Providing 3D Guidance and Improving the Music-Listening Experience in Virtual Reality Shooting Games Using Musical Vibrotactile Feedback

Yusuke Yamazaki* Shoichi Hasegawa†

Tokyo Institute of Technology

ABSTRACT

In this study, we aim to improve the experience of virtual reality (VR) shooting games by employing a 3D haptic guidance method using necklace-type and belt-type haptic devices. Such devices help to modulate the vibrations generated by and synchronized with musical signals according to the azimuth and height of a target in 3D space, which is expected to improve the gaming experience by providing 3D guidance and enhancing the music-listening experience. For the first step, we evaluated the method's potential by conducting an experiment in which participants were asked to shoot a randomly spawned target moving in 3D VR space. The experiment applied four conditions: the proposed method (Haptic), displaying 3D radar (Vision) to represent the visualization method, no guidance (None), and a combination of Haptic and Vision (VisHap). Outcomes related to the success rate and accomplishment time (of the shooting task), the number of head rotations, and participant responses to a follow-up questionnaire revealed that Haptic performed significantly better than None but was inferior to Vision, indicating that the proposed method succeeded in terms of effectively providing 3D guidance. VisHap performed roughly as well as Vision and was preferred to other conditions in most cases, indicating the general usefulness of the proposed method. Meanwhile, the findings from the questionnaire suggest that although the modular vibrations improved the music-listening experience during the shooting task, the impact on the overall gaming experience is unclear. This warrants further research.

Index Terms: Human-centered computing—Haptic devices; Human-centered computing—Virtual reality; Human-centered computing—Sound-based input / output

1 INTRODUCTION

One of the attractions of virtual reality (VR) games is that players can intuitively interact with their 3D surroundings through embodied movement. However, the limited field of view permitted by existing head mount displays (HMDs) can constrain VR game design. To address this limitation, researchers have proposed methods including superimposing a user interface (UI), such as a 3D radar and arrows that display out-of-view objects [4, 11, 12, 14], using haptic feedback, such as vibration, pressure, or skin stretching [7, 10, 15, 22, 28, 39, 42], and combining multiple modalities of audio-visual and audio-tactile sensation [3, 26, 38].

In addition, improving player experience (PX) has been identified as important for VR games. Among the many factors involved in PX is the background music in video games (hereafter referred to as "game music"), the significance of which many researchers are beginning to recognize [6, 23, 24, 29, 31, 47]. For example, Cassidy *et al.* [6] have observed that playing one's preferred music during a driving game improves the gaming score, and Zhang *et al.* [47] have reported that game music increases the player's level of immersion in the game. Meanwhile, in an example of haptic feedback, researchers have also reported that stimulating vibrations generated by and synchronized with music (musical vibrations) can enhance the music-listening experience compared to audio-only listening [19, 27].

This background research motivated us to develop the notion of improving the PX of VR games by simultaneously utilizing 3D guidance and enhancing the music-listening experience. More specifically, we previously proposed that modulating the amplitude of the stimulated music-associated vibrations according to amplitude based on the target's location (henceforth, modulated musical vibrations) can improve both 2D navigation and the music-listening experience [44]. In this paper, we extend the proposed method to 3D guidance and test the following hypotheses to explore our theory, selecting 3D radar [13] as the visualization method for comparison.

- H1: Stimulating modulated musical vibrations allows players to find a moving target in out-of-view 3D space.
- H2: Stimulating modulated musical vibrations gives players the same ability to detect an out-of-view target as displaying a 3D radar.
- H3: Displaying a 3D radar in conjunction with stimulating modulated musical vibrations reduces the time that players spend looking at the 3D radar.
- H4: Stimulating modulated musical vibrations improves PX by enhancing the experience of listening to game music.

2 RELATED WORK

2.1 3D guidance methods using haptic feedback

Researchers have proposed haptic-based 3D guidance methods that use small vibrators and skin stretching.

2.1.1 Vibrotactile 3D guidance

Günther *et al.* [15] proposed a glove-type device with nine built-in vibrators on the hand's dorsum and one on the palm to help guide the user's hand to an arbitrary location in 3D space by controlling vibration patterns. Fiannaca *et al.* [10] attempted 3D localization by modulating the frequency and pulse delay of the built-in vibrator of a commercial game controller held in both hands. Although this method helped to guide the participant's hand to arbitrary 3D coordinates, the process was rather time-consuming (approximately 13 s), which posed a practical challenge. Thus, Kaul *et al.* [22] subsequently proposed the HapticHead, a cap-type device featuring 20 small vibrators. Their experimental results demonstrate that the HapticHead allows users to locate an object in 3D space rapidly (2.6 s) and accurately (96.4%), outperforming the results obtained for spatial audio guidance (6.9 s and 54.2%). Meanwhile, Oliveira *et al.* [7] attached seven vibrators to the face cover (forehead) and headband (side of the head) of the HMD and successfully presented elevation and azimuth angles to the target by changing the vibration pattern and frequency.

*e-mail: yus988@haselab.net
†e-mail: hase@haselab.net

2.1.2 Skin-stretching 3D guidance

Tsai *et al.* [39] proposed the wristband device GuideBand, which applies a 3D tensile force to the forearm to enable 3D guidance and haptic feedback in VR applications. Elsewhere, Wang *et al.* [42] mounted and controlled six skin-stretching modules on the face cover of an HMD to create a sense of weight, inertia, impact, and left-right and vertical guidance. Finally, Nakamura *et al.* [28] proposed Virtual Whiskers, which involved attaching two small robotic arms to the HMD to stimulate the user's cheeks. They showed that Virtual Whiskers could guide players with an average accuracy of around 2.76° in azimuth and 7.32° in elevation.

However, to the best of our knowledge, no studies have attempted to achieve 3D guidance and improve the music-listening experience simultaneously. Moreover, because these studies have all focused solely on static targets, their usefulness for dynamic targets, such as enemies in VR shooting games, remains unknown.

2.2 Improving PX with game music

In the context of PC games, many studies have reported that music influences PX and in-game behaviors such as immersion and enjoyment [6, 23, 24, 29, 31, 47]. However, according to Rogers *et al.* [32], although game music in VR games accelerates participants' time perception, it does not affect PX, including immersion [32, 33]. Nonetheless, it should be noted that they did not attempt to enhance the music-listening experience using haptic stimuli.

The synergy between music listening and haptic stimuli has attracted increasing attention, with several companies now selling haptic devices to enrich music listening, including Woojer [43], SUBPAC [35], and Hapbeat [16]. Several researchers have also reported the effectiveness of stimulating musical vibrations, especially on low-frequency bands, for the music-listening experience. For instance, Merchel *et al.* [27] had participants sit in a whole-body vibration device before asking them to evaluate the music-listening experience with and without musical vibrations. The participants highly rated the condition of pop music with musical vibrations, which featured a strong frequency component of bass sounds. Furthermore, using SUBPAC M2X, a backpack with a large built-in linear vibrator, as the haptic device, Hove *et al.* [19] demonstrated that the low-frequency band of musical vibrations enhanced participants' sense of groove when listening to the music,

However, to the best of our knowledge, only Carroll *et al.* [5] have investigated the effects of musical vibrations on PX in VR games. They had participants play a VR rhythm game (Beat Saber [2]) using a vest-type device (bHaptics Inc., TactSuit X 40) with 40 small built-in vibrators and reported no difference in immersion or PX. Although their results seem to support the findings of Rogers *et al.* [32], they are also questionable. This is because the built-in vibrators of TactSuit X 40 are all eccentric rotating mass vibration motors that cannot sufficiently output the low-frequency musical vibration band, limiting the capacity of Carroll *et al.*'s [5] study to properly represent musical vibrations. This makes it quite possible that using a proper haptic device to stimulate musical vibrations would produce rather different PX results in the context of VR games.

3 Proposal

This paper proposes a 3D haptic guidance method that primarily improves the PX of VR shooting games by both achieving 3D guidance and improving the music-listening experience by modulating musical vibrations according to the height and azimuth angle of a target.

3.1 Haptic device

We use Hapbeat, which utilizes the tension-based vibration generation mechanism proposed by Yamazaki *et al.* [45]. Hapbeat can transmit high-amplitude, low-frequency vibrations over a wide body area by converting the rotation of a DC motor's shaft into the translational motion of a satin ribbon in contact with the body. Because this method stimulates two body areas to articulate the target's height (h in Fig. 3), we use a necklace-type Hapbeat [46] for the upper body and a belt-type Hapbeat for the lower body (Fig. 1(a, d)). Each device features two built-in motors and can independently stimulate modulated musical vibrations in the left-right ribbon and body contact areas by inputting separate audio signals. Although the modulation algorithm described in the following subsections can be applied to any actuator that can independently control vibration frequency and amplitude, such as a linear vibrator, we chose Hapbeat for the following reasons. Hapbeat can output over the vibration amplitude of 10 m/s^2 in a range of around 10 to 400 Hz [46] and is suitable for stimulating musical vibrations due to its wide dynamic range of amplitude over a wide frequency range. Furthermore, its wide dynamic range enables greater variation in vibration amplitude, making it easier for participants to distinguish the variation.

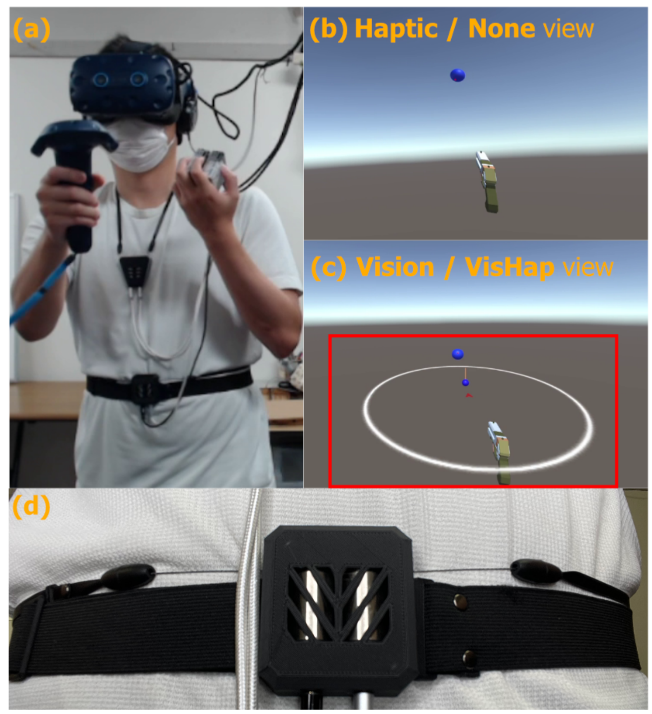

Figure 1: Experimental environment. (a): Photo of a participant during the experiment. (b): A participant's view in Haptic and None. (c): A participant's view in Vision and VisHap. The red border defines the radar area for "gaze at radar area" explained in Section 4.8. (d): Photo of the belt-type Hapbeat.

3.2 Modulation method

The coordinate systems used for the proposed method appear in Fig. 2 and 3, while the modulation equation appears in Eq. 1, and a specific example of it appears in Fig. 4.

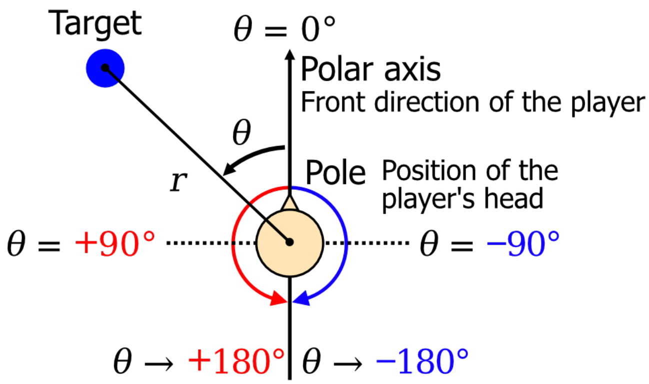

Figure 2: Description of polar coordinates in our proposal. Quoted from our previous study [44].

$$G(\theta,h) = CA_{\mathrm{T,\,B}}(h)A_{\mathrm{L,\,R}}(\theta) \quad (0 \leq G(\theta,h) \leq 1) \tag{1}$$
$$A_{\mathrm{T}}(h) = 0.5(1+h) \ (-1 \leq h \leq 1)$$
$$A_{\mathrm{B}}(h) = 0.5(1-h) \ (-1 \leq h \leq 1)$$
$$A_{\mathrm{L}}(\theta) = \begin{cases} -\theta/90^\circ - 1 & (-180^\circ \leq \theta \leq -90^\circ) \\ \theta/90^\circ + 1 & (-90^\circ \leq \theta \leq 90^\circ) \\ -\theta/90^\circ + 3 & (90^\circ \leq \theta \leq 180^\circ) \end{cases}$$
$$A_{\mathrm{R}}(\theta) = \begin{cases} \theta/90^\circ + 3 & (-180^\circ \leq \theta \leq -90^\circ) \\ -\theta/90^\circ + 1 & (-90^\circ \leq \theta \leq 90^\circ) \\ \theta/90^\circ - 1 & (90^\circ \leq \theta \leq 180^\circ) \end{cases}$$

where h and θ are variables in Fig. 3, θ is the azimuth angle from the player to the target, and h is the target height (y-coordinate) normalized by radius. C is an arbitrary value that determines vibration magnitude, which should be adjusted as $G(\theta,h) < 1$ in any case. Each subscript of functions A represents L: Left; R: Right; T: Top; B: Bottom. The four gain values (TL, TR, BL, BR) to be input to the top-bottom left-right actuators are calculated accordingly. In this paper, $G_{\mathrm{TL,\,TR}}$ values are input to the left and right motors of the necklace-type Hapbeat, while $G_{\mathrm{BL,\,BR}}$ values are input to that of the belt-type Hapbeat.

For direction presentation, the left-right amplitudes are modulated based on the azimuth angle from the player to the target (θ in Fig.3). This modulation method resembles the approach adopted in our previous study, [44], differing only in that it does not distinguish between cases where the target is in front of or behind the player. As such, the player cannot discriminate between front and back using only the vibration stimulus. However, this should not be considered a problem because the player can see the target in front of them. For height presentation, the amplitude ratio of the upper (necklace-type) and lower (belt-type) body area is modulated based on the target's height (y-coordinate). These specifications indicate that the player can locate the target by turning their head in the direction from which they strongly perceive the vibration stimulus.

4 Evaluation

To test the hypotheses presented in Section 1, we conducted an experiment in which participants were tasked with locating and shooting targets in 3D space. The experiment employed four guidance conditions and two target movement conditions.

4.1 Participants

Twelve participants ([Male:Female] = [9:3], [20s:30s:40s] = [9:2:1]) took part in the experiment. All participants were healthy, with no abnormalities in their sensory modalities. Before taking part, each participant signed a consent form based on human research ethics. The per-person duration of the experiment was approximately one hour, and none of the participants received any payment.

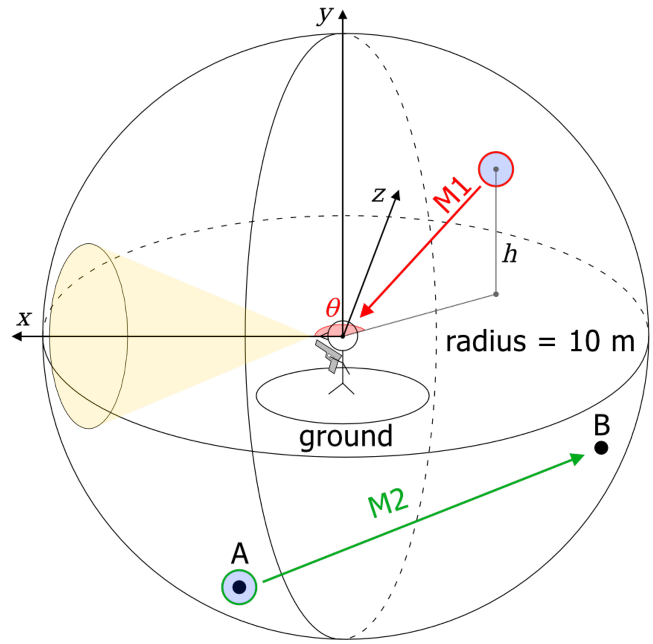

Figure 3: Coordinate system and experimental environment of the proposed method. Details of the polar coordinate system appear in Fig. 2. The red and green arrows show one example of the M1 and M2 movement described in Section 4.5.

4.2 Apparatus

The application used in this experiment was created using the game engine Unity (version 2020.3.22f1) and executed on a gaming PC (CPU: Intel(R) Core(TM) i7-9700K, RAM: 16 GB, GPU: NVIDIA GeForce RTX 2070 SUPER). A VIVE Pro Eye (HTC Corporation), which is capable of eye tracking, was used for the HMD. Small, inexpensive headphones (Audio-Technica Corp., ATH-S100, 32 Ω impedance) were used for audio playback. The necklace-type and belt-type Hapbeat described in Section 3.1 were used to stimulate

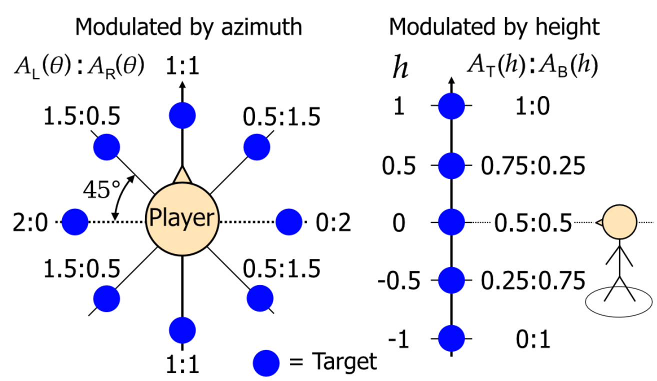

Figure 4: Specific examples of modulation by Eq.1. The left shows direction-based modulation, while the right shows height-based modulation.

modulated musical vibrations.

4.3 Stimuli

The music track “Creation the State of Art,” one of the songs used in the VR synesthesia shooting game Rez Infinite [9], was used throughout the experiment. The playback interval ran from the beginning of the track (0:00) to the end of the experiment (maximum 4:35). Given that this interval contains continuous bass beats, Hapbeat constantly stimulated modulated musical vibrations for participants.

The researcher adjusted the volume of the audio and vibration stimuli such that participants could sufficiently perceive and enjoy listening to the music. The volume was the same for all participants. The following measurement values, using a sinusoidal signal of 80 Hz and -3 dBFS, provide information enabling the reproduction of these aspects. For the audio volume, the input to the headphones (impedance 32 Ω) was 93.3 mV_{RMS}. For the vibration volume, the voltage applied to the motor's terminal was 1.95 V_{RMS} when the gain value in Eq. 1 was 0.5. As a broad indication of the transmitted vibration's intensity, the researcher wore the two Hapbeat devices and measured the vibration intensity by attaching an accelerometer (NXP Semiconductors, MMA7361LC) to the coupling. The measured 3-axis composite values were 60 m/s^2 for the necklace-type device and 64 m/s^2 for the belt-type device.

4.4 Virtual Environment

The experiment's virtual environment appears in Fig. 3. The HMD tracked each participant's head position and rotation, enabling their front direction in the virtual environment to be synchronized with the real position. The experimental task involved using a virtual handgun to locate and shoot a target moving in a uniform linear motion. To aid the player in this task, the handgun emitted a laser beam as an extension of the muzzle. When the participant directed the muzzle toward the target (a blue sphere 50 cm in radius), a red laser dot would appear on the target. The participant could shoot the target by pressing the trigger button of the controller with their index finger. The laser beam was always enlarged by a cylinder collider with a radius of 1 m and an axis corresponding to the muzzle's direction. This enabled the participant to shoot the target by directing the muzzle in its approximate direction. The only shooting feedback was the controller's vibration, with no sound effects played.

During the experiment, the participant's gaze (vertical and horizontal coordinates), head posture (HMD position and rotation), and elapsed time (s) were recorded. The Tobii G2OM (Gaze-2-Object-Mapping) library [37], which is bundled with the Tobii XR SDK provided by Tobii, was used to obtain gaze coordinates. These data were recorded at approximately 33 ms intervals and at the end of each shooting trial, as described in Section 4.5.

4.5 Conditions

The experiment applied four guidance conditions and two target movement conditions, producing a total of eight condition combinations. For each condition, 55 target shooting trials were conducted, and 50 were recorded, with the first five trials used as warm-ups. For each trial, the time limit was 5 s, during which the target would move in a uniform linear motion (1.6 m/s). After the participant shot the target (success) or the 5 s elapsed (fail), the target for that trial would disappear, and the next trial would begin immediately with the appearance of the next target. The targets appeared randomly from the coordinates on the sphere presented in Fig. 3, with the nth and $(n+1)$th coordinates at least 10 m apart. This meant that the $(n+1)$th target was as far beyond the participant's view as possible immediately after the nth trial. To prevent different appearance orders for different participants from affecting the results, the random appearance positions were saved, and all participants conducted the trials with the same appearance order for each condition.

4.5.1 Guidance conditions

We conducted the following four guidance conditions for hypothesis testing in [H1]–[H4].

- None: a condition with no guidance.
- Visual: a condition in which the target is presented visually. Fig. 1(c) presents the 3D radar used. According to studies by Bork *et al.* [4] and Gruenfeld *et al.* [13], this 3D radar makes it easy for participants to understand the task at hand and requires only a small display area.
- Haptic: a condition of the method proposed in Section 3. This condition includes no visualization, meaning that the UI is the same as in None (Fig. 1(b)).
- VisHap: a condition that combines the Visual and Haptic conditions as described in previous subsections. In this condition, participants can explore the target both visually and tactilely.

The comparison between None and Haptic verifies [H1], the comparison between Haptic and Vision verifies [H2], the comparison between Vision and VisHap verifies [H3], and the comparison between the presence of the musical vibration condition (Haptic or VisHap) and the absence of the musical vibration condition (None or Vision) verifies [H4]. To establish the experimental order, we considered order effects for Vision and Haptic and divided the participants into two groups: one would complete the experiment in the order Vision, Haptic, VisHap, None, and the other would complete the experiment in the order Haptic, Vision, VisHap, None. VisHap was positioned third for both groups because participants that experienced VisHap before Haptic might spend more time attempting to decipher the radar (due to the novelty of the haptic condition), thereby potentially not validating [H3] correctly. Given that None was assumed to be the most difficult condition, it was positioned last to enable the participants to be well-trained at the task before finally completing it without assistance.

4.5.2 Target movement conditions

Based on a study by Grunefeld *et al.* [13], the experiment was conducted under the following two conditions to evaluate the effect of target movement:

- M1: a condition for moving from the appearance position to the player's head (Fig. 3 red arrow). The appearance point is an arbitrary point on a sphere with a radius of 10 m.
- M2: a condition for moving from the appearance position (point A) to another position (point B) (Fig. 3 green arrow). Points A and B were randomly selected points on a sphere with a radius of 10 m, but the distance between them exceeded 8 m.

Considering the order effect while emphasizing the comparison between guidance conditions with the same movement, we divided the participants into two groups: a group that first performed the four guidance conditions under M1 and then performed the same experiments under M2 and a group that performed the four guidance conditions under M2 and then performed the same experiments under M1.

4.6 Questionnaire

We conducted a questionnaire survey to evaluate the music-listening experience and the subjective impact of the guidance conditions on finding the target. The five questions listed below were asked using a 7-point Likert scale with the following explanations for each value: 0—strongly disagree, 1—disagree, 2—slightly disagree, 3—neither/nor, 4—slightly agree, 5—agree, 6—strongly agree.

- Q1 This music evokes the sensation of wanting to move some part of my body.
- Q2 Listening to this music gives me pleasure.
- Q3 I can find the target easily.
- Q4 I can find the target intuitively.
- Q5 I enjoyed the experience.

Note that Q1 and Q2 were adopted from a study by Senn *et al.* [34] that assessed the suitability of questions used to evaluate players' musical groove.

4.7 Procedure

First, the researcher explained the shooting task and described the guidance with Haptic and Vision. Participants then stood wearing the HMD and the two Hapbeat devices before practicing the task under each guidance condition. As part of a tutorial, all targets appeared under the same experimental conditions (but at a different appearance point) until participants understood them. Upon completing the tutorial, participants were given time to remove their HMD and take a break if necessary. Then, participants again wore the HMD and calibrated the eye tracking of VIVE Pro Eye. The HMD was not removed after calibration, instead remaining in place until the end of the experiment. Next, the researcher started the experiment in the order described in Section 4.5. The participants navigated the experiment by completing it under one condition, answering the questionnaire, and then moving on to the next condition, and so on. After completing the experiment under all conditions, the researcher interviewed the participants about their questionnaire responses. During the interviews, the researcher attempted to elicit participants' thoughts without trying to guide them, thus enabling them to modify their scores at their own will.

4.8 Result

The results of the experimental behavior logs appear in Fig. 5. Success rate indicates the percentage of trials (out of 50 trials) in which the participant shot the target within 5 s. Clear time indicates the mean time spent shooting the target in successful trials. Rotation amount is the sum of the participant's head rotations (sum of roll, pitch, and yaw) during the experiment. Although the appearance position of the target in each condition differed, which meant the minimum rotation required to locate the target also differed (in the range of 142–165 rad), we have ignored this difference because, at its maximum, it is 23 rad, meaning the effect on the results is unlikely to be substantial.

The eye gaze data, including within-participant comparisons between Vision and VisHap, appear in Fig. 9. Gaze at radar area indicates the percentage of time the participants spent looking at the area of the 3D radar (the red border in Fig. 1(c)) while they were seeking a target (when the target's figure projected on the HMD was smaller than the hemisphere). The results of the questionnaire appear in Fig. 6.

4.8.1 Statistical hypothesis testing

Statistical hypothesis tests were conducted on datasets of within-participant differences to test hypotheses [H1]–[H4]. Within-participant differences were calculated by subtracting each participant's result according to the combination of conditions. For example, the Haptic-None case subtracts the result for None from that for Haptic, such that a positive difference indicates that the result for Haptic exceeded that for None. Then, based on the null hypothesis of "no difference in representative values between the two groups" at the significance level $\alpha = 0.05$, a Wilcoxon signed-rank test was conducted for each pair. The within-participant differences and test

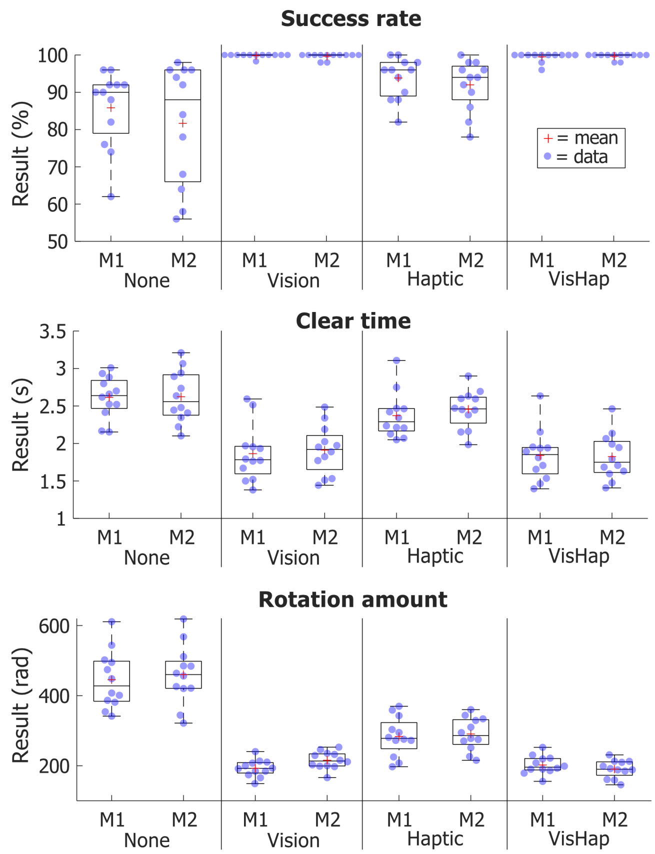

Figure 5: Quantitative results of behavior log. Blue circles represent the mean value for each participant. The red plus signs represent the mean, while the rest is equivalent to a general box-and-whisker plot. These are all the same for Fig. 6–12.

results appear in Fig. 7–12. Note that the P values in these figures do not consider a familywise error. In the following, we define families of tests when multiple null hypotheses must be rejected simultaneously to verify the hypothesis, referring to the guidelines proposed by Veazie *et al.* [41]. MATLAB R2022a was used as the statistical processing software.

In testing [H1] (Haptic vs. None) and [H2] (Haptic vs. Vision), we targeted the ten items shown in Fig. 7 and 8. In this case, because it was not necessary for all items to show a significant difference simultaneously, we did not adjust familywise errors. Instead, to check for the possibility of false positives when testing ten items simultaneously, we calculated q-values using the Benjamini-Hochberg (BH) method with the false discovery rate (FDR) = 0.05 and added in the figures. In testing [H3], we targeted only "Gaze at radar area", as Fig. 9 shows. Because it was not necessary for the results to show a significant difference simultaneously, we did not adjust familywise errors.

In testing [H4], as Fig. 11 shows, we focused more on the answers to the questionnaire items that asked about music experience (Q1 or Q2) and PX (Q5) between guidance conditions with and without musical vibration: Haptic-None (H-N in Table 1), Haptic-Vision (H-V), VisHap-None (VH-N), and VisHap-Vision (VH-V). To verify [H4], the within-participant differences for Q1 or Q2 and Q5 had to demonstrate significance simultaneously. As Table 1 shows, we defined four families for each pair of guidance conditions, one for M1 and one for M2, for a family comprising the two null hypotheses Q1 and Q5 and the two null hypotheses Q2 and Q5. For each family,

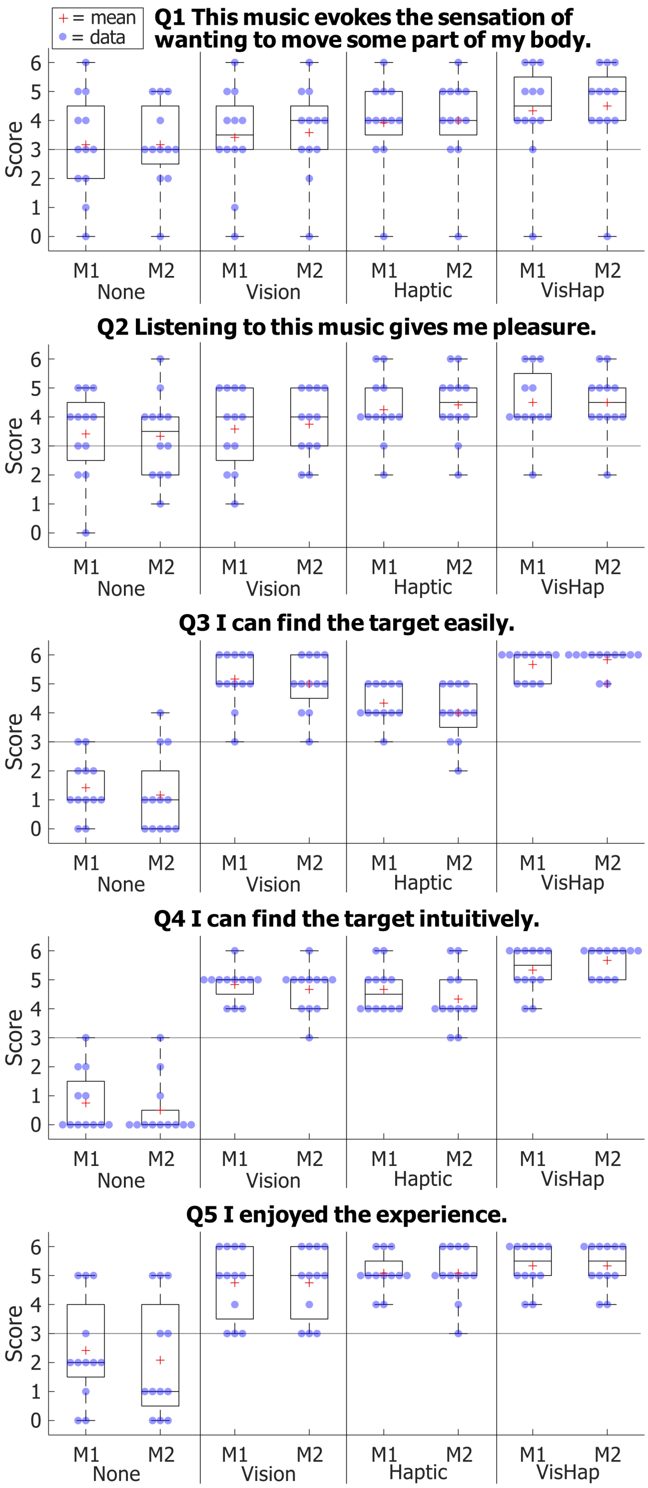

Figure 6: Scores of questionnaires.

Table 1: Simultaneous test results in H4

	H-N	H-V	VH-N	VH-V
M1-Q1∧Q5	*n.s.*	*n.s.*	∗	*n.s.*
M1-Q2∧Q5	∗	*n.s.*	∗	*n.s.*
M2-Q1∧Q5	∗	*n.s.*	∗	*n.s.*
M2-Q2∧Q5	∗	*n.s.*	∗	*n.s.*

we adjusted the P-value using the Holm method [1, 18]. In Table 1, we show the cases where two null hypotheses in a family are rejected simultaneously as ∗ and those not rejected as *n.s.*.

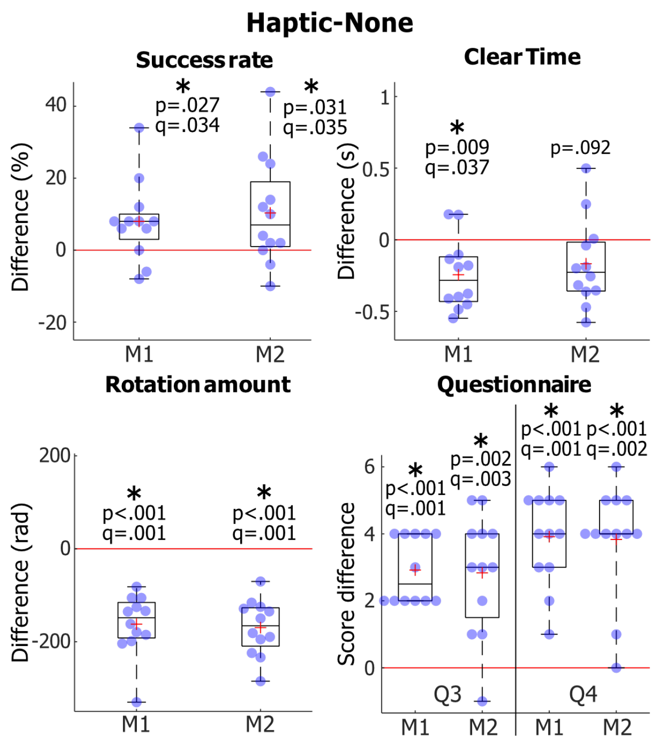

Figure 7: Within-participant differences in Haptic-None. A positive value indicates Haptic > None. Mainly target at verifying [H1]. The asterisk above the p-value indicates a significant difference. Please refer to section 4.8.1 when calculating p- and q-values.

5 Discussion

This study's experimental results demonstrate that modulated musical vibration stimulation using the proposed method helped to guide players toward a moving target in a VR 3D space. Although it is less comprehensible than the 3D radar, combining the two methods has no adverse effects on guidance and improves the music-listening experience. The following subsections discuss the experimental findings in terms of the hypotheses presented in Section 1, the findings from the experiments, and the limitation of this study.

5.1 H1: Haptic vs None

We consider [H1] to be supported. Fig. 7 shows higher success rates, shorter clear times, and less head rotation under Haptic than None. The difference in the amount of rotation is particularly striking, reflecting the fact that, in Haptic, participants likely sought the target by following the guidance provided by vibratory stimulation, whereas, in None, they sought the target by randomly moving their head. This trend was also observed in the questionnaire results, with all but one person in rating Haptic higher in response to Q3

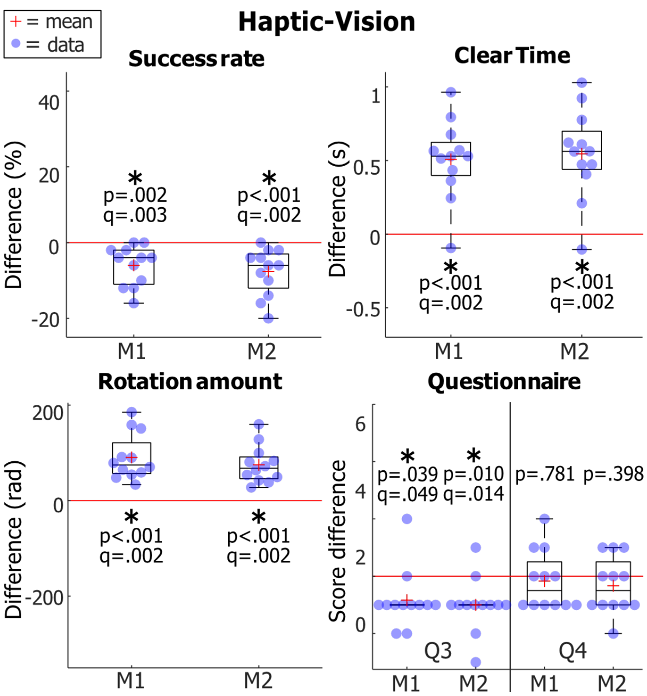

Figure 8: Within-participant comparison in Haptic-Vision. A positive value indicates Haptic > Vision. Mainly target at verifying [H2].

and Q4. The hypothesis test results show significant differences in nine items, with M2 clear time representing an exception. The q-value derived from the BH method was below an FDR = 0.05 for the same items, indicating that these nine items are likely to show significant differences simultaneously. These results also suggest that the proposed method helped users find the target.

5.2 H2: Haptic vs Vision

We consider [H2] to not be supported. Fig. 8 shows lower success rates, longer clear times, and more head rotation under Haptic than Vision. Meanwhile, the questionnaire results in Fig. 8 show that ten participants rated Haptic lower than Vision in response to Q3. The hypothesis test results demonstrate significant differences for eight items, with the Q4-related items representing exceptions. The q-value was below FDR = 0.05 for the same items, indicating that these eight items are likely to show significant differences simultaneously. Hence, we conclude that the proposed method is inferior to the 3D radar in terms of comprehensibility.

However, a within-participant comparison of Q4, which asked whether the guidance was intuitive, showed no significant difference. Therefore, the required cognitive resources may be comparable between the proposed method and the 3D radar approach.

5.3 H3: VisHap vs Vision

We consider [H3] to be supported. Fig. 9 shows that the gaze at the radar area was reduced in VisHap compared to the gaze in Vision in the case of M2. Therefore, in the case of relatively complex movement patterns, as represented by M2, combining the proposed method and the 3D radar reduces the time spent gazing at the radar.

The following paragraphs discuss the findings from the comparison between VisHap and Vision. The questionnaire results suggest that all but one participant considered VisHap to be either just as helpful or more helpful for finding the target than Vision for both M1 and M2 based on the results representing responses to Q3 and Q4 presented Fig. 6. Therefore, it seems unnecessary to consider

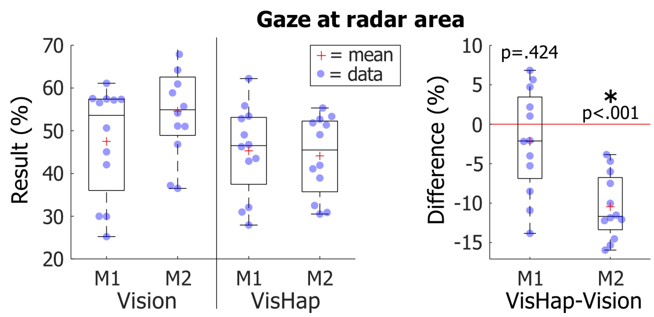

Figure 9: Results of gaze at radar area. The right graph shows a within-participant comparison in VisHap-Vision. Mainly target at verifying [H3].

the possibility of the proposed method interfering with participant localization. When discussing VisHap in the interview, three participants commented that "[The] vibration[s] helped me when it was difficult to find [the target] with [the] 3D radar, especially when the target [was] extremely up and down." Moreover, two other participants noted that "I roughly grasp[ed] the location with vibration[s] and confirm[ed] the exact location with [the] 3D radar." These observations indicate that modulated musical vibrations contribute to target exploration. This is consistent with the meta-analysis by Prewett *et al.* [30] that reported that conventional vibrotactile cues contributed to improved task performance, indicating the utility of using modulated musical vibrations as a tactile cue.

However, Fig. 10 shows no significant difference in the behavior logs except regarding the amount of rotation under M2, an effect size that is small compared to the effect sizes observed for Haptic-None and Haptic-Vision. It is quite possible that the 3D radar is a powerful 3D guidance method, with the guidance effect of Haptic potentially being negligible, and in fact, two participants commented that they closely monitored the 3D radar and did not pay attention to the vibration stimuli. However, the experimental environment and task were very simple, such that participants had no difficulty performing the task even if they were gazing at the 3D radar. This could be another reason for the slight difference between VisHap and Vision. Therefore, evaluating a task that longer 3D radar gazing time negatively impacts task accomplishment can deepen our knowledge of the interaction between haptic and visual in 3D guidance.

5.4 H4: PX and music-listening experience

We consider [H4] to have not been properly tested. From Table 1, the combination of Haptic-None and VisHap-None improved both the music-listening experience and PX, while Haptic-Vision and VisHap-Vision did not. If [H4] is correct, it should hold for all comparisons between the conditions with and without musical vibrations. Fig. 6 shows that the scores for Q5 under None were negative, whereas the score under Vision was positive compared to Haptic and VisHap. Therefore, it is likely that their evaluation of the PX focused not on the music-listening experience but on the ease of finding the target, meaning that they focused on whether they could accomplish the shooting task smoothly. In such a case, it is quite possible that using a more easily comprehensible signal instead of music—such as a sine wave—would improve the Q5 evaluation, indicating that the result cannot definitively prove our claim that improving the music-listening experience contributes to the PX.

Focusing only on evaluating the music-listening experience, Fig. 11 demonstrates that the Q1 and Q2 scores for the condition with musical vibration were higher overall but that no significant differences were observed in Haptic-Vision except for Q2 of M2. The task's low difficulty level (i.e., the ease of finding the target) likely affected

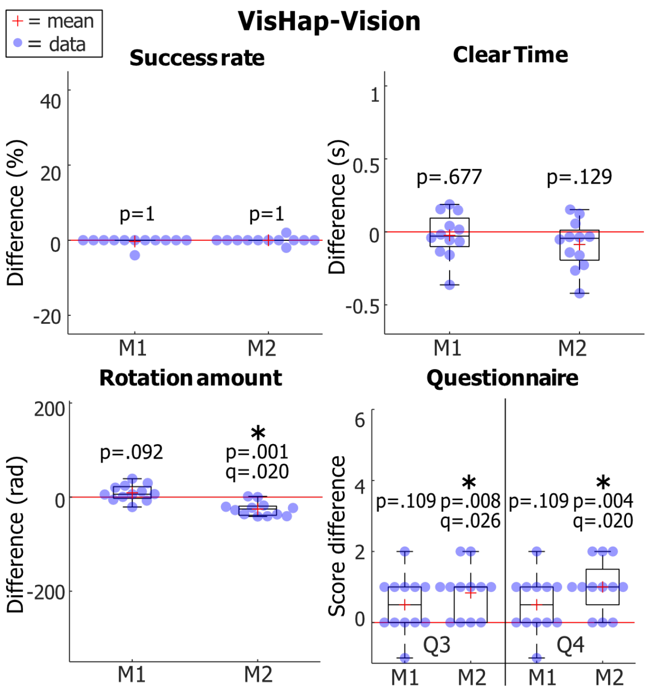

Figure 10: Within-participant comparison in VisHap-Vision. A positive value indicates VisHap > Vision.

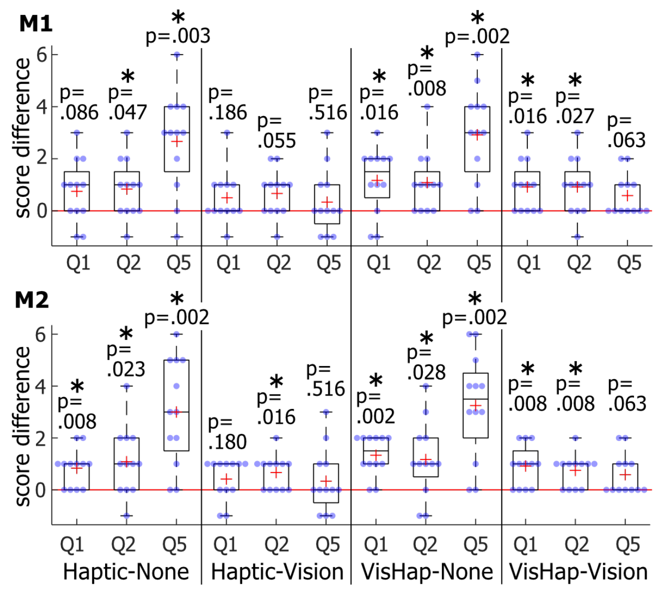

Figure 11: Within-participant comparison of Q1, Q2, and Q5. The top graph shows the M1 case, while the bottom graph shows the M2 case. Mainly target at verifying [H4].

evaluations of the listening experience. During the interview, nine participants observed that "concentrating on the task distract[ed] my mind from the music. [It was] hard to concentrate on both." This suggests that cognitive saturation likely occurred, resulting in the same tendency as that observed in Roger *et al.*'s study [32,33]. Thus, it cannot be denied that the task's low difficulty level influenced the results of Q1 and Q2 under Haptic-None. However, given the scores for Q1 and Q2 under VisHap were significantly higher than those under Vision with similar task difficulty (Fig. 10), it can be said with some confidence that modulated musical vibrations contribute to an improved listening experience.

To evaluate [H4] in future studies, Roger *et al.*'s exploratory case study [32] could serve as a useful reference. They asked 12 participants to play a VR game under two conditions; the first condition did not modulate game music; the second condition modulated the game music's speed and volume according to the events in the game (adaptive music condition). Most participants reported in interviews that they paid more attention to game music in the adaptive music condition, suggesting that music linked to game events may influence PX in VR games. Therefore, redesigning the task to strengthen the relationship between game music, experimental environment, and task score is necessary to adequately investigate [H4]. This might include designing the target's movement and task score relative to the rhythm of the game music.

5.5 M1 vs M2: the effect of target movement

We contend that the combination of Vision and Haptic will make guidance somewhat easier if a target moves in such a way that it changes direction, as in the case of M2. Fig. 12 shows that M2 led to significantly greater head rotation than M1 in Vision, whereas M1 involved significantly greater head rotation than M2 in VisHap, suggesting that the 3D radar often misleads the participant. In addition, participants spent significantly longer gazing at the radar area under M2 than under M1 in Vision, suggesting that it took longer to see the radar and locate the target. Therefore, in the case of M2, the 3D radar is less comprehensible, suggesting that the proposed method may have supplemented the 3D radar.

Nonetheless, Gruenefeld *et al.* [13] reported that participants had a slightly (not significant) better grasp of the trajectory of a moving object on the 3D radar in M2 than in M1, contradicting this paper's findings. One possible reason for this is that although participants were stationary in that study, our participants were required to move their heads quickly, which caused the 3D radar to rotate quickly, making it difficult to capture the change in the azimuth angle due to the M2 condition.

5.6 Limitations

Regarding the comparison between VisHap and Vision, the 3D radar's appearance, display position, and size should be assessed. Three participants explicitly stated that the radar served only as a distraction and suggested improvements regarding the color of the target on the radar and the display position of the radar. Although task evaluation and the cognitive load caused by visualization methods for out-of-view objects have been evaluated [17], few studies have considered sensory evaluation. This calls for further research. Furthermore, 3D radar may not be appropriate in conjunction with haptic guidance. For example, Jo *et al.* [21] proposed the AroundPlot method, which places out-of-view objects in the corners of the visual field. Although, on its own, this approach provides less information than a 3D radar does, combining AroundPlot with haptic guidance may well present sufficient information to the user while reducing the cognitive load. For example, Lin *et al.* [25] proposed a method in which visual and haptic feedback are combined in robot teleoperation. Therefore, it is worth considering the possibility of evaluating visualization methods on the premise of using haptic feedback.

Notably, the evaluation of the PX was too simplistic due to the experiment primarily focusing on whether the proposed method achieved both 3D guidance and an improved music-listening experience. Although various methods for evaluating PX have been proposed [8, 20, 36, 40], conducting this kind of evaluation for all of this paper's conditions was not practical due to the burden on participants and the impact on other evaluation items. Therefore, future research should limit the experimental conditions to, for ex-

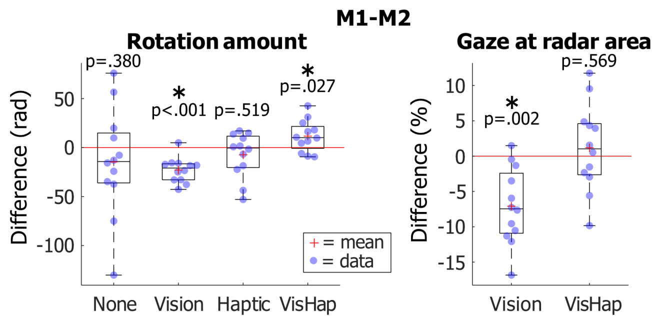

Figure 12: Within-participant comparisons of the behavior log between the target movement conditions. Only items demonstrating significant differences have been excerpted, with neither success rate, clear time, nor any questionnaire item exhibiting any significant difference under any guidance conditions.

ample, VisHap-Vision, conduct a detailed evaluation of the PX, and redesign the experimental task as described in Section 5.4.

6 Conclusion

This paper aimed to enhance the PX of VR shooting games by achieving 3D guidance and improving the music-listening experience. For the first step, we proposed and evaluated a haptic 3D guidance method that involved stimulating musical vibrations that are modulated based on the target's position using two haptic devices: a necklace-type Hapbeat and a belt-type Hapbeat. The experimental results suggest that the proposed method can guide players toward the location of a moving target in 3D space using only tactile stimuli. Modulated musical vibrations were also shown to enhance the music-listening experience during the shooting task. Although the proposed method is less comprehensible than a 3D radar, combining the two approaches could improve guidance comprehension and enhance the PX.

However, the simplicity of the experimental environment and task prevented us from properly evaluating the causal relationship between the music-listening experience and PX or the contribution of tactile stimuli to comprehension when the 3D radar and modulated musical vibrations were presented simultaneously. Therefore, we would like to redesign the experiment to address these concerns properly.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers JP17H01774, JP20H04220.

References

[1] M. Aickin and H. Gensler. Adjusting for multiple testing when reporting research results: the bonferroni vs holm methods. *American journal of public health*, 86(5):726–728, 1996.
[2] BEAT GAMES. Beat Saber - VR rhythm game. https://beatsaber.com/. (Accessed Oct. 13, 2022).
[3] N. Binetti, L. Wu, S. Chen, E. Kruijff, S. Julier, and D. P. Brumby. Using visual and auditory cues to locate out-of-view objects in head-mounted augmented reality. *Displays*, 69:102032, 2021.
[4] F. Bork, C. Schnelzer, U. Eck, and N. Navab. Towards efficient visual guidance in limited field-of-view head-mounted displays. *IEEE transactions on visualization and computer graphics*, 24(11):2983–2992, 2018.
[5] M. Carroll and C. Yildirim. The effect of body-based haptic feedback on player experience during vr gaming. In *International Conference on Human-Computer Interaction*, pp. 163–171. Springer, 2021.
[6] G. Cassidy and R. A. MacDonald. The effects of music on time perception and performance of a driving game. *Scandinavian journal of psychology*, 51(6):455–464, 2010.
[7] V. A. de Jesus Oliveira, L. Brayda, L. Nedel, and A. Maciel. Designing a vibrotactile head-mounted display for spatial awareness in 3d spaces. *IEEE transactions on visualization and computer graphics*, 23(4):1409–1417, 2017.
[8] A. Drachen, L. E. Nacke, G. Yannakakis, and A. L. Pedersen. Correlation between heart rate, electrodermal activity and player experience in first-person shooter games. In *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*, pp. 49–54, 2010.
[9] Enhance Experience Inc. Rez Infinite. https://rezinfinite.com/. (Accessed Oct. 13, 2022).
[10] A. Fiannaca, T. Morelli, and E. Folmer. Haptic target acquisition to enable spatial gestures in nonvisual displays. In *Proceedings of graphics interface 2013*, pp. 213–219. 2013.
[11] U. Gruenefeld, A. E. Ali, S. Boll, and W. Heuten. Beyond halo and wedge: Visualizing out-of-view objects on head-mounted virtual and augmented reality devices. In *Proceedings of the 20th international conference on human-computer interaction with mobile devices and services*, pp. 1–11, 2018.
[12] U. Gruenefeld, D. Ennenga, A. E. Ali, W. Heuten, and S. Boll. Eyesee360: Designing a visualization technique for out-of-view objects in head-mounted augmented reality. In *Proceedings of the 5th symposium on spatial user interaction*, pp. 109–118, 2017.
[13] U. Gruenefeld, I. Koethe, D. Lange, S. Weiß, and W. Heuten. Comparing techniques for visualizing moving out-of-view objects in head-mounted virtual reality. In *2019 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 742–746. IEEE, 2019.
[14] U. Gruenefeld, D. Lange, L. Hammer, S. Boll, and W. Heuten. Flyingarrow: pointing towards out-of-view objects on augmented reality devices. In *Proceedings of the 7th ACM international symposium on pervasive displays*, pp. 1–6, 2018.
[15] S. Günther, F. Müller, M. Funk, J. Kirchner, N. Dezfuli, and M. Mühlhäuser. Tactileglove: Assistive spatial guidance in 3d space through vibrotactile navigation. In *Proceedings of the 11th pervasive technologies related to assistive environments conference*, pp. 273–280, 2018.
[16] Hapbeat LLC. Hapbeat. https://hapbeat.com/. (Accessed Oct. 13, 2022).
[17] Y. Harada and J. Ohyama. Quantitative evaluation of visual guidance effects for 360-degree directions. *Virtual Reality*, 26(2):759–770, 2022.
[18] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, pp. 65–70, 1979.
[19] M. J. Hove, S. A. Martinez, and J. Stupacher. Feel the bass: Music presented to tactile and auditory modalities increases aesthetic appreciation and body movement. *Journal of Experimental Psychology: General*, 149(6):1137, 2020.
[20] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, 66(9):641–661, 2008.
[21] H. Jo, S. Hwang, H. Park, and J.-h. Ryu. Aroundplot: Focus+ context interface for off-screen objects in 3d environments. *Computers & Graphics*, 35(4):841–853, 2011.
[22] O. B. Kaul and M. Rohs. 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, 2017.
[23] C. Klimmt, D. Possler, N. May, H. Auge, L. Wanjek, and A.-L. Wolf. Effects of soundtrack music on the video game experience. *Media Psychology*, 22(5):689–713, 2019.
[24] L. Levy, R. Solomon, M. Gandy, and R. Catrambone. The rhythm's going to get you: Music's effects on gameplay and experience. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, pp. 607–612, 2015.
[25] T.-C. Lin, A. U. Krishnan, and Z. Li. Comparison of haptic and augmented reality visual cues for assisting tele-manipulation. In *2022 International Conference on Robotics and Automation (ICRA)*, pp. 9309–9316. IEEE, 2022.

[26] A. Marquardt, C. Trepkowski, T. D. Eibich, J. Maiero, E. Kruijff, and J. Schöning. Comparing non-visual and visual guidance methods for narrow field of view augmented reality displays. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3389–3401, 2020.

[27] S. Merchel and M. E. Altinsoy. Auditory-tactile experience of music. In *Musical Haptics*, pp. 123–148. Springer, Cham, 2018.

[28] F. Nakamura, A. Verhulst, K. Sakurada, and M. Sugimoto. Virtual whiskers: Spatial directional guidance using cheek haptic stimulation in a virtual environment. In *Augmented Humans Conference 2021*, pp. 141–151, 2021.

[29] C. Plut and P. Pasquier. Music matters: An empirical study on the effects of adaptive music on experienced and perceived player affect. In *2019 ieee conference on games (cog)*, pp. 1–8. IEEE, 2019.

[30] M. S. Prewett, L. R. Elliott, A. G. Walvoord, and M. D. Coovert. A meta-analysis of vibrotactile and visual information displays for improving task performance. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(1):123–132, 2011.

[31] K. Rogers, M. Jörg, and M. Weber. Effects of background music on risk-taking and general player experience. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pp. 213–224, 2019.

[32] K. Rogers, M. Milo, M. Weber, and L. E. Nacke. The potential disconnect between time perception and immersion: Effects of music on vr player experience. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, pp. 414–426, 2020.

[33] K. Rogers, G. Ribeiro, R. R. Wehbe, M. Weber, and L. E. Nacke. Vanishing importance: studying immersive effects of game audio perception on player experiences in virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018.

[34] O. Senn, T. Bechtold, D. Rose, G. S. Câmara, N. Düvel, R. Jerjen, L. Kilchenmann, F. Hoesl, A. Baldassarre, and E. Alessandri. Experience of groove questionnaire: Instrument development and initial validation. *Music Perception: An Interdisciplinary Journal*, 38(1):46–65, 2020.

[35] SUBPAC INC. Subpac - the new way to experience sound: Feel it. `https://subpac.com/`. (Accessed Oct. 13, 2022).

[36] P. Sweetser and P. Wyeth. Gameflow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)*, 3(3):3–3, 2005.

[37] Tobii. Tobii G2OM. `https://vr.tobii.com/sdk/solutions/tobii-g2om/`. (Accessed Oct. 13, 2022).

[38] C. Trepkowski, A. Marquardt, T. D. Eibich, Y. Shikanai, J. Maiero, K. Kiyokawa, E. Kruijff, J. Schöning, and P. König. Multisensory proximity and transition cues for improving target awareness in narrow field of view augmented reality displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(2):1342–1362, 2021.

[39] H.-R. Tsai, Y.-C. Chang, T.-Y. Wei, C.-A. Tsao, X. C.-y. Koo, H.-C. Wang, and B.-Y. Chen. Guideband: Intuitive 3d multilevel force guidance on a wristband in virtual reality. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.

[40] V. Vanden Abeele, L. E. Nacke, E. D. Mekler, and D. Johnson. Design and preliminary validation of the player experience inventory. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, pp. 335–341, 2016.

[41] P. J. Veazie. When to combine hypotheses and adjust for multiple tests. *Health services research*, 41(3p1):804–818, 2006.

[42] C. Wang, D.-Y. Huang, S.-w. Hsu, C.-E. Hou, Y.-L. Chiu, R.-C. Chang, J.-Y. Lo, and B.-Y. Chen. Masque: Exploring lateral skin stretch feedback on the face with head-mounted displays. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 439–451, 2019.

[43] Woojer USA Inc. Woojer. `https://www.woojer.com/`. (Accessed Oct. 13, 2022).

[44] Y. Yamazaki and S. Hasegawa. Navigation method enhancing music listening experience by stimulating both neck sides with modulated music vibration. Dec. 2022.

[45] Y. Yamazaki, H. Mitake, and S. Hasegawa. Tension-based wearable vibroacoustic device for music appreciation. In *International conference on human haptic sensing and touch enabled computer applications*, pp. 273–283. Springer, 2016.

[46] Y. Yamazaki, H. Mitake, and S. Hasegawa. Implementation of tension-based compact necklace-type haptic device achieving widespread transmission of low-frequency vibrations. *IEEE Transactions on Haptics*, 2022.

[47] J. Zhang and X. Fu. The influence of background music of video games on immersion. *Journal of Psychology & Psychotherapy*, 5(4):1, 2015.