sessions (3 direction cues × 2 modes of difficulties) for each participant. Their order was counterbalanced with a Latin square.

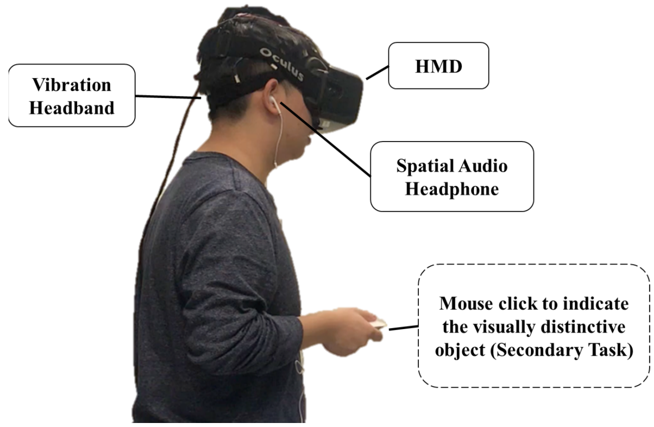

Figure 6: Experiment setup.

3.3 Procedure

In each session, the participant started with a training of finding five sets of primary targets and visually distinctive targets (one set at a time), to get familiar with the tasks and the particular type of direction cue. After training, the participants proceeded to the actual test session. There was a compulsory break after each session, and the participant needed to finish the user-experience questionnaire and the NASA TLX questionnaire [8]. The experiment took around one hour per participant. Each participant performed 120 trials: 3 direction cues (visual, auditory, vibrotactile) × 2 modes of difficulties (easy, hard) × 20 targets. As a reward, each participant received a 7USD shopping coupon.

3.4 Results

We measured the completion time of the primary task, the range of head movement, the accuracy of the secondary task, the perceived workload, and the user-experience questionnaire, as the dependent measurements. Table 1 shows the descriptive results. We performed two-way ANOVA on these dependent measurements, taking the type of the directional cue and the mode of difficulty as the independent factors. The results showed that the type of the directional cue significantly affected the participants' performance: the range of head movement ($F(2, 2874) = 63.33$, $p < 0.001$), the completion time for the primary task ($F(2, 2874) = 20.18$, $p < 0.001$), and the accuracy of the secondary task ($F(2, 210) = 4.68$, $p < 0.05$). Post-hoc pairwise comparison showed the visual cue yielded significantly less completion time and smaller range of head movement than the tactile cue ($p < 0.005$) and the auditory cue ($p < 0.001$). The tactile cue was significantly faster than the auditory cue ($p < 0.001$), while there is no significant difference between the tactile and the visual cue. For the secondary task, the vibrotactile cue was significantly more accurate than the auditory cue ($p < 0.05$), and it was slightly but not significantly more accurate than the visual cue.

Two-way ANOVA revealed a significant interaction effect of the type of the directional cue and the level of difficulty on the accuracy of the secondary task ($F(2, 210) = 18.63$, $p < 0.0005$). Post-hoc pairwise comparison showed that in the easy mode, the auditory cue was significantly less accurate than both the vibrotactile and the visual cue, and the vibrotactile cue was slightly but not significantly more accurate than the visual cue. In the hard mode, the vibrotactile cue

Table 1: Average task-completion time, range of head movement, accuracy of secondary searching task, and workload score among the three modalities. SD is standard deviation.

		Completion Time (s)	Range of Head Movement (degree)	Accuracy of Secondary Task (%)	NASA TLX Workload Score
Easy Mode	Auditory	12.39 (SD = 0.42)	356.4 (SD = 18.7)	68.40% (SD = 0.08)	45.83 (SD = 3.66)
	Vibrotactile	11.30 (SD = 0.35)	369.5 (SD = 19.1)	87.30% (SD = 0.14)	41.15 (SD = 3.80)
	Visual	9.92 (SD = 0.18)	305.7 (SD = 13.3)	82.00% (SD = 0.16)	37.31 (SD = 3.83)
Hard Mode	Auditory	14.34 (SD = 0.29)	350.4 (SD = 18.3)	75.40% (SD = 0.17)	48.85 (SD = 3.80)
	Vibrotactile	11.94 (SD = 0.42)	372.8 (SD = 16.9)	83.80% (SD = 0.05)	42.02 (SD = 3.67)
	Visual	10.23 (SD = 0.48)	323.5 (SD = 11.0)	68.40% (SD = 0.12)	48.47 (SD = 3.95)

was significantly more accurate than the visual cue, and there was no significant difference between the accuracy of the vibrotactile cue and the auditory cue. Pairwise comparison further showed that the visual cue was significantly more prone to error in the hard mode than in the easy mode of the secondary task ($p < 0.05$), while there was no significant difference between the accuracy in the easy mode and in the hard mode for both the vibrotactile and the auditory cues. This suggested that the visual cue was affected more by the increment of the difficulty on the secondary task.

The perceived workload was calculated into the range of 0 - 100 by averaging the raw NASA TLX data collected through the questionnaire, and the descriptive results are plotted in Fig. 12. This method has been proved to increase experimental validity [3]. One-way ANOVA revealed that the score of the perceived work load with the visual cue significantly increased in the hard mode than in the easy mode ($F(1, 46) = 4.55$, $p < 0.05$), while there was no significant difference between the perceived workload in these two modes for the other two types of direction cues.

The qualitative comments from the participants further supported these statistical findings. For the visual cue, the participants commented, "My eyes felt really tired.", "Sometimes I felt dizzy by looking at the arrows and searching the polygon object", "It's very easy to miss the distinctive object." For the vibrotactile cue, they mentioned, "It's very natural to tell the direction. It's like someone tapping you.", but "It's hard to feel the vibration when I am moving my head.", "After a long time, it's hard for me to feel the vibration." For the auditory cue, they mostly found it "confusing to tell left or right", and they "needed to be really focused".

In sum, these results suggested that (a) the vibrotactile directional cue is more effective than the auditory cue in the multi-task VR scenario, and (b) as the intensity of the visual content and the complexity of the tasks increase, the vibrotactile cue could result in a comparable and even better user performance than the visual cue does.

4 Limitation & Future Work

One possible explanation for the demonstrated advantage of the vibrotactile cue over the visual cue for the accuracy of the identification task and the perceived overall workload is that, unlike the visual cue, the vibrotactile cue is often under-utilised in conveying structured, specific task-relevant information. In this study, there was no other simultaneous task information being conveyed through the tactile modality. The user performance in the multi-task VR scenario could vary with simultaneous tactile feedbacks (e.g. vibration, temperature, and force) indicating different tasks. In addition, the user stood still in our current experimental setup. Situating the users in a moving condition (e.g., simulating the driving experience over a bumpy road) may affect the perception of the vibrotactile feedback, and result in different performance.

Secondly, the directional cue paused when the target of the primary task appeared in the camera view of the HMD, meaning the participant needed to search the target within this view without any cue, which could influence the performance. However, this issue could be solved with a sufficient number of randomized control trials [5] (i.e. 360 trials for each of the six conditions in our study).

Lastly, the participants were asked to identify primitive 3D objects (i.e. sphere, cone, and tetrahedron), to which they may not be familiar. This might affect their performance. In real-life scenarios, the users are usually trained for years to identify certain objects. For example, an experienced driver can quickly spot road signs along the street. As the future work, we would like to study the effectiveness of different directional cues in the real-life simulation settings (e.g. driving and piloting).

5 Conclusion

In this study, we compared the effectiveness of the visual, auditory, and vibrotactile directional cue on multiple simultaneous visual-searching tasks in an HMD-based immersive VR environment. Generally, the vibrotactile cue was significantly faster and more accurate than the auditory cue. This could be explained by the fact that the visual modality and the auditory visual are usually linked cognitively [2]. Although the visual cue outperformed the other two in terms of the task-completion time and the range of head movement, the vibrotactile cue was significantly more accurate than the visual cue in the hard mode of the secondary task. Our findings also revealed that the increment of the task difficulty affected the visual cue significantly more than it did to the other two modalities. These results can be applied to various multi-task and visually intensive VR applications (e.g. gaming, driving, and piloting) in which the user experience and performance could be vulnerable with many simultaneous tasks.

Acknowledgments

This research was partially supported by grant from the Centre for Applied Computing and Interactive Media (ACIM) of School of Creative Media, the Strategic Research Grants (Project No. 7005021 & 7005172), the Teaching Development Grang (Project No. 6000623 & 6000639), City University of Hong Kong, and grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21200216).

References

[1] Beattie, D., Baillie, L., Halvey, M., & McCall, R. (2014, October). What's around the corner?: enhancing driver awareness in autonomous vehicles via in-vehicle

spatial auditory displays. In Proceedings of the 8th nordic conference on human-computer interaction: fun, fast, foundational (pp. 189-198). ACM.
[2] Burke, J. L., Prewett, M. S., Gray, A. A., Yang, L., Stilson, F. R., Coovert, M. D., ... & Redden, E. (2006, November). Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In Proceedings of the 8th international conference on Multimodal interfaces (pp. 108-117). ACM.
[3] Bustamante, E. A., & Spain, R. D. (2008, September). Measurement invariance of the Nasa TLX. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 52, No. 19, pp. 1522-1526). Sage CA: Los Angeles, CA: SAGE Publications.
[4] Cassinelli, A., Reynolds, C., & Ishikawa, M. (2006, October). Augmenting spatial awareness with haptic radar. In Wearable Computers, 2006 10th IEEE International Symposium on (pp. 61-64). IEEE.
[5] Chalmers, T. C., Smith, H., Blackburn, B., Silverman, B., Schroeder, B., Reitman, D., & Ambroz, A. (1981). A method for assessing the quality of a randomized control trial. Controlled clinical trials, 2(1), 31-49.
[6] Danna, J., Fontaine, M., Paz-Villagrán, V., Gondre, C., Thoret, E., Aramaki, M., ... & Velay, J. L. (2015). The effect of real-time auditory feedback on learning new characters. Human movement science, 43, 216-228.
[7] Gilliland, K., & Schlegel, R. E. (1994). Tactile stimulation of the human head for information display. Human Factors, 36(4), 700-717.
[8] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in psychology, 52, 139-183.
[9] Hayward, V., & MacLean, K. E. (2007). Do it yourself haptics: part I. IEEE Robotics & Automation Magazine, 14(4).
[10] Hopkins, K., Kass, S. J., Blalock, L. D., & Brill, J. C. (2017). Effectiveness of auditory and tactile crossmodal cues in a dual-task visual and auditory scenario. Ergonomics, 60(5), 692-700.
[11] Kaul, O. B., & Rohs, M. (2016, May). Haptichead: 3d guidance and target acquisition through a vibrotactile grid. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (pp. 2533-2539). ACM.
[12] Kaul, O. B., & Rohs, M. (2017, May). 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 (pp. 3729-3740). ACM.
[13] Kaul, O. B., Meier, K., & Rohs, M. (2017, September). Increasing Presence in Virtual Reality with a Vibrotactile Grid Around the Head. In IFIP Conference on Human-Computer Interaction (pp. 289-298). Springer, Cham.
[14] Kerdegari, H., Kim, Y., & Prescott, T. J. (2016, July). Head-Mounted Sensory Augmentation Device: Comparing Haptic and Audio Modality. In Conference on Biomimetic and Biohybrid Systems (pp. 107-118). Springer International Publishing.
[15] Marsalia, A. (2013). Evaluation of vibrotactile alert systems for supporting hazard awareness and safety of distracted pedestrians (Doctoral dissertation).
[16] Rash, C. E., Russo, M. B., Letowski, T. R., & Schmeisser, E. T. (2009). Helmet-mounted displays: Sensation, perception and cognition issues. ARMY AEROMEDICAL RESEARCH LAB FORT RUCKER AL.
[17] ępakov, O., Rantala, J., & Isokoski, P. (2015, June). Sequential and simultaneous tactile stimulation with multiple actuators on head, neck and back for gaze cuing. In World Haptics Conference (WHC), 2015 IEEE (pp. 333-338). IEEE.
[18] Turchet, L., Burelli, P., & Serafin, S. (2013). Haptic feedback for enhancing realism of walking simulations. IEEE transactions on haptics, 6(1), 35-45.
[19] van Erp, J. (2001). Tactile navigation display. Haptic human-computer interaction, 165-173.
[20] Ward, M., Barde, A., Russell, P. N., Billinghurst, M., & Helton, W. S. (2016, September). Visual cues to reorient attention from head mounted displays. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 60, No. 1, pp. 1574-1578). Sage CA: Los Angeles, CA: SAGE Publications.
[21] Way, T. P., & Barner, K. E. (1997). Automatic visual to tactile translation. i. human factors, access methods and image manipulation. IEEE Transactions on rehabilitation engineering, 5(1), 81-94.
[22] Wickens, C. D. (1991). Processing resources and attention. Multiple-task performance, 1991, 3-34
[23] Wuertz, J., Alharthi, S. A., Hamilton, W. A., Bateman, S., Gutwin, C., Tang, A., ... & Hammer, J. (2018, April). A Design Framework for Awareness Cues in Distributed Multiplayer Games. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (p. 243). ACM.