Spherical Layout with Proximity-Based Multimodal Feedback for Eyes-Free Target Acquisition in Virtual Reality

BoYu Gao[1], Yujun Lu[1], HyungSeok Kim[2], Byungmoon Kim[3], and Jinyi Long[1(✉)]

[1] Jinan University, Guangzhou 510632, Guangdong, China
{bygao,jinyil}@jnu.edu.cn, yujunlu@stu2018.jnu.edu.cn
[2] Konkuk University, Seoul 05092, South Korea
hyuskim@konkuk.ac.kr
[3] Adobe Research, San Jose, CA 95110-2704, USA
bmkim@adobe.com

Abstract. Eyes-free interaction can reduce the frequency of headset rotation and speed up the performance via proprioception in Virtual Environments (VEs). However, proprioception cues make it difficult and uncomfortable to select the targets located at further distance. In VEs, proximity-based multimodal feedback has been suggested to provide additional spatial-temporal relation for 3D selection. Thus, in this work, we mainly study how such multimodal feedback can assist eyes-free target acquisition in a spherical layout, where the target size is proportional to the horizontally egocentric distance. This means that targets located at further distance become bigger allowing users to acquire easily. We conducted an experiment to compare the performance of eyes-free target acquisition under four feedback conditions (none, auditory, haptic, bimodal) in the spherical or cylinder layout. Results showed that three types of feedback significantly reduce acquisition errors. In contrast, no significant difference was found between spherical and cylinder layouts on time performance and acquisition accuracy, however, most participants prefer the spherical layout for comfort. The results suggest that the improvement of eyes-free target acquisition can be obtained through proximity-based multimodal feedback in VEs.
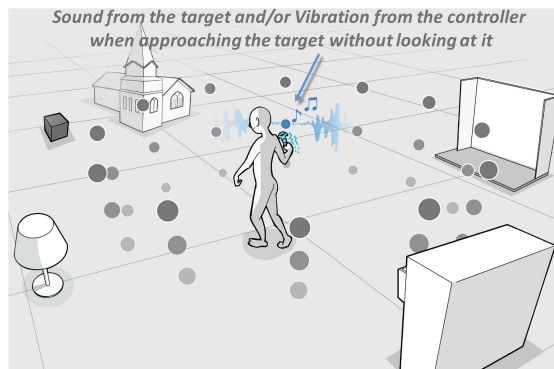
Keywords: Proximity-based multimodal feedback ·
Eyes-free target acquisition · Spatial layout · Virtual Environments

1 Introduction

Recent work has shown the importance and potential of eyes-free target acquisitions in Virtual Reality (VR) [1, 2]. For example, it can effectively reduce the frequency of headset rotation and greatly improve the efficiency of interaction (e.g. painting, blind typing [2]). The mostly used cue to leverage such an ability is the proprioception (*a sense of the relative position of one' own parts of the body*) [1, 3, 4]. For example, Yan et al. [1] found that the users mainly rely on proprioception to quickly acquire the

targets without looking at them in VR. However, this eyes-free interaction with proprioception cues makes difficult and uncomfortable to select the targets located at further distance. Such limitations lead to longer time performance and more efforts for eyes-free target acquisition in VR.

In this work, we focus on alleviating this issue to improve the performance of the eyes-free target acquisition in VR. Commonly, use of additional feedback to assist eyes-free interaction is a promising solution (e.g. Earpod [8]). In VR, the multimodal feedback has been suggested to provide additional temporal or spatial information [6], particularly for 3D selections. For example, Ariza et al. [11] explored the effects of proximity-based multimodal feedback (*the intensity of feedback depends on the spatial-temporal relations between input devices and the virtual target*) on 3D selections in immersive VEs, and found the feedback types significantly affect the selection movement [11]. Thus, we hypothesize that the proximity-based multimodal feedback that provides additional spatial-temporal information could further improve the performance of eyes-free target acquisition in VR. When the user's controller approaches the target, the sound or the vibration is progressively given to inform the movement or the acquisition of target (See Fig. 1).

Fig. 1. This illustrates that the participant can acquire the blue target without looking at it (the participant looks at the dark cube), with the help of additional proximity-based multimodal feedback in VR (the distance between the target and the controller is mapped with the intensity of auditory and/or haptic cues). (Color figure online)

To the best of our knowledge, there is no work focusing on such investigation. We therefore create the proximity-based auditory (pitch) or/and haptic (intensity of vibration) feedback in both the spherical and the cylinder layouts. Both allow more items to display, and the items in the cylinder layout have the same horizontally egocentric distance. In the spherical layout, the item size is proportional to the horizontally egocentric distance, so that the items located at the further become larger to acquire. The items at the higher and lower locations become closer to the users so that they can acquire it easily and comfortably.

To evaluate the proposed approach, we conducted a Mixed ANOVA experimental design to measure the between-subject factor (layouts) and the within-subject factor (feedback conditions) on the eyes-free target acquisition task.

The contributions of this work are as follows:

- This is the first work to explore the proximity-based multimodal feedback (auditory, haptic, auditory and haptic) for the eyes-free target acquisition in VR.
- We designed more comfortable spherical layout and compared the performance of eyes-free target acquisition with the cylinder layout in VEs.
- Based on the results, we suggested that use of the proximity-based multimodal feedback to keep a balance between the trial completion time and the acquisition accuracy for the eyes-free target acquisition in VEs.

2   Related Work

2.1   Spatial Layouts for Target Acquisition

Target acquisition is one of the common tasks in VEs [12]. To improve the efficiency of target acquisition in VEs, it involves many factors, for example, the spatial depth between target and hand [19], the size of target [20], and the perception of the space [21]. Fitts' law [22] is a well-known model to predict selection time performance for a given target distance and size. It is necessary to consider these factors when designing the construction of the spatial layout in VEs. For example, a spatial layout for effective task switching on head-worn displays is called the *personal cockpit* [23], it allows users to quickly access the targets. In addition, Ens et al. [24] proposed a layout manager to leverage spatial constancy to efficiently access the targets. The target distance and size were carefully controlled to improve the efficiency in these experiments.

In particular, the spatial layout itself can provide additional information - the depth variation. For example, Gao et al. [18] proposed the amphitheater layout with *egocentric distance-based item sizing (EDIS),* and found that the small and medium *EDIS* can give efficient target retrieval and recall performance, compared to circular wall layout [17]. Yan et al. [1] explored the target layout (a kind of *personal cockpit, the distance between the item and virtual camera is the same*) for eyes-free target acquisition around the body space, and found that the distance between higher/lower rows and the body make it uncomfortable to acquire the targets. In this work, we propose a spherical layout, so that participants can acquire the targets located at higher and lower rows comfortably and easily. Little work has been investigated the comparison of spherical and circular layouts for eyes-free target acquisition in the VE.

### 2.2   Multimodal Feedback for Target Acquisition

Previous research has shown the importance of multimodal feedback for selection guidance in 2D graphical user interfaces and gestural touch interfaces [7, 9, 10, 25]. However, increasing the quality of the visual feedback does not necessarily improve user performance [14]. In this work, we therefore mainly consider the additional auditory and haptic feedback instead of visual feedback.

Continuous auditory feedback can improve the gestural touch performance via frequency and sound [13, 25]. For example, Gao et al. [25] presented that the gradual continuous auditory feedback contributes to the performance of trajectory-based finger gestures in 2D interfaces. The distance and orientation between the target and the user can be given via the spatial auditory in VEs [6]. In addition, proximity-based multimodal feedback [11], in which the sensory stimuli intensity is matching with the spatial-temporal relationship, can provide better performance of 3D selection in VR. In particular, the binary feedback performed better than continuous feedback in terms of faster movement and higher throughput. However, the eyes-free manner is different from the eyes-engaged in terms of information perception. We believe the eyes-free target acquisition requires more information than 3D selection. We therefore focus on improving the performance of eyes-free target acquisition via continuous proximity-based multimodal feedback in the spatial layouts.

3   Target Layouts with Multimodal Feedback

In this work, we adopted the findings from Yan et al. [1] to construct the target layout. For example, the comfortable distance to acquire the target is 0.65 m for the users. The radius of the target is 0.1 m. We create three rows, each row has 12 spheres (3 × 12).

To allow better comfort when acquiring targets, we propose a spherical layout, which makes the distance between the higher/lower rows of targets and the body's chest closer. The radius is 0.45 m for both rows, and the radius of the middle row is 0.65 m. For the cylinder layout, the radius of the three rows is 0.65 m (See Fig. 2).
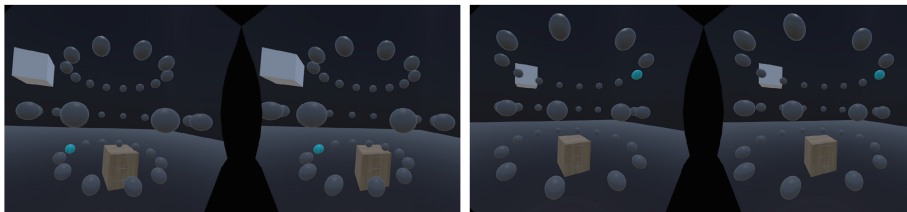
Fig. 2. The items size is proportional to the horizontally egocentric distance in the spherical layout (left), so that the further targets become larger and the user feels comfortable to acquire them, while the items in the cylinder layout have the same horizontally egocentric distance (right). Note that the white cube indicates the direction that the participant looked at when acquiring the target in an eyes-free way. The center of the layout is located at the user's chest.

To create the spatial-temporal relation between the location of target and movement of controller, we predefined the activation area as a sphere with 0.2 m radius for multimodal feedback (Note that the radius of blue target is 0.1 m). Such 0.2 m distance is used to give prior notification of the movement to the user. Park et al. [9] also showed that the preemptive-based continuous auditory feedback gave better performance for 3D hand gestures on circular menu selection. For example, if the controller

gets closer to the center of the target (See Fig. 3: d is smaller than 0.2 m), the pitch of spatial auditory feedback from the target or the intensity of haptic feedback from the controller tends to be the larger to inform the movement. So the participant was informed that the controller is being approached to the target. When the controller intersects with the center of the target, the pitch of sound or intensity of vibration remains the largest.
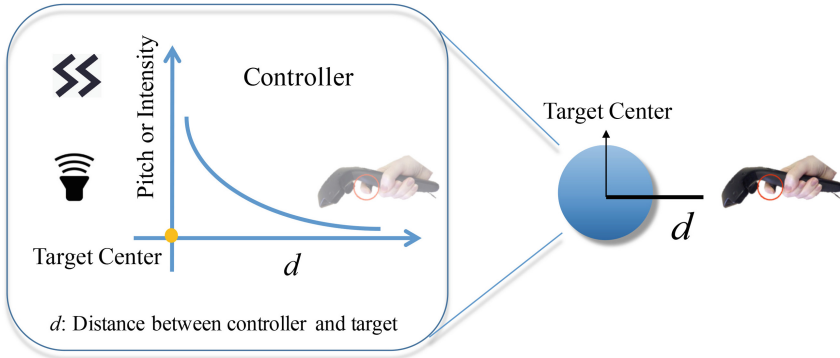
Fig. 3. The relation between the pitch of spatial auditory/the vibration intensity and the spatial distance d (between target and controller). The radius of the blue sphere (target) was 0.1 m. We defined the d as 0.2 m in the experiment. When approaching the target (smaller than 0.2 m), the pitch of sound from the target and the intensity of vibration from the controller is increased. (Color figure online)

4   Experiment

In this experiment, we utilized the 2×4 Mixed ANOVA experimental design, the between-subject factor was the spatial layout (cylinder layout vs. spherical layout), while the within-subject factor was the feedback condition (none, auditory only, haptic only, both auditory and haptic). The hypothesis formulated in this experiment are:

> **H1:** The spherical layout could allow the participants to have better acquisition performance and comfortable experience, compared to the cylinder layout.
> **H2:** The proximity-based multimodal feedback could assist the correct phase during eyes-free target acquisition in VR.

4.1   Participants

The 24 subjects (mean age = 22.2, SD = 2.43 years, the number of female was 12) were recruited from the local campus. All of the participants had normal or corrected-to-normal vision. None of them had experience using VR devices, and they were assigned into two groups randomly. The number of female in each group was balanced. Twenty of them have right hand for dominant and balanced into two groups. Six participants do not play games. The remaining participants play video games for 2 or 3 hour per week.

4.2   Material

The experiment was performed with HTC Vive Pro [26], which allows the user to navigate with lower-latency head tracking, and a refresh rate of 90 Hz. The device has two screens, one per eye, each one having a resolution of 1440 * 1600. The field of view of the device is 110°. The main machine has the capability of Inter core i7-8700, CPU (3.2 GHz), 16 GB RAM, Geforce GTX-1060 graphics card. The platform of development was Unity 2018 with C# language, where we implemented the cylinder and spherical layouts for eyes-free target acquisition in the VE. Target acquisition was implemented with one of the two controllers [27]. The participants were informed to press the trigger on the controller to acquire the target. The target was rendered as the blue. The selection interactions were the same in every condition.

For the assigned target, we defined the spatial distance predefined d as 0.2 m. In the program, the largest pitch value was 1, and the strongest intensity was 4000, the relationship between spatial distance and feedback intensity was illustrated in Fig. 3. As mentioned above, the radius of the circular layout, the horizontally egocentric distance between the target and the participants' main body, was 0.65 m. For the spherical layout, we reduced the radius to 0.45 m for the higher/lower rows, and the radius of middle row was 0.65 m.

4.3   Task Design

Thanks to Yan et al. [1], we adopted the similar task design and procedure with them. However, the goal of this work is to explore the effects of additional multimodal feedback on assisting the performance, it is not necessary to ask participants to acquire every target in the layout. We therefore randomly selected the representative locations (3 targets) from low, medium, and high rows respectively among 36 targets.

For each target, the participants were informed to rotate the body towards the white cube when acquiring it, twelve rotations in total (See Fig. 2). The order of rotations was random. During each rotation, the location of target was the same while the participant changed the orientation. It is important to note that the color of the target was changed to the same with other items once the experiment started, so that the participants could not notice the location of target during body rotation. The participants were asked to grasp the 3D controller with the same grip pose. To ensure the eyes-free approach, the observer was asked to look at the participant and the monitor for each acquisition. The total number of trials in the experiment were 3456 (24 subjects * 3 targets * 12 directions * 4 feedback conditions). The participants were asked to focus on acquiring the specific targets as quickly and accurately as possible with the 3D controller under each of feedback conditions in the assigned layout. The orders of the feedback conditions were randomly assigned to each participant.

4.4   Procedure

The experiment consisted of three phases. In the preparation phase, the participants were asked to fill in the personal form, and informed to get familiar with the VR environment and the interaction method, to acquire the targets several times using the controller until they felt confident.

In the experimental phase, the participant was asked to memorize the spatial location of each target, by means of practicing acquiring it in an eyes-free way several times. Then the participant started the eyes-free target acquisition. After each acquisition, the white cube was automatically rotated to the next direction randomly, which guided the participant to rotate the body towards the cube. Only the duration of acquisition was recorded as the trial completion time. The spatial and angular offset errors were also recorded. After 12 rotations for one target, they were allowed to have a 2-min break. After each of feedback conditions, they were asked to fill in the NASA-TLX questionnaire. In total, each participant finished 144 trials (3 targets * 12 directions * 4 feedback conditions).

After the experiment, the participants were informed to do a subjective interview. The whole experimental phases lasted about one hour. Each participant can obtain 4-dollar payment.
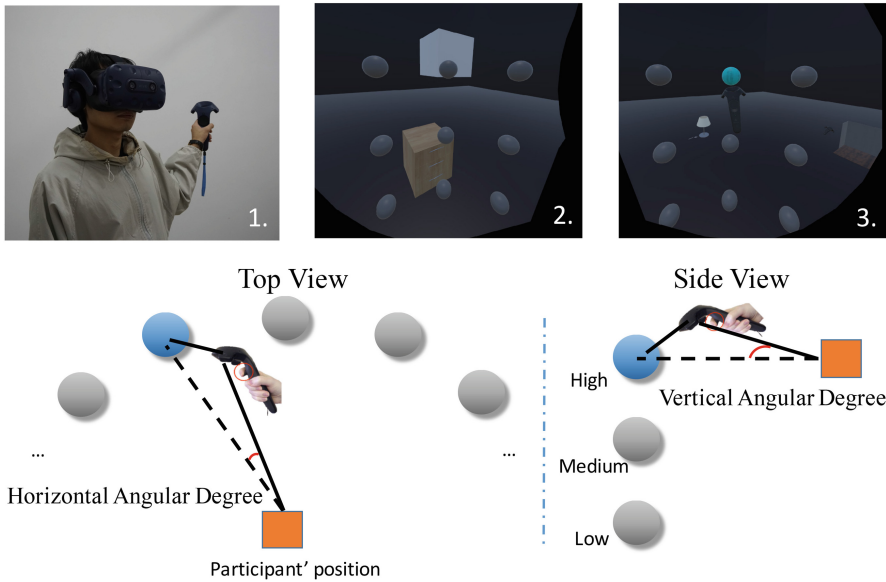
Fig. 4. The participant acquires the target without looking at it. 2: The participant looks at the white cube while acquiring the target. 3: The participant's hand approaches the target (the blue ball) using the controller. The bottom images illustrate the horizontal (from top view) and vertical (from side view) angular offset degree between controller and the target (bottom). (Color figure online)

4.5   Metrics

We utilized the following metrics, *trial completion time*, *spatial offset* [1], *angular offset* [1], and *subjective questionnaire* (NASA-TLX [5]), to measure the eyes-free target acquisition performance under the different feedback conditions. We defined the duration from the starting of acquiring each target to the acquisition confirmation as a

trial completion time. Spatial offset was defined as the Euclidean distance between the acquisition point and the target's actual position, *Angular offset* included horizontal and vertical axis in degrees (See Fig. 4). It indicated the directions of the acquisition points were shifted from the actual location. For example, if the acquisition point was the upwards, as shown in Fig. 4, the vertical offset was positive.

5   Results

The Mixed RM-ANOVA with the post-hoc test (*Least Significant Difference*) was employed to analyze the objective measures (mean trial completion time, spatial offset, angular offset), and the *Mann-Whitney U test* and the *Kruskal-Wallis H test* for subjective rating. If the ANOVA' s sphericity assumption was violated (Mauchly's test $p < .05$), Greenhouse–Geisser adjustments were therefore performed.

5.1   Trial Completion Time

Figure 5 summarizes the mean trial completion time under four feedback conditions in both the cylinder and the spherical layouts. No significant difference was found between the cylinder (1447 ms, SD: 57) and the spherical layouts (1423 ms, SD: 57) on the trial completion time ($F_{1, 22} = .767$, $p = .09 > .05$).

As expected, there was a significant difference for four *feedback conditions* (*None*: 1145.7 ms, SD: .258; *H*: 1356.1, SD: .279; *A & H*: 1586.3, SD: .345; *A*: 1652.8, SD: .292) on the trial completion time ($F_{3,66} = 18.442$, $p < .001$). Participants spent a bit longer time under three feedback conditions than the *none* condition. The post-doc test revealed that there were significant differences ($p < .05$) for all pairs of feedback conditions except the pair (*auditory* vs. *bimodal* ($p = .423$)).
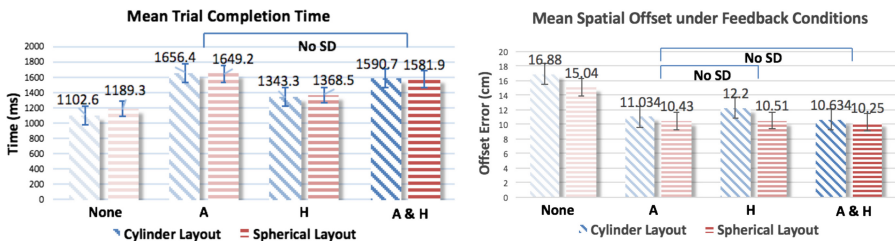

Fig. 5. Mean trial completion time under four feedback conditions (*None: no auditory and haptic feedback; A: auditory feedback only; H: haptic feedback only; A & H: both auditory and haptic feedback*) ($F_{3, 66} = 18.442$, $p < .001$) in both cylinder and spherical layouts (1447 ms vs. 1423 ms, $F_{1, 22} = .767$, $p = .09 > .05$) are illustrated, No SD represents no significant difference between two conditions, other pairs of comparisons differed significantly at $p < .05$ (Left). Mean spatial offset errors under four feedback conditions ($F_{1.968, 43.297} = 23.491$, $p < .001$, $\eta^2 = .516$) in both cylinder and spherical layouts (12.5 cm vs. 11.6 cm, $F_{1, 22} = 1.009$, $p = .326 > .05$) are illustrated (Right).

Participants spent more time under the *auditory* ($p < .001$) and *bimodal* feedback condition ($p < .05$), compared to *haptic*. Most probably, the participants did not need to confirm the acquisition under *none* condition. In other words, they did not know the acquisition was correct or not. They relied on the proprioceptive cues only. In addition, there was no interaction effect for the *feedback conditions* and the *layouts* ($F_{3, 66} = .174$, $p = .914$).

5.2   Spatial Offset

As with the trial completion time, there was no significant difference between *two layouts* (*Cylinder*: 12.5 cm, SD: .006; *Spherical*: 11.6 cm, SD: .006) on the spatial offset ($F_{1, 22} = 1.009$, $p = .326 > .05$). Figure 5 shows the spatial offsets under the *feedback conditions* (*None*: 16.19 cm, SD: .051; *H*: 11.44 cm, SD: .019; *A*: 10.68 cm, SD: .025; *A & H*: 9.97 cm, SD: .019), a significant difference was found ($F_{1.968, 43.297} = 23.491$, $p < .001$, $\eta^2 = .516$). The post-hoc test showed that participants performed much better under each of three types of sensory feedback (*A, H, A & H*) than under the *none* condition ($p < .001$). The *bimodal* condition gave better spatial accuracy than the *haptic* condition ($p < .05$). However, there was no significant difference between *auditory* and *haptic* ($p = .307$), *auditory* and *bimodal* (*A & H*) feedback ($p = .224$). No interaction effect between the feedback conditions and the layouts was found ($F_{3, 66} = .033$, $p = .99$).
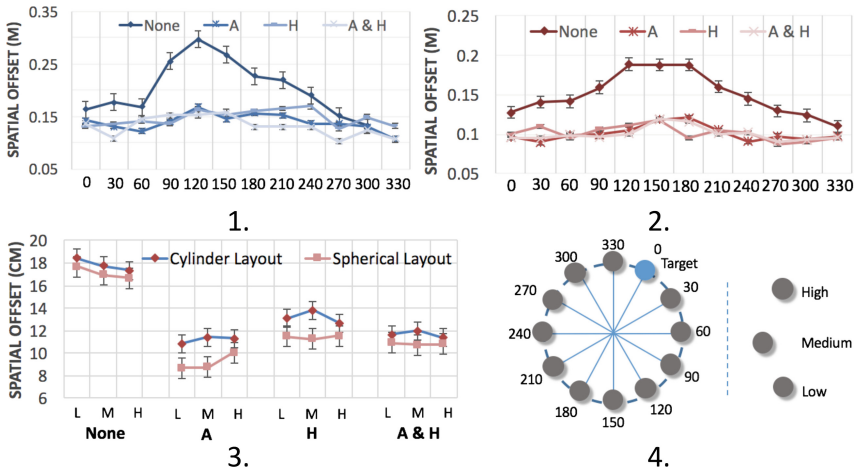

Fig. 6. 1: Mean spatial offset errors at the different horizontal degrees in the cylinder layout. 2: Mean spatial offset errors at the different horizontal degrees ($F_{11, 253} = 91.9$, $p < .001$) in the spherical layout. 3: Mean spatial offset errors at the vertical degrees (L represents targets at the low row, M represents it at the medium row, H represents it at the high row. $F_{2, 46} = 16.72$, $p < .001$) under four feedback conditions ($F_{3, 69} = 19.56$, $p < .05$) in both layouts. 4: The position of target is at 0°, when it is changed, then the 0° is changed accordingly.

However, we additionally included the mean spatial offset errors for the *horizontal rotation degrees* under the four feedback conditions in both layouts, as shown in Fig. 6 (1 and 2). We defined the position of the target as the 0°. Note that we recorded the order of rotations for each target acquisition. There was a significant difference for the horizontal rotation degree on the spatial offset ($F_{11,\ 253} = 91.9$, $p < .001$). As expected, participants performed much better under the three feedback conditions in the both layouts ($F_{3,\ 69} = 11.96$, $p < .001$), compared to the *None* condition. In particular, the additional feedback greatly improved the performance at the horizontal degree of 90, 120, 150, 180, and 210 (See Fig. 6). However, there was no significant difference among the three sensory feedback conditions (*A vs. H: p* = .92; *A vs. A & H: p* = .9; *H vs. A & H: p* = .91).

As with the horizontal degree, a significant difference for the vertical position (low, medium, high) on the spatial offset was also found ($F_{2,\ 46} = 16.72$, $p < .001$). Figure 6 (3) shows the mean spatial offset at the low, medium and high position under four feedback conditions in the cylinder layout and the spherical layout respectively. The participants performed much better under the additional feedback conditions ($F_{3,\ 69} = 19.56$, $p < .05$), compared to the *None* condition. No significant differences were found for the three feedback conditions with the Post-hoc test.
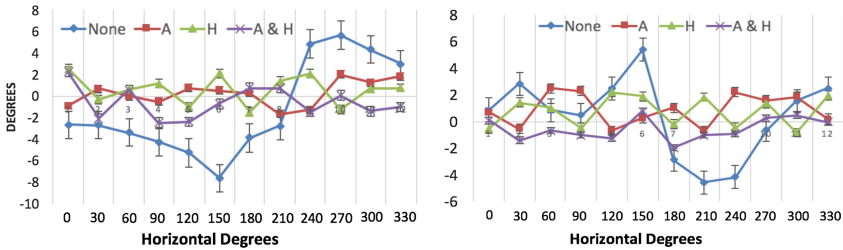
5.3 Angular Offset

We first calculated the absolute angular offset without considering the direction, as shown in Fig. 7. The absolute angular offset indicates the difference between target angular degree and pointing angular degree. Different from the trial completion time and spatial offset, there was a significant difference for the *layouts* (*Cylinder*: 6.316°, SD: .424; *Spherical*: 7.766°, SD: .64) on the angular offset ($F_{1,\ 22} = 5.848$, $p < .05$). Post-hoc tests showed that a significant difference was found under the haptic ($p < .001$) and the bimodal ($p < .05$) feedback conditions. Participants could greatly reduce the angular offset under the three *feedback conditions* respectively ($F_{1.364,\ 30.015} = 28.239$, $p < .001$, $\eta^2 = .562$). In particular, the auditory feedback gave better performance than the haptic feedback ($p < .05$).
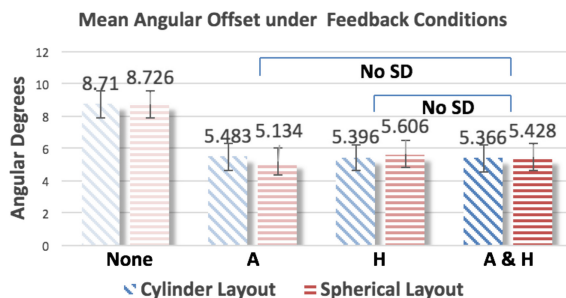
Fig. 7. Mean absolute angular offset under four feedback conditions ($F_{1.364,\ 30.015} = 28.239$, $p < .001$, $\eta^2 = .562$) in both cylinder and spherical layouts ($F_{1,\ 22} = 5.848$, $p < .05$) are illustrated.

To further understand the direction of angular offset, Fig. 8 summarizes that the horizontal rotation degree affected the horizontal angular offset of the acquisition ($F_{11, 253} = 13.23$, $p < .001$), the sensory feedback significantly reduced the horizontal angular offset in both layouts. However, no significant difference was found between two layouts ($F_{1, 23} = .103$, $p = .349$).

Fig. 8. Mean horizontal angular offset at the different horizontal degrees under four feedback conditions in the cylinder layout (left) and spherical layout (right) ($F_{1, 23} = .103$, $p = .349$) are illustrated.

5.4   Subjective Evaluation

The mean score of each dimension in both layouts are shown in Table 1. We found that no significant difference for the *feedback conditions* on the *temporal* ($H = 1.03$, $p < .14$) and *frustration* ($H = .146$, $p = .21$). The participants made more *effort*

Table 1. Mean subjective rating (SR) of NASA-TLX questionnaire for the four feedback conditions in both layouts (C represents the cylinder layout, and S represents the spherical layout). Note that the Mann-Whitney U test is used to compare the layouts, while the Kruskal-Wallis H test is used to compare the feedback conditions. Significant differences were marked with bold.

	None		*A*		*H*		*A&H*		*Mann-Whitney U Test*	*Kruskal-Wallis H Test*
	C	*S*	*C*	*S*	*C*	*S*	*C*	*S*		
Mental	5.3	4.33	4.8	4	4.4	3.5	4.3	3.45	**6.3**	**5.6**
	(1.56)	(2.06)	(1.68)	(1.65)	(1.42)	(1.623)	(1.34)	(1.56)	**$P < .05$**	**$P < .05$**
Physical	5.83	5	5.75	5.083	5.5	4.583	5.25	4.25	**4.9**	**7.97**
	(1.85)	(1.65)	(1.05)	(2.15)	(1.16)	(1.56)	(1.13)	(2.34)	**$P < .05$**	**$P < .05$**
Temporal	6	5.16	5.5	4.8	6	4.83	5.7	4.7	1.03	.21
	(1.15)	(1.80)	(1.43)	(1.67)	(2.51)	(1.85)	(1.64)	(1.84)	$p = .14$	$p = .35$
Performance	5.3	5.6	6.5	6.7	6.37	6.67	7	7.2	**10.523**	**31.08**
	(1.23)	(1.12)	(.90)	(.80)	(1.37)	(1.27)	(.85)	(.81)	**$p < .05$**	**$p < .001$**
Effort	5.67	5.67	4.67	4.92	4.17	5.083	4.083	4.91	**4.56**	**16.35**
	(1.23)	(1.61)	(1.15)	(1.62)	(1.19)	(1.31)	(1.24)	(1.37)	**$p < .05$**	**$p < .001$**
Frustration	3.66	2.31	3.5	2.23	3.25	1.92	3.75	2.31	**11.1**	.146
	(1.61)	(1.25)	(1.93)	(1.09)	(1.48)	(1.03)	(1.71)	(1.03)	**$p < .001$**	$p = .21$

($H = 16.35$, $p < .001$) while resulted in the worst *performance* ($H = 31.08$, $p < .001$) with the *none* condition. In addition, the participants required less *mental* ($H = 5.6$, $p < .05$) and *physical* ($H = 7.97$, $p < .05$) demand for the *bimodal (A & H)* condition to have the best performance, compared to *none* and *auditory* feedback conditions.

In terms of layouts, there was no significant difference for the *temporal* ($U = 1.03$, $p = .14$). The participants reported that the spherical layout made them comfortable and easy to acquire the targets. It requires less *mental* ($U = 6.3$, $p < .05$) and *physical* ($U = 4.9$, $p < .05$) for the spherical layout than the cylinder layout. In the spherical layout, the horizontally egocentric distance for the targets located at low and high rows was smaller than that the targets in the cylinder layout, so participants can easily acquire them without rotating the arm or shoulder in the spherical layout. Subjectively, they can achieve better *performance* ($U = 10.523$, $p < .05$) with less *effort* ($U = 4.56$, $p < .05$) in the spherical layout, although there was no significant difference between two layouts on time and accuracy. Most probably, such a spherical layout can provide additional depth information, which makes the participants feel less blocking. In addition, every participant reported that they did not feel VR motion sickness after the experiment.

6 Discussion

From the experimental results, although no significant difference between two layouts, more participants preferred the spherical layout subjectively. For the feedback conditions, auditory, haptic, and auditory & haptic greatly improved the accuracy of eyes-free target acquisition in VR, rather than the trial completion time (See Figs. 5, 6, 7 and 8). Overall, the haptic feedback allowed the participants to faster acquire target than auditory feedback, while auditory feedback can provide better acquisition accuracy compared to haptic feedback. The bimodal feedback could keep a balance between the acquisition time and the acquisition accuracy basically. Subjective evaluation also reflected the result that 15 participants preferred the bimodal feedback conditions. Thus, H2 hypothesis could be accepted from the experimental results.

6.1 How to Assist Eyes-Free Target Acquisition?

From the existing work [1], the suggestion for the target UI for eyes-free target acquisition was that "*the horizontal ranges over 150 and 180 degrees (the rear region) resulted in poor performance in these dimensions* [1] ". Our results showed that with the help of auditory or bimodal feedback, the accuracy of target acquisition from such a range was significantly improved (See Fig. 6). The participant reported that "*this proximity-based feedback greatly informed the movement and correction of eyes-free target acquisition for the back side target*".

Second, the 'area cursor' [15] and 'bubble cursor' [16] could be used to improve the acquisition accuracy according to the mean spatial offset error (around 2–20 cm). However, the proximity-based multimodal feedback can help reduce the spatial offset error to around 11 cm (See Fig. 5). This could help the designer to utilize such a guideline to decide the size of the bubble cursor.

Third, the acquisition accuracy increased under the proximity-based sensory feedback. It revealed that the proximity-based multimodal feedback can improve the correction phase, however, the trial completion time increased a bit with the help of three types of proximity-based multimodal feedback. Overall, the results suggest that the proximity-based bimodal feedback (both auditory and haptic) can keep a balance between the trial completion time and the acquisition accuracy. This is consistent with Ariza et al. [11], the bimodal is better than unimodal feedback for reduced error rates in 3D selection tasks.

6.2   Limitation and Future Work

In this experiment, we chosen the spatially continuous proximity-based multimodal feedback due to the eyes-free target acquisition requires more spatial-temporal information from the continuous feedback, while the 3D selection allows the participants to look at the target, probably less spatial-temporal information is enough. In the future, we will compare the effects of binary and continuous proximity-based multimodal feedback on eyes-free target acquisition to confirm the difference empirically. Then we will obtain the solid conclusion about the effects of the feedback type on the eyes-free target acquisition in VR.

In this work, we also leave some work to be studied in the future, for example, the effects of non-dominant hand on the eyes-free target acquisition in VR, in the real world, we always focus on performing the main task with our dominant hand, and acquiring the items using the non-dominant. We will observe how different the performance of non-dominant from the dominant hand.

7   Conclusion

This work mainly investigated the effects of proximity-based multimodal feedback on eyes-free target acquisition between two spatial layouts in VR. No significant difference was found between two spatial layouts in terms of objective measures. However, the participants preferred spherical layout subjectively over the cylinder layout for comfort (See Table 1). As expected, auditory, haptic, and bimodal feedback greatly improved the accuracy of eyes-free target acquisition in VR. The results showed that the proximity-based bimodal feedback could keep a balance between the trial completion time and acquisition accuracy basically. This research suggests the improvement of eyes-free target acquisition in VR via the proximity-based multimodal feedback.

References

1. Yan, Y., Yu, C., Ma, X., Huang, S., Iqbal, H., Shi, Y.: Eyes-free target acquisition in interaction space around the body for virtual reality. In: CHI 2018, no. 42, Montréal, Canada. ACM, April 2018
2. Lu, Y., Yu, C., Yi, X., Shi, Y., Zhao, S.: BlindType: eyes-free text entry on handheld touchpad by leveraging thumb's muscle memory. ACM Interact. Mob. Wearable Ubiquitous Technol. **1**(2), 24 (2017). Article 18
3. Mine, M., Brooks Jr, F., Sequin, C.: Moving objects in space: exploiting proprioception in Virtual-Environment interaction. In: Proceedings of SIGGRAPH 1997, pp. 19–26. ACM Press, New York, USA (1997)
4. Chen, X., Schwarz, J., Harrison, C., Mankoff, J., Hudson, S.: Around-body Interaction: sensing and interaction techniques for proprioception-enhanced input with mobile devices. In: MobileHCI 2014, Toronto, Canada, pp. 287–290, September 2014
5. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, pp. 139–183. Elsevier, Amsterdam (1988)
6. Ariza, O., Lange, M., Steinicke, F., Bruder, G.: Vibrotactile assistance for user guidance towards selection targets in VR and the cognitive resources involved. In: 2017 IEEE Symposium on 3D User Interfaces (3DUI), pp. 95–98, March 2017
7. Cockburn, A., Brewster, S.: Multimodal feedback for the acquisition of small targets. Ergonomics **48**(9), 1129–1150 (2005)
8. Zhao, S., Dragicevic, P., Chigell, M., Balakrishnan, R., Baudisch, P.: earPod: eyes-free menu selection with touch input and reactive audio feedback. In: CHI, San Jose, CA, USA, pp. 1395–1404. ACM, May 2007
9. Park, Y., Kim, J., Lee, K.: Effects of auditory feedback on menu selection in hand-gesture interfaces. IEEE Multimed. **22**(1), 32–40 (2015)
10. Gao, B., Kim, H., Lee, H., Lee, J., Kim, J.: Use of sound to provide occluded visual information in touch gestural interfaces. In: CHI Extended Abstracts, Seoul, South Korea, pp. 1277–1282. ACM, April 2015
11. Ariza, O., Bruder, G., Katzakis, N., Steinicke, F.: Analysis of proximity-based multimodal feedback for 3D selection in immersive virtual environments. In: IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Germany, pp. 327–334, March 2018
12. Argelaguet, F., Andujar, C.: A survey of 3D object selection techniques for virtual environments. Comput. Graph. **37**(3), 121–136 (2013)
13. Oh, U., Branham, S., Findlater, L., Kane, S.K.: Audio-based feedback techniques for teaching touchscreen gestures. ACM Trans. Access. Comput. **7**, 9 (2015)
14. Poupyrev, I., Ichikawa, T., Weghorst, S., Billinghurst, M.: Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. Comput. Graph. Forum **17**(3), 41–52 (1998)
15. Worden, A., Walker, N., Bharat, K., Hudson, S.: Making computers easier for older adults to use: area cursors and sticky icons. In: CHI, New York, NY, USA, pp. 266–271. ACM (1997)
16. Grossman, T., Balakrishnan, R.: The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In: CHI, New York, NY, USA, pp. 281–290. ACM (2005)
17. Gao, B., Kim, H., Kim, B., Kim, J.: Artificial landmarks to facilitate spatial learning and recalling for the curved visual wall layout in virtual reality. In: IEEE BigComp 2018, Shanghai, China, pp. 475–482, January 2018

18. Gao, B., Kim, B., Kim, J., Kim, H.: Amphitheater layout with egocentric distance-based item sizing and landmarks for browsing in virtual reality. Int. J. Hum.-Comput. Interact. **35**(10), 831–845 (2019)
19. Gerig, N., Mayo, J., Baur, K., Wittmann, F., Riener, R., Wolf, P.: Missing depth cues in virtual reality limit performance and quality of three dimensional reaching movements. PLoS ONE **13**(1), e0189275 (2018)
20. Yamanaka, S., Miyashit, H.: Modeling the steering time difference between narrowing and widening tunnels. In: CHI, San Jose, CA, USA, pp. 1846–1856. ACM (2016)
21. Loomis, J.M., Philbeck, J.W.: Measuring spatial perception with spatial updating and action. In: Behrmann, M., Klatzky, R.L., Macwhinney, B. (eds.) Embodiment, Ego-Space, and Action, pp. 1–43. Psychology Press, New York (2008)
22. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. J. Exp. Psychol. **47**(6), 381–391 (1954)
23. Ens, B., Finnegan, R., Irani, P.P.: The personal cockpit: a spatial interface for effective task switching on head-worn displays. In: CHI, Toronto, Canada, pp. 3171–3180. ACM (2014)
24. Ens, B., Ofek, E., Bruce, N., Irani, P.: Spatial constancy of surface-embedded layouts across multiple environments. In: SUI, Los Angeles, CA, USA, pp. 65–68. ACM (2015)
25. Gao, B., Kim, H., Lee, H., Lee, J., Kim, J.: Effects of continuous auditory feedback on drawing trajectory-based finger gestures. IEEE Trans. Hum.- Mach. Syst. **48**(06), 658–669 (2018)
26. HTC VIVE Pro. https://www.vive.com/us/product/vive-pro/. Accessed 20 Nov 2018
27. VRTK. https://vrtoolkit.readme.io/docs. Accessed 20 Nov 2018