Exploring Head-based Mode-Switching in Virtual Reality

Rongkai Shi*, Nan Zhu†, Hai-Ning Liang‡
Department of Computing
Xi'an Jiaotong-Liverpool University

Shengdong Zhao§
NUS-HCI Lab
National University of Singapore

ABSTRACT

Mode-switching supports multilevel operations using a limited number of input methods. In Virtual Reality (VR) head-mounted displays (HMD), common approaches for mode-switching use buttons, controllers, and users' hands. However, they are inefficient and challenging to do with tasks that require both hands (e.g., when users need to use two hands during drawing operations). Using head gestures for mode-switching can be an efficient and cost-effective way, allowing for a more continuous and smooth transition between modes. In this paper, we explore the use of head gestures for mode-switching especially in scenarios when both users' hands are performing tasks. We present a first user study that evaluated eight head gestures that could be suitable for VR HMD with a dual-hand line-drawing task. Results show that move forward, move backward, roll left, and roll right led to better performance and are preferred by participants. A second study integrating these four gestures in Tilt Brush, an open-source painting VR application, is conducted to further explore the applicability of these gestures and derive insights. Results show that Tilt Brush with head gestures allowed users to change modes with ease and led to improved interaction and user experience. The paper ends with a discussion on some design recommendations for using head-based mode-switching in VR HMD.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—Empirical studies in HCI;

1 INTRODUCTION

Modes represent distinct user interface settings, where two modes lead to different results from the same input, and are manifested by how an interface responds to inputs [22]. They are common in all types of user interfaces (UI). Mode-switching is the transition from one mode to another. However, different modes and switching between them, to an extent, can lead to errors, increased confusion, unnecessary restrictions, and added complexity in a UI [22]. Having a practical, efficient, and easily-performed mode-switching mechanism is important for most interfaces, especially for Virtual Reality and Augmented Reality head-mounted displays (VR/AR HMD) which have more interactive possibilities and complexity.

Researchers have utilized specific hand-held devices or hand motions for mode-switching for different devices, especially for those with pen- and touch-based interfaces (e.g., [9, 13, 30, 33]). However, mode-switching in immersive VR/AR HMD is relatively underexplored. Applications in VR HMD, just like those in other platforms, can have multiple modes. For example, a 3D-painting application (e.g., [18]) can contain several modes for interacting with it, such as painting and erasing, while in a 3D-modeling application (e.g., [43]) interaction could involve object selection and manipulation, each

*e-mail: rongkai.shi19@student.xjtlu.edu.cn
†e-mail: nan.zhu18@student.xjtlu.edu.cn
‡*Corresponding author*; e-mail: haining.liang@xjtlu.edu.cn
§e-mail: zhaosd@comp.nus.edu.sg

of which has a distinct effect from the same input. Having an efficient mode-switching technique or mechanism can be supportive for scenarios that involve more than one mode. The existing research tended to focus on and suggested using hand-based interactions for mode switching [29, 31]. However, hand-based mode-switching techniques are not suitable or ideal for many VR applications, for example, when users need to use both hands for interaction, which is common in HMD [12]. Unlike 2DUI, which have more constraints because of their input devices (e.g., mouse or pen), VR/AR provide more ways of interaction and degrees of freedom where users' hands are frequently used [12]. As such, a hands-free approach for mode switching can be useful in many VR applications to support a richer and more continuous flow of interaction.

Voice control is one hands-free approach commonly supported by HMD. However, it has poor performance in noisy environments [12, 19] and is unsuitable in a public place or even a private space that is shared with others. Foot interaction is another hands-free mechanism and has been studied as an input modality in VR (e.g., [17,36]). It can be intuitive and helpful but it also requires extra detection devices that tend to be difficult to set up [7, 17, 34]. Similarly, eye-based interaction requires additional trackers inside the HMD, can only provide limited possibilities, is sensitive to the accuracy of the eye tracker, and can lead to eye fatigue and errors [16, 17, 43].

Head-based gestures can be a natural and cost-efficient hands-free approach that suits "hands-busy" scenarios such as those often found in painting, modeling, or gaming applications. In general, using head gestures in HMD has following benefits: (1) users are familiar with the head gestures, such as to nod, raise, or turn the head; (2) head gestures are easy, cheap, and fast to perform; (3) head movements can be tracked with high accuracy by most of the current HMD devices without any additional hardware [39, 42]. As such, head gestures might be able to aid mode-switching in a fast, low-cost, and error free manner. To best of our knowledge, there has not been prior research that has formally and systematically explored the use of head gestures for mode-switching in VR HMD. In this research, we aim to investigate what potential head gesture(s) can lead to better performance and usability for mode-switching in VR (Study One) and how they perform in a real application (Study Two).

We first evaluated eight head gestures that were selected based on three design criteria to ensure they are potentially suitable for mode-switching in VR: move forward, move backward, pitch up, pitch down, yaw left, yaw right, roll left, and roll right (see Fig. 1). We conducted two user studies to evaluate their usability and performance. In the first study, we collected and compared the performance and users' subjective feedback of head gestures in an extended line-drawing task which was designed based on the certified "*subtraction method*" [4, 13, 30, 31, 33]. Results from this study show that move forward, move backward, roll left, and roll right had better performance and were preferred by the participants. In the second study, we further evaluated these four gestures in a actual VR painting application where these gestures helped achieve quick and smooth switch between the painting tools, allowing an enhanced continuous interaction and user experience with the application.

Overall, this paper makes three main contributions: (1) it provides a first systematic exploration of using head gestures for mode-switching in VR HMD; (2) it reports an effective integration of four
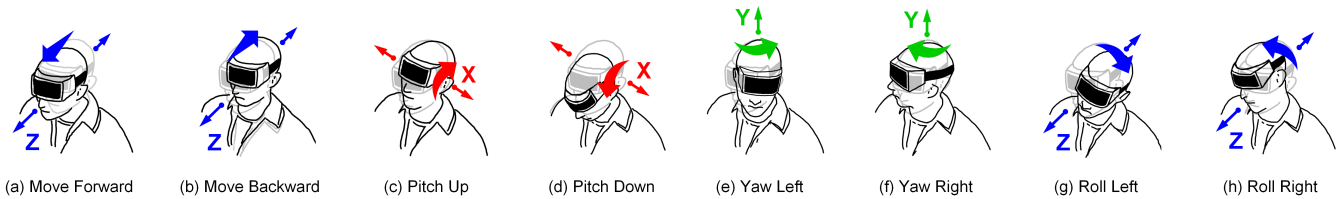
Figure 1: Investigated head-based gestures: (a) move forward, (b) move backward, (c) pitch up, (d) pitch down, (e) yaw left, (f) yaw right, (g) roll left, (h) roll right.

head gestures in an actual application to show their applicability in allowing fast, smooth, and continuous mode-switching; (3) it offers a set of design considerations for incorporating head gestures for mode-switching in VR HMD.

2 RELATED WORK

2.1 Mode-Switching in 2D User Interfaces

The early research investigated mode switching in desktop computers. Dillon et al. [4] measured the time cost and errors for command selection. They laid the foundation of the "*subtraction method*" which can derive the true time spent for switching commands. Sellen et al. [25] discussed and suggested user-maintained mechanism to reduce mode-switching errors during text editing. Kabbash et al. [10] evaluated two-handed input in a compound line-drawing and color selection task. Based on these prior works, Li et al. [13] compared five techniques for mode switching in pen-based interfaces. They measured the mode-switching time by subtracting the time to complete single-color pie-crossing tasks from those when alternately switching between two colors. By following this approach, Hinckley et al. [9] investigated the performance of mode switching using a user-maintained mechanism for tablet devices. In addition, Tu et al. [33] and Surale et al. [30] followed the subtraction method and used a modified pie-crossing task to compare mode-switching techniques for pen- and touch-based interfaces, respectively. The subtraction method provides a precise measure to quantify the time cost of mode-switching and has been adapted and examined by above-mentioned work. Thus, we followed this proven, robust approach to design and conduct the experiments in our study.

2.2 Mode-Switching and Related Studies in HMD

VR or AR HMD provide larger interaction space compared to 2DUI. With the support of 3D input technologies, HMD leverage the use of controller, hand gestures, head gestures, body gestures, and eye gaze [12]. Specifically, hand gestures to do mode-switching have been discussed in HMD. Surale et al. [31] presented two studies to evaluate barehand mid-air mode-switching techniques for VR HMD. They followed the subtraction method and compared dominant and non-dominant hand gestures in a VR line-drawing task. As for AR HMD, Smith et al. [29] performed an evaluation of five mode-switching techniques and found that the non-preferred hand was fast and led to lower errors. They further evaluated the scalability of non-preferred hand mode-switching in their recent study [28].

The above prior work has primarily focused on exploring mode-switching techniques in controlled experimental settings. However, there has been limited exploration of such techniques in actual HMD applications. Some studies dealing with tasks in VR applications, such as menu selection [18, 20], task-switching [5, 14], disambiguation [3], though have involved the concept of mode-switching as part of their research, did not conduct controlled evaluations. In this research, we first conducted a systematic evaluation of mode-switching using head motions in a more controlled setting. We then incorporated the most suitable and best rated motions to a real

VR application to assess their performance and usability in more uncontrolled and realistic scenario.

2.3 Head-based Interaction

As part of our body language, head movements are natural and deliberate, and due to this, prior work has leveraged head-based interactions in HMD. Head-gaze interaction is one of the most common approaches, which has been widely studied for target selection (e.g., [2, 11, 17, 21]) and text entry (e.g., [35, 37, 41]) in VR/AR HMD. As head movements can be performed accurately yet effectively, head-gaze interaction has been used as a standalone approach and combined with or supported by other modalities, such as eye-based interaction [2, 11, 15, 21, 26, 27, 35]. Head-gaze has also been used, for example, to interact in games [1], map interfaces [8], and to construct new interaction methods [6, 26, 27, 38].

Head-based gestures are another form of head-based interaction. They can be sensed and captured by the built-in inertial measurement unit sensor of smartphones or HMD. These gestures have been applied for different scenarios. Yi et al. [40] leveraged head gestures for user authentication. Rudi et al. [23] used nod and shake to interact with maps on smart glasses. Tregillus [32] tracked head tilting (xz plane) for omnidirectional navigation in mobile VR. Yu et al. [42] explored the possibility of using move forward/backward (z axis) as a hands-free interaction approach in VR HMD, which has been applied for text entry recently [16]. Yan et al. [39] conducted an elicitation study and derived a head gesture set for HMD devices. They further evaluated its performance and concluded that head gestures could work as a supplementary input approach for HMD.

In summary, prior work has proposed and examined using head gestures in HMD but outside of mode-switching. This study is inspired by this prior research and has adapted a proven method and task scenario to evaluate the use of head-based mode-switching gestures in VR HMD (Study One) and took a step further to assess the mode-switching gestures in a actual VR application with more realistic scenario (Study Two). To do this, we first sift out suitable head gestures based on specific design criteria.

3 HEAD-BASED MODE-SWITCHING

Head gestures are basic human movements and are quick, low-cost and fast to perform. There are a wide range of head gestures that can be translated into commands. Inspired by [31], we defined the following three design criteria (DC) to narrow the scope for head-based mode-switching gestures in VR HMD.

- **DC1. User-maintainable**. A user-maintained mechanism helps users maintain awareness of the system state, in which potentially reduces mode errors [9, 25]. Such awareness may also contribute to quickly correct a false positive if it happens. The head gestures should be able to be maintained by the users when the mode is enabled. This criterion rejects the gestures that are hard to perform or maintain.

- **DC2. Simple and fast**. Considering that users may be performing dual-hand interactions, their current task may already

involve significant amount of cognitive and physical efforts. Thus, the mode-switching action need to be simple and straightforward so that it is possible to perform simultaneously with the existing tasks. In addition, it should be fast to allow a seamless integration with the other task(s).

- *DC3.* **Independent**. A mode-switching action should be independent of previous tracking states and should not rely on time-based actions [31]. This guarantees an explicit transition between states happens during mode-switching. Thus, gestures involving dwelling, combining different actions, or repeating an action are excluded.

Based on above criteria, we derive 8 head gestures (in 4 pairs) that are potentially suitable for mode-switching in VR HMD, as shown in Fig. 1. Pitch up/down, yaw left/right, and roll left/right are based on rotational movements along x, y, z axis, respectively. They are commonly performed in daily life and have been studied in prior work (e.g., [23]). On the other hand, move forward/backward are based on positional movements along z axis. According to [42], it is useful to allow users to lean slightly forward/backward to perform head motions along the z axis because it is more natural and helps with their balance. We also followed this approach in our studies.

We conducted a pilot study with 6 participants to collect head movement data including the head position (Euler distance in 3 axes) and head orientation (Euler angles of pitch, yaw, roll). The participants performed each gesture five times. We then segmented each gesture and extracted the statistical characteristics of each gesture's potential position and orientation patterns in a certain duration. The information is used to setup each gesture's recognition. Our approach achieved near 100% accuracy in gesture recognition using another 6 users in the pilot test.

4 STUDY ONE: FEASIBILITY EVALUATION

The aim of this user study was to investigate the following research question: *which head gesture(s) can lead to better user performance and feedback in VR HMD mode-switching tasks?* To do this, we present a comparison of the above-mentioned eight head gestures in an extended line-drawing task in VR where they are used to switch between two colors when drawing the lines.

4.1 Participants and Apparatus

Sixteen participants (8 females) from a local university volunteered to join this experiment. Participants were on average 21 years old ($M = 20.94$, $SD = 1.88$). According to the data from the pre-experiment questionnaire, all of them were right-handed, and had normal or corrected-to-normal vision and had no issues distinguishing colors. Six of them had used or interacted with VR HMD before the experiment. All participants had no physical discomfort and did not have problems moving their head to the mentioned positions and orientations.

An HTC VIVE Pro VR HMD was used to provide the immersive virtual environments (VE). It has a resolution of 1440×1600 px per eye, a refreshing rate of 90 Hz, and a $110°$ field of view. It was connected to an Intel Core i7 processor PC with NVIDIA Quadro P5200 graphics card. The VE was written in C# using the Unity3D platform (version 2019.2.3f1). VIVE controllers were used to interact with the VE.

4.2 Task

In this study, we extended the VR line-drawing task described in [31] by asking participants to draw two lines with both hands at the same time. This task represented an extreme situation where users had to perform a non-trivial task with two hands simultaneously. The head gestures were used to change line colors (blue/red) which represented different modes.
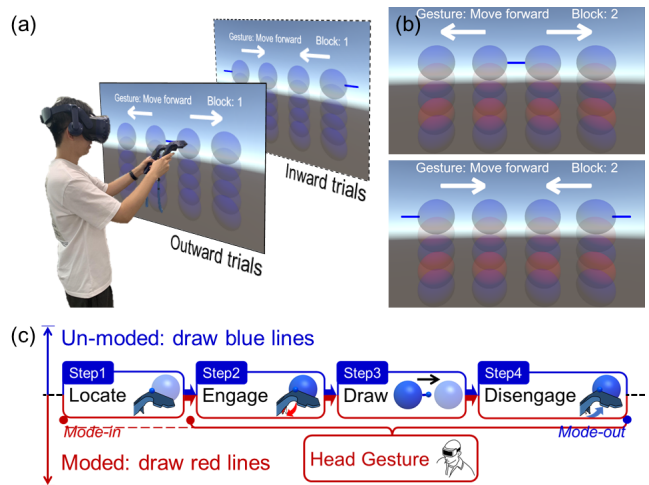

Figure 2: Experiment setup. (a) An example of a *baseline* block which only requires to draw blue lines between the spheres. (b) *Outward* (upper) and *inward* trials (lower) in a *compound* block. (c) The line-drawing process in a trial. The upper part (in blue) shows an *un-moded* line-drawing process, and the lower part (in red) shows a *moded* line-drawing process. The head gesture can be performed before engaging with the drawing activity, and should be maintained before disengaging from it.

At the beginning of the experiment, five groups of semi-transparent spheres were presented to participants in the VE (see Fig. 2a,b). Each group involved four spheres that had the same size (diameter = 80 mm) and was on the same horizontal line. The distance between two adjacent spheres was also 80 mm. The distance was chosen to allow participants see all the spheres and the text hints placed in front of them. Participants were required to draw two lines from the two inner spheres to the two outer spheres (*outward trials*) or to do the reverse (*inward trials*). To reduce any possible confusion, we put a line between the inner spheres for the outward trials and two lines outside the outer spheres for the inward trials. We also put two arrows above the spheres to show the direction of movement (see Fig. 2a,b).

In each trial, the process of drawing could be divided into four steps (see Fig. 2c): (Step1) placing two controllers on the starting spheres; (Step2) engaging with the drawing task by pressing the trigger in each hand; (Step3) performing the drawing activity until the line reached the ending spheres; (Step4) releasing the triggers to disengage and indicate the end of the drawing task. If participants engaged or disengaged outside the specified spheres, an error was logged. Participants then needed to redo this trial. The spheres would turn opaque if the controllers were inside to provide visual feedback to users.

We defined two types of task blocks: a *baseline* block and a *compound* block, which were based on the subtraction method [4, 13, 30, 31, 33]. Both types of blocks had two parts: the first part was to do 5 outward trials and the second part was 5 inward trials. In a baseline block, only a default mode was involved, and the spheres were in blue in all 10 trials (Fig. 2a). As such, participants were only required to draw *un-moded*, blue lines. In a compound block, the spheres were in blue in the first, third, and last trial of each part, but in red in the second and fourth trials (Fig. 2b). For the former case, like in a baseline block, participants only needed to draw the un-moded lines (see the upper part in Fig. 2c). While for the latter case, participants had to switch the mode with the specified mode-switching head gesture and draw *moded*, red lines (see the lower part in Fig. 2c). The head gesture should be performed before engaging with the drawing activity (before Step2). As suggested in

previous studies [9, 25, 31], we used a user-maintained mechanism in this study. That is, users needed to hold the head gesture while drawing (Step2, 3, 4). There were text hints provided above the spheres to show the current head gesture and block (see Fig. 2a,b). If participants failed to switch in or out the mode, an error was logged, and they were required to redo this trial. All trials were performed in a standing position.

4.3 Design and Procedure

The experiment included three sessions: (1) an introduction session to collect participants' demographics and background information and to brief them of study purpose, (2) an experimental session with eight sub-sessions for each head gesture, and (3) a post-experiment questionnaire session to collect their feedback. We used a within-subjects design with head gesture (GESTURE) as the independent variable. The order of GESTURE was counterbalanced using a Latin Square design. For each GESTURE condition, participants were requested to complete 9 blocks. The first block was a baseline task and then a compound task, alternating until the ninth block [13]. The first two blocks were training blocks for participants to familiarize themselves with the head gesture and drawing process, and the remaining seven were formal blocks used in the analysis. Considering the duration of the experiments, we only asked participants to draw horizontal lines (outwards and inwards). When participants completed all the blocks for a GESTURE condition, they were given time to rest. A questionnaire about subjective ratings for the just-used head gesture was given during the break. After participants completed all the 8 conditions, we ask them to complete a post-questionnaire that asked them to provide a ranking of their preferred head gestures.

Based on this design, there were 8 GESTURE × 9 blocks of tasks × 10 line-drawing trials (5 outward + 5 inward) = 720 times of line-drawing for each participant, including 160 trials for training and 560 formal trials for analysis.

4.4 Evaluation Metrics

The dependent variables were *mode-switching time* and *error rates* in compound blocks. We used BLOCK to denote the compound blocks. BLOCK is a repeated measure in the analysis. The calculation of mode-switching time and the types of errors are explained in the following sub-sections. In addition, we used questionnaires to collect participants' subjective feedback, including ratings to each head gesture and a ranking of the most preferred gestures.

4.4.1 Mode-Switching Time

For each part of a block, there were three cycles involved, as shown in Fig. 3. The first cycle started from the moment the spheres were visible until the triggers were released after drawing the lines for the first group of spheres. The second cycle began immediately after, and ended when the triggers were released after drawing the lines for the third group. And, finally, the third cycle began immediately after, ending when the triggers were released after drawing the lines for the fifth group.

The second and third cycles in each part of a block were defined as *full cycles* [13,30,31]. A full cycle in a compound block contained a complete mode-switching process including switching into a mode using the specified head gesture, completing a moded line-drawing task, switching back to the baseline mode, completing an un-moded line-drawing task. The first cycle guaranteed an un-moded line-drawing task performed before a full cycle. As such, there were 4 full cycles in each block (2 from outward trials, 2 from inward trials).

We used the time of error-free trials for calculation. The mean time of error-free full cycles in a block was defined as *average full cycle duration*. The mode-switching time for each of the 3 compound blocks was computed by subtracting the mean of the two adjacent baseline blocks' average full cycle duration from the
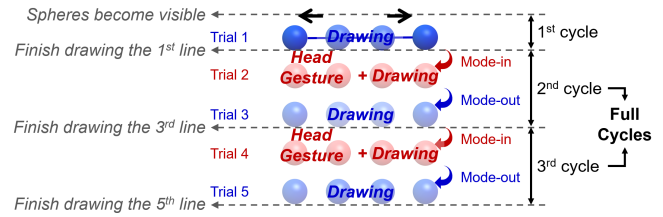

Figure 3: An example of the three cycles from outward trials in a compound block. A full cycle contains a complete mode-switching process: switching mode *in* to draw *moded* lines and switching mode *out* to draw *un-moded* lines.

compound block's average full cycle duration. In total, we derived 24 mode-switching time per participant (8 GESTURE × 3 BLOCK).

4.4.2 Error Rates

Like this prior study [31], we defined four error types. First, a *start error* happened if participants failed to engage in the line-drawing activity; that is, participants pressed the triggers outside the starting spheres. Second, an *end error* occurred if participants failed to draw the lines to the specified end locations. In other words, participants released the triggers outside the ending spheres. Third, a *mode-in error* happened if participants failed to use head gestures to draw the lines in red spheres, indicating a failure of switching into a mode. Finally, a *mode-out error* happened if participants used a mode-switching gesture to draw the lines in the blue spheres, which would be a failure of switching back to baseline mode. Each of these error types were counted if they took place. In each BLOCK, an error rate was the ratio of the number of errors to the total number of trials, and an *overall error rate* was the sum of all four error rates.

4.4.3 Subjective Feedback

After finishing each GESTURE condition, participants were asked to rate the just-used head gesture in terms of ease of learning, ease of use, accuracy, speed, neck fatigue and VR sickness. A 5-point continuous scale was used, with 1 being the most negative perception and 5 being the most positive. These questions were adapted from prior mode-switching studies [13, 20, 29–31]. At the end of the experiment, we asked participants to rank their preferred head gestures based on the experience of using them for mode-switching in VR and provide their reasons.

4.5 Results

We used Excel (MS Office 365) and SPSS (version 26.0) for data processing and analysis. Before analyzing the data, we identified outliers of error-free line-drawing times in full cycles that were outside 3 standard deviations of the mean. We removed 1.56% of the error-free trials in the full cycles. The remaining trials were used to calculate the mode-switching time with the explained strategy.

Shapiro-Wilk tests showed that mode-switching time and overall error rates were normally distributed ($p > .05$). Thus, we performed Repeated Measures Analysis of Variance (RM-ANOVA) tests for these two measurements, Greenhouse-Geisser ($\varepsilon < .75$ in our results) corrections when sphericity assumption was not met, and Bonferroni-adjusted post-hoc tests where applicable. Start, end, mode-in, and mode-out error rates were not normalised data according Shapiro-Wilk test ($p < .05$). As such, non-parametric Friedman tests were used for these measurements with GESTURE as the factor, and if a significant main effect was found, post-hoc analysis with Wilcoxon signed rank tests was conducted with Bonferrorni correction.

4.5.1 Mode-Switching Time

RM-ANOVA test did not show a significant GESTURE × BLOCK interaction effect on mode-switching time ($p > .05$), indicating the
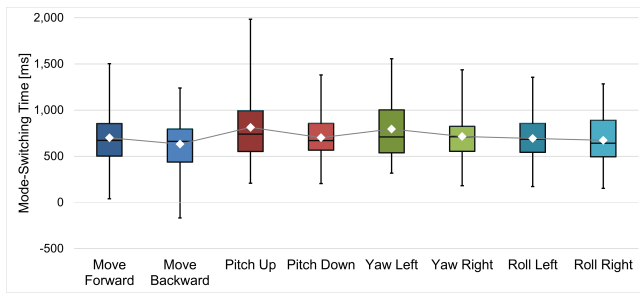
Figure 4: Boxplot of mode-switching time (in milliseconds) by GESTURE. White diamonds indicate mean mode-switching time for each head gesture.
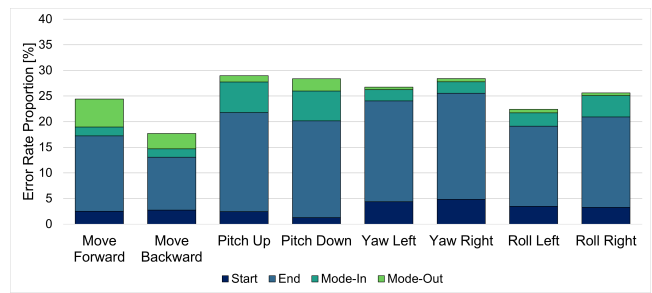

Figure 5: Proportion of each type of error rates by GESTURE. The sum of four types of error rates in a bar is the overall error rate for that head gesture.

stability of time performance among the blocks. This is in line with prior work [13, 30, 31, 33]. Fig. 4 shows the results of mode-switching time for each head gesture. There was no significant main effect for GESTURE on mode-switching time ($p > .05$).

4.5.2 Error Rates

An error rates matrix summarizing the mean and standard deviation of error rates for each gesture is shown in Table 1. RM-ANOVA test did not show a significant GESTURE × BLOCK interaction effect on overall error rates ($p > .05$), indicating the stability of accuracy among the blocks. There was a main effect for GESTURE on overall error rates ($F(7,105) = 2.281$, $p = .033$, $\eta^2 = .132$). However, the Bonferroni-corrected post-hoc tests did not yield significant differences ($p > .05$).

Fig. 5 illustrates the proportion of specific error rates by GESTURE. On average, end errors accounted for the largest weight (68%). The proportion of start errors (12%) and mode-in errors (13%) were close, while mode-out errors had the smallest share (7%). Results of the Friedman test showed no significant main effect for GESTURE on start error rates ($p > .05$). Average end error rates took up the largest proportion. The Friedman test yielded a significant main effect for GESTURE on end error rates ($\chi^2(7) = 31.318$, $p < .001$). Post-hoc tests showed that move backward (10.34%) had significant lower end error rates than pitch up (19.35%), pitch down (18.86%), yaw left (19.66%), yaw right (20.68%), and roll right (17.66%) ($p \leq .001$).

There was a significant main effect for GESTURE on mode-in error rates ($\chi^2(7) = 21.388$, $p = .003$). Post-hoc tests showed that move forward had significant lower mode-in error rates (1.74%) than pitch down (5.84%, $p = .006$). Moreover, mode-in error rates of move backward (1.62%) were significant lower compared to pitch up (5.94%, $p = .002$) and pitch down (5.84%, $p = .001$). There was a significant main effect for GESTURE on mode-out error rates ($\chi^2(7) = 40.215$, $p < .001$). Post-hoc tests showed that yaw left (0.49%) had lower rates than move forward (5.46%, $p = .001$), move backward (3.01%, $p = .003$), and pitch down (2.37%, $p = .004$). Besides, in addition to being higher than yaw left, move forward had a higher rates than yaw right (0.66%, $p = .001$), roll left (0.71%, $p = .005$), and roll right (0.53%, $p < .001$).

4.5.3 Subjective Ratings, Ranking, and Feedback

The results of the subjective ratings are listed in Fig. 6. Friedman tests on subjective ratings showed significant main effects in ease of use ($\chi^2(7) = 15.048$, $p = .035$), accuracy ($\chi^2(7) = 21.272$, $p = .003$) and speed ($\chi^2(7) = 20.336$, $p = .005$), but no significant effects in the remaining measures ($p > .05$). In terms of ease of use, post-hoc tests did not yield any significant differences among the head gestures ($p > .05$). For accuracy, move forward was rated more accurate than pitch down ($Z = -2.877$, $p = .004$) and yaw right ($Z = -2.913$, $p = .004$). Moreover, move backward was rated more accurate than yaw

right ($Z = -2.810$, $p = .005$). As for speed, post-hoc tests indicated that yaw left was rated faster than yaw right ($Z = -2.828$, $p = .005$).

Fig. 7 shows the distribution of participants' rankings of the head gestures after the trials. Participants tended to rank them based on gesture pairs. Paired gestures represented reverse pattern which could be a simple and intentional input [39] in case that more than one switching is required among multiple modes. Move forward/backward were ranked as the best head gestures for mode-switching in VR, closely followed by roll left/right. Pitch up/down and yaw left/right were ranked lower (see Fig. 7). Participants reported that, though they were able to complete the tasks, they could not see the whole task space when using pitch up/down or yaw left/right.

4.6 Discussion

We investigated using potential head gestures to perform mode-switching in a dual-hand line-drawing task when using a VR HMD. Our study revealed insights into the user performance and feedback between these head gestures.

Results did not yield a significant fast or slow head gesture for switching between modes in VR HMD. The time cost for performing these gestures was similar. All the eight filtered head gestures are onefold (**DC3**) and simple to perform (**DC1**). We used error-free trials for measuring mode-switching time. The average time was less than 850 ms for all motions, indicating that using head gestures for mode-switching in VR HMD is efficient if a gesture can be successfully performed. Our results meet **DC2**.

We received high end error rates regardless of the head gestures ($M = 17.11\%$) compared to prior VR mode-switching study [31]. In other words, participants faced difficulty in finishing the lines to the specified positions. One main reason might be the dual-hand line-drawing task we used was more challenging than the one-hand line-drawing task in VR [31] or pie-crossing tasks in 2DUI [13, 30, 33]. Our setting represents an extreme situation where both hands are occupied. For such case, an explicit advantage of using head gestures for mode-switching is that its interaction is hands-free and does not rely on extra tracking devices.

Move forward/backward were the two head gestures that participants gave the highest rankings and good ratings. Also, these two gestures had low overall error rates. The difference mainly comes from end errors where move backward and forward had the lowest and second lowest end error rates (see Table 1). This was because the spheres were always in participants' front view, as they did not need them to do rotational movements when performing these two gestures. Move forward/backward had less mode-in errors but more mode-out errors than others, especially for move forward which had significantly higher mode-out errors. This implies that participants did not fully return to the default position, and thus did not switch back to the default mode. As mentioned, when performing these two head movements, the users' body may move slightly. Since our task

Table 1: Mean and standard deviation of overall, start, end, mode-in, and mode-out error rates (%) by GESTURE. The lowest and second lowest error rates are highlighted in green and light green, while the highest and second highest error rates are highlighted in orange and light orange.

Head Gesture	Overall Error Rate	Start Error Rate	End Error Rate	Mode-in Error Rate	Mode-out Error Rate
Move Forward	$24.42_{\pm15.35}$	$2.52_{\pm5.09}$	$14.70_{\pm11.33}$	$1.74_{\pm3.81}$	$5.46_{\pm9.74}$
Move Backward	$17.71_{\pm15.00}$	$2.74_{\pm5.49}$	$10.34_{\pm11.68}$	$1.62_{\pm3.36}$	$3.01_{\pm5.27}$
Pitch Up	$28.98_{\pm20.98}$	$2.45_{\pm5.07}$	$19.35_{\pm14.80}$	$5.94_{\pm9.56}$	$1.24_{\pm3.80}$
Pitch Down	$28.36_{\pm17.78}$	$1.29_{\pm3.12}$	$18.86_{\pm13.67}$	$5.84_{\pm8.43}$	$2.37_{\pm5.46}$
Yaw Left	$26.73_{\pm21.42}$	$4.39_{\pm8.65}$	$19.66_{\pm18.04}$	$2.20_{\pm4.26}$	$0.49_{\pm2.77}$
Yaw Right	$28.43_{\pm21.14}$	$4.85_{\pm8.69}$	$20.68_{\pm14.56}$	$2.24_{\pm4.49}$	$0.66_{\pm2.73}$
Roll Left	$22.42_{\pm18.28}$	$3.47_{\pm5.39}$	$15.65_{\pm14.37}$	$2.58_{\pm6.52}$	$0.71_{\pm2.99}$
Roll Right	$25.60_{\pm14.31}$	$3.27_{\pm4.88}$	$17.66_{\pm12.04}$	$4.14_{\pm6.09}$	$0.53_{\pm2.19}$
Total	$25.33_{\pm18.42}$	$3.12_{\pm6.11}$	$17.11_{\pm14.2}$	$3.29_{\pm6.35}$	$1.81_{\pm5.17}$

occupied high cognitive resources, it was possible that participants were more focused on head movements, so they tried to move their head back but their body did not completely return. This can be easily minimized by fine-tuning the threshold to a slightly lower value for the returning motion.

The objective results of roll left/right showed a sense of balance with little obvious positive and negative difference with other gestures in terms of mode-switching time and error rates. They received comparable high ranking and ratings with move forward/backward. Participants did not provide explicit reason for their ratings, especially in relation to move forward/backward. They said they were based on personal feelings of the gestures that were "intuitive" to make. In short, our results show an overall strong case for move forward/backward and roll left/right to be used for mode-switching in VR HMD.

Pitch up/down, and yaw left/right received lower rank from participants (see Fig. 7) because they could not see the task area when performing these head gestures. Due to this situation, they perceived they had more mistakes with these gestures and gave lower ratings on accuracy compared to move forward/backward and roll left/right (see Fig. 6). These results do not support the use of pitch up/down, and yaw left/right in the scenarios that have a continuous demand of users' visual attention on a certain area or direction of the VE. It is interesting to further evaluate these gestures in the case that the primary task would follow users' head movements; for example, performing selection in a menu which can be attached to users' view. Nevertheless, it is also possible to avoid these gestures because move forward/backward and roll left/right are still applicable to cases where the virtual objects move along with users' head motions/view.

In summary, our empirical comparison of the eight head gestures in a dual-hand line-drawing task point to four most suitable gestures to do mode-switching in VR HMD. They are move forward, move backward, roll left and roll right. These gestures have few error rates and better user's ratings and rankings compared to other head gestures.

There are two aspects to note about Study One: (1) we only tested horizontal line-drawing to avoid long experiments which would lead to extra fatigue; and (2) while we did not asked participants not to move, they completed the tasks in a relatively static position and moved only for adjusting the distance between them and spheres. These two aspects led us to use a 3D-painting application to test these four gestures because it would involve omni- and multi-directional line-drawing and locomotion. Thus, with the initial insights from Study One, we conducted Study Two to further assess the four head gestures in a real VR painting application and investigate their applicability and integration in a natural and representative task.

5 STUDY TWO: INTEGRATION OF THE HEAD GESTURES AND THEIR APPLICABILITY IN AN ACTUAL APPLICATION

The second study aimed at evaluating how the four head gestures can be integrated in an actual application to allow users to switch modes while doing a somewhat more realistic task. We integrated the head gestures (move forward, move backward, roll left, and roll right) into a 3D-painting VR application to allow quick switches between the painting tools. We collected participants' subjective feedback to assess the gestures' performance, suitability, and usability.

5.1 Participants and Apparatus

A total of 12 unpaid participants (5 females) were recruited to do this experiment. They were aged between 19 and 29 years ($M = 22.92$, $SD = 3.15$). All of them were right-handed. Five of them draw or paint frequently (i.e., at least once a month). Nine participants were regular VR users and four reported prior experience of drawing in VR. We used the same apparatus as in the previous study.

5.2 Application Scenario

We used a 3D-painting scenario to further evaluate the four head-based mode-switching gestures. An open-source application, Tilt Brush[1] was used to provide the virtual painting environment. Participants were requested to draw a courtyard in 3D space. We suggested that participants draw a sun, a house, and a tree but the actual appearance of these objects can be freeform (see Fig. 8a,b). The controller held by dominant hand was for drawing and non-dominant one was a tool panel, which had three-cornered interfaces including *Tools*, *Color Picker*, and *Brushes*, as shown in Fig. 8c. During the drawing process, participants were free to use the built-in tools on these interfaces. We integrated the four head gestures into the application. They could be used to achieve quick switch between drawing and erasing, and between drawing free lines and drawing straight lines; that is, switch to use an eraser (and back to drawing) or to draw with the help of a straight edge (and back to free drawing). Both eraser and straight edge are placed on *Tools* interface by default. We offered participants an option to choose two preferred head gestures for the just-mentioned two types of switching after experiencing and comparing them in the VE. The integrated head gestures did not replace the default operation. In other words, participants could still switch the mode by selecting the options from the tool panel, if they wanted to (see Fig. 8c).

5.3 Design and Procedure

We first collected participants' demographic and background information via a questionnaire. Then, participants were briefed about the research aim and introduced to painting in VR using Tilt Brush

---

[1] https://www.tiltbrush.com/

	Head Gesture	Frequencies	Median	Friedman
Ease of Learning (hard - easy)	Move Forward		5	
	Move Backward		4.5	
	Pitch Up		4	
	Pitch Down		4	$\chi^2(7) = 13.451,$
	Yaw Left		4	$p = .062$
	Yaw Right		4	
	Roll Left		4	
	Roll Right		5	
Ease of Use (hard - easy)	Move Forward		4	
	Move Backward		4	
	Pitch Up		4	
	Pitch Down		4	$\chi^2(7) = 15.048,$
	Yaw Left		4	$p = .035*$
	Yaw Right		3.5	
	Roll Left		4	
	Roll Right		4	
Accuracy (low - high)	Move Forward		4	
	Move Backward		4	
	Pitch Up		3	
	Pitch Down		3	$\chi^2(7) = 21.272,$
	Yaw Left		3	$p = .003**$
	Yaw Right		3	
	Roll Left		4	
	Roll Right		3	
Speed (slow - fast)	Move Forward		4	
	Move Backward		4	
	Pitch Up		3.5	
	Pitch Down		4	$\chi^2(7) = 20.336,$
	Yaw Left		3	$p = .005**$
	Yaw Right		3	
	Roll Left		4	
	Roll Right		4	
Neck Fatigue (high - low)	Move Forward		2.5	
	Move Backward		3	
	Pitch Up		2.5	
	Pitch Down		4	$\chi^2(7) = 8.926,$
	Yaw Left		2	$p = .258$
	Yaw Right		2	
	Roll Left		3	
	Roll Right		3	
VR Sickness (high - low)	Move Forward		4	
	Move Backward		4	
	Pitch Up		4	
	Pitch Down		4	$\chi^2(7) = 5.938,$
	Yaw Left		4	$p = .547$
	Yaw Right		4	
	Roll Left		4	
	Roll Right		4	

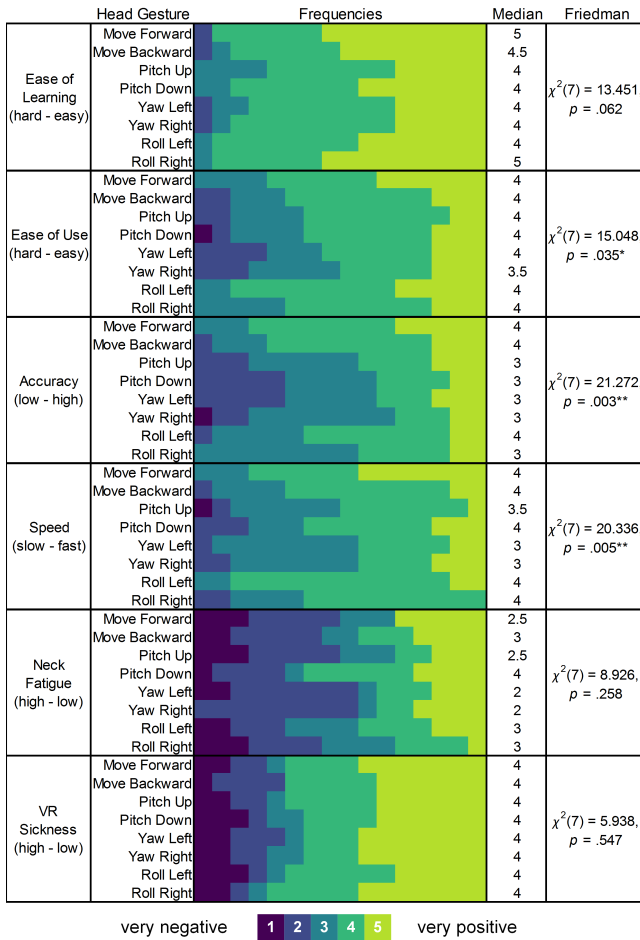very negative [1 2 3 4 5] very positive

Figure 6: Subjective ratings and Friedman test results, indicating qualitative differences between the head gestures (* and ** represent significant results at .05, and .01, respectively).
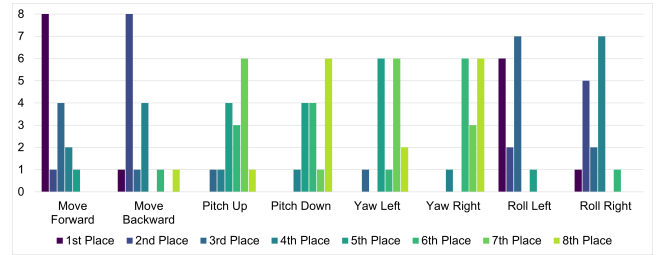

Figure 7: Distribution of participants' post rankings of the head gestures.

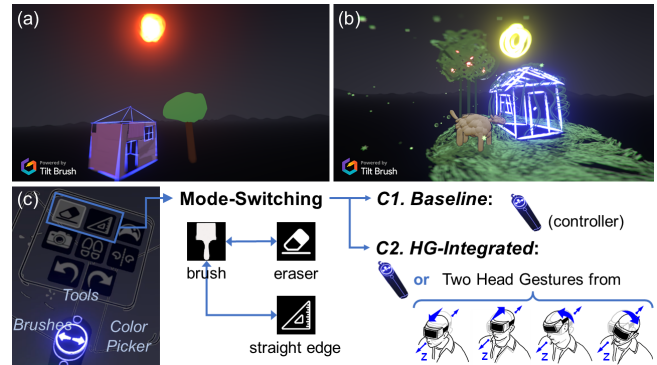

Figure 8: (a) and (b) Two art work from participants (P6, P9). (c) Experimental conditions. *C1.Baseline*: use controller to switch. *C2.HG-Integrated*: allow quick switch using head gestures. Note that eraser and straight edge are both located in *Tools* interface.

with a short tutorial about the usage of its common tools. They got at most 10 minutes to familiarize themselves with the operations in the VE. After this, we introduced the four head gestures to them. Participants could try the head gestures and use them to switch between tools in the application. They needed to select two of gestures that they thought practical and suitable for switching between modes. Then the formal experiment started. It had two sessions, one using the non-modified version of the application (*Baseline*), and another using the version integrated with head gestures (*HG-Integrated*). Their order was alternated. Each session lasted about 10 minutes, where participants could make their own creation that were related to the topic of a courtyard. After completing each session, we gave them a short version of User Experience Questionnaire (UEQ) [24]. At the end of the experiment, we conducted a structured interview asking their experience with the two versions.

5.4 Results and Discussion

5.4.1 UEQ results

Table 2 summarizes the results from UEQ in terms of pragmatic, hedonic, and overall quality. The average scores in *HG-Integrated* are all higher than in *Baseline*. Results from Wilcoxon signed rank tests showed that the differences in UEQ scores between *HG-Integrated* and *Baseline* were significant in terms of *overall* quality ($Z = -2.394$, $p = .017$) and *hedonic* quality ($Z = -2.937$, $p = .003$), but not significant in *pragmatic* quality ($p > .05$). The results indicate that

using head gestures to support mode-switching in the VR drawing application provides better user experience.

5.4.2 Selection of Head Gestures

In this study, we allowed participants to select their two preferred head gestures and match them with the two mode-switching process after experiencing them in the application. Such gesture-to-action mapping represents a sense of customization that is commonly seen in software applications.

Most participants ($N = 9$) selected paired gestures. Seven of them chose roll left/right for eraser or straight edge. Four commented that move forward/backward were difficult or uncomfortable to perform while drawing. Two participants (P3, P6) worried that they might lose their balance while performing or maintaining move forward/backward. On the other hand, two participants (P4, P7) chose move forward/backward. P4 mentioned that: *"move forward/backward were more intuitive and easier to perform."* While P7 felt hard to maintain roll left/right.

Three participants (P1, P8, P10) selected mixed settings. They all chose roll left for straight edge and move forward for eraser. They felt roll left/right might affect drawing due to the rotation of the vision, and thus did not use them for straight edge. In addition, they thought roll left was more natural and intuitive compared to roll

Table 2: Mean and standard deviation of the results from UEQ for pragmatic, hedonic, and overall quality. Higher scores indicate better user experience.

Condition	Pragmatic	Hedonic	Overall
Baseline	$0.29_{\pm1.25}$	$-0.29_{\pm1.19}$	$0_{\pm1.13}$
HG-Integrated	$1.06_{\pm1.00}$	$1.42_{\pm0.71}$	$1.24_{\pm0.72}$

right. They commented that move forward was "*more convenient*" and "*easier*" than move backward. In general, these comments were in line with the subjective feedback from Study One—users have different perceptions on the head gestures, leading to unique preferences.

5.4.3 Participants' Comments

All 12 participants commented that they felt mode-switching in *HG-Integrated* was more efficient than *Baseline*. They used "*continuous*" and "*smooth*" to describe the drawing with the former, and "*cumbersome*" and "*interruptive*" for the baseline. The head-based gestures improved the mode-switching procedure by saving the time to locate the target mode and default mode before and after using it. Participants (P6, P9, P10) noted that head-based mode-switching was particularly efficient if the current interface was Brushes or Color Picker (see Fig. 8), in which by default an additional switch between the upper-level interfaces was required. P2 and P4 mentioned that direct switching using head gestures helped to get more chances for resting their non-dominant hands, reducing arm fatigue.

Almost all participants ($N = 11$) agreed that learning the head-based mode-switching approach in the drawing scenario was effortless. This extends the results we obtained from Study One showing that the head gestures for mode-switching were easy to learn (see Fig. 6). P5 said that the customization of the head gestures before the formal experiment improved the learnability and memorability.

Some participants ($N = 6$) believed the experience and the efficiency of using head gestures to do mode switching in the application could be further improved if the gestures can be more resistant to false positives, though they did not happen frequently as we only observed this in a handful of cases. To complete the 3D-painting task, participants inevitably needed to walk around, which we allowed them to do, leading to more locomotion than standing without much motion like in Study One, though we did not specifically asked them not to move. This shows that head gestures can work reliably even when participants are in motion. Also, three participants (P4, P7, P12) suggested having visual cues so that they could know the switching of modes. Providing visual cues, voice prompts, or tactile feedback can help users be aware of the current mode.

Two participants (P3, P10) thought their judgement or estimation of objects in some application scenarios might be influenced while performing roll left/right. They also thought that with some more practice, this should not be a major issue. Similarly, P5 and P11 commented that head gestures may be more difficult to perform when they need to stand on their toes or squat. They said that these activities may not be frequent in VR applications and, at the end, would not affect the main usefulness of head gestures to do mode switching.

6 SUMMARY AND DESIGN RECOMMENDATIONS

In this section, we summarize the results from the two user studies, and distill design recommendations for incorporating head-based gestures in VR HMD.

Head gestures can support mode switching in VR. We derived four head gestures that can be suitable for mode-switching in VR from the first controlled experiment. To understand their suitability and usability for actual applications, we conducted a second user study where these gestures were embedded in an existing application. Based on the results from two studies, incorporating head gestures for mode-switching in VR is an efficient hands-free approach to provide a "continuous" and "smooth" experience. Given that head gestures can be detected effectively and performed easily [39, 42], they complement other modalities and tasks.

Head gestures should be customized by end-users. We noticed that participants had different preferences on head gestures from the first study (see Fig. 7). As such, we gave participants in the second study opportunities to try and select preferred gestures based on their experience before the actual use. We received positive feedback about customizing their own gestures. Therefore, allowing users to select their preferred gestures would be helpful for improving their learnability and usability.

Selecting suitable head gestures according to the use case. In the first study, we found that move forward/backward had better performance than roll left/right and were more preferred by participants for mode-switching in VR. Despite this, a small number of participants in the second study selected to help them make mode switches. This choice could be due to the different nature of the tasks in the experiments. The more controlled setting in the first study represent a single task, while the painting task in the second study involved multitasking which was more dynamic and complex. In this sense, designers should consider how the unique features of the head gestures might affect user experience in a specified application.

7 LIMITATIONS AND FUTURE WORK

This research has the following limitations, which could represent avenues for future work. First, we used a simple strategy for sensing and capturing head gestures. Though it had high accuracy, it can be further improved by optimizing the algorithm and testing different sensing parameters. This was not the concern of the current study but it could be relevant for a wider adoption of head gestures in other types of HMD. Second, we only focused on a user-maintained mechanism for mode-switching. Though a system-maintained mechanism may be error prone [9, 25], it could still be worthy to examine its comparative performance and usability. Furthermore, it could be also interesting to design and evaluate a hybrid method for real applications. Third, we mainly investigated head-based switching between two modes. While some scenarios could involve more than two modes, the two paired gestures could cover a wider range of practical scenarios in VR applications. In the future, we plan to explore the feasibility and applicability of having more than two modes for each pair of head gestures. One approach is to have multiple modes placed in a continuum. By using gestures like move forward/backward, one can traverse through the available modes and returning to a neutral position represents the selection of the mode. Forth, we focused on exploring and evaluating different head gestures for hands-free mode switching. We plan to conduct user studies to further compare the usability of head-based mode switching with other hands-free approaches (like gaze) in the future.

8 CONCLUSION

In this paper, we explored the use of head gestures for mode switching in Virtual Reality head-mounted displays (VR HMD). We first derived head gestures that are potentially suitable based on three design criteria. We then conducted a user study to systematically compare their efficiency, accuracy, and usability. In a second study, we further investigated the usability of the four most usable head gestures in an actual 3D-painting application. Results from two studies show that move forward, move backward, roll left and roll right offer relatively better performance and received the best user ratings. They also show that they can be easily integrated into an application and learned by users. Based on our results, we conclude that head gestures can support mode switching in VR and represent a suitable approach that is hands-free, fast, accurate, and low-cost in current VR HMD. Our work contributes a first systematic exploration of using head gestures for mode switching in VR HMD, a set of suitable gestures, and several design recommendations for their use in VR applications.

REFERENCES

[1] R. Atienza, R. Blonna, M. I. Saludares, J. Casimiro, and V. Fuentes. Interaction techniques using head gaze for virtual reality. In *Proc. TENSYMP*, pp. 110–114, 2016. doi: 10.1109/TENCONSpring.2016.7519387

[2] J. Blattgerste, P. Renner, and T. Pfeiffer. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In *Proc. COGAIN*. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3206343.3206349

[3] D. L. Chen, R. Balakrishnan, and T. Grossman. Disambiguation techniques for freehand object manipulations in virtual reality. In *Proc. IEEE VR*, pp. 285–292. IEEE Press, New York, NY, USA, 2020. doi: 10.1109/VR46266.2020.00048

[4] R. F. Dillon, J. D. Edey, and J. W. Tombaugh. Measuring the true cost of command selection: Techniques and results. In *Proc. CHI*, pp. 19–26. Association for Computing Machinery, New York, NY, USA, 1990. doi: 10.1145/97243.97247

[5] B. M. Ens, R. Finnegan, and P. P. Irani. The Personal Cockpit: A spatial interface for effective task switching on head-worn displays. In *Proc. CHI*, pp. 3171–3180. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2556288.2557058

[6] A. Esteves, D. Verweij, L. Suraiya, R. Islam, Y. Lee, and I. Oakley. SmoothMoves: Smooth pursuits head movements for augmented reality. In *Proc. UIST*, pp. 167–178. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3126594.3126616

[7] Y. Felberbaum and J. Lanir. Better understanding of foot gestures: An elicitation study. In *Proc. CHI*, pp. 1–12. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173908

[8] I. Giannopoulos, A. Komninos, and J. Garofalakis. Natural interaction with large map interfaces in VR. In *Proc. PCI*. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3139367.3139424

[9] K. Hinckley, F. Guimbretiere, P. Baudisch, R. Sarin, M. Agrawala, and E. Cutrell. The Springboard: Multiple modes in one spring-loaded control. In *Proc. CHI*, pp. 181–190. Association for Computing Machinery, New York, NY, USA, 2006. doi: 10.1145/1124772.1124801

[10] J. Kabbash, W. Buxton, and A. Sellen. Two-handed input in a compound task. In *Proc. CHI*, pp. 417–423. Association for Computing Machinery, New York, NY, USA, 1994. doi: 10.1145/259963.260425

[11] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billinghurst. Pinpointing: Precise head- and eye-based target selection for augmented reality. In *Proc. CHI*, pp. 1–14. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3173574.3173655

[12] J. J. J. LaViola, E. Kruijff, R. P. McMahan, D. A. Bowman, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, USA, 2017.

[13] Y. Li, K. Hinckley, Z. Guan, and J. A. Landay. Experimental analysis of mode switching techniques in pen-based user interfaces. In *Proc. CHI*, pp. 461–470. Association for Computing Machinery, New York, NY, USA, 2005. doi: 10.1145/1054972.1055036

[14] D. Lindlbauer, A. M. Feit, and O. Hilliges. Context-aware online adaptation of mixed reality interfaces. In *Proc. UIST*, pp. 147–160. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3332165.3347945

[15] X. Lu, D. Yu, H.-N. Liang, and J. Goncalves. iText: Hands-free text entry on an imaginary keyboard for augmented reality systems. In *Proc. UIST*. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474788

[16] X. Lu, D. Yu, H.-N. Liang, W. Xu, Y. Chen, X. Li, and K. Hasan. Exploration of hands-free text entry techniques for virtual reality. In *Proc. ISMAR*, pp. 344–349. IEEE Press, New York, NY, USA, 2020. doi: 10.1109/ISMAR50242.2020.00061

[17] K. Minakata, J. P. Hansen, I. S. MacKenzie, P. Bækgaard, and V. Rajanna. Pointing by gaze, head, and foot in a head-mounted display. In *Proc. ETRA*. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3317956.3318150

[18] C. Park, H. Cho, S. Park, Y.-S. Yoon, and S.-U. Jung. HandPoseMenu: Hand posture-based virtual menus for changing interaction mode in 3D space. In *Proc. ISS*, pp. 361–366. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3343055.3360752

[19] D. Pearce and H.-G. Hirsch. The aurora experimental framework for the performance evaluation of speech recognition systems under noisy conditions. In *Proc. ICSLP*. International Speech Communication Association, 2000.

[20] K. Pfeuffer, L. Mecke, S. Delgado Rodriguez, M. Hassib, H. Maier, and F. Alt. Empirical evaluation of gaze-enhanced menus in virtual reality. In *Proc. VRST*. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3385956.3418962

[21] Y. Y. Qian and R. J. Teather. The eyes don't have it: An empirical comparison of head-based and eye-based selection in virtual reality. In *Proc. SUI*, pp. 91–98. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3131277.3132182

[22] J. Raskin. *The Humane Interface: New Directions for Designing Interactive Systems*. ACM Press/Addison-Wesley Publishing Co., USA, 2000.

[23] D. Rudi, I. Giannopoulos, P. Kiefer, C. Peier, and M. Raubal. Interacting with maps on optical head-mounted displays. In *Proc. SUI*, pp. 3–12. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2983310.2985747

[24] M. Schrepp, A. Hinderks, and J. Thomaschewski. Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6):103–108, 2017. doi: 10.9781/ijimai.2017.09.001

[25] A. J. Sellen, G. P. Kurtenbach, and W. A. Buxton. The prevention of mode errors through sensory feedback. *Human–Computer Interaction*, 7(2):141–164, 1992. doi: 10.1207/s15327051hci0702_1

[26] L. Sidenmark and H. Gellersen. Eye&Head: Synergetic eye and head movement for gaze pointing and selection. In *Proc. UIST*, pp. 1161—-1174. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3332165.3347921

[27] L. Sidenmark, D. Potts, B. Bapisch, and H. Gellersen. *Radi-Eye: Hands-Free Radial Interfaces for 3D Interaction Using Gaze-Activated Head-Crossing*. Association for Computing Machinery, New York, NY, USA, 2021.

[28] J. Smith, I. Wang, W. Wei, J. Woodward, and J. Ruiz. Evaluating the scalability of non-preferred hand mode switching in augmented reality. In *Proc. AVI*. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3399715.3399850

[29] J. Smith, I. Wang, J. Woodward, and J. Ruiz. Experimental analysis of single mode switching techniques in augmented reality. In *Proc. GI*. Canadian Human-Computer Communications Society, Waterloo, CAN, 2019. doi: 10.20380/GI2019.20

[30] H. B. Surale, F. Matulic, and D. Vogel. Experimental analysis of mode switching techniques in touch-based user interfaces. In *Proc. CHI*, pp. 3267–3280. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025865

[31] H. B. Surale, F. Matulic, and D. Vogel. Experimental analysis of barehand mid-air mode-switching techniques in virtual reality. In *Proc. CHI*, pp. 1–14. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300426

[32] S. Tregillus, M. Al Zayer, and E. Folmer. Handsfree omnidirectional VR navigation using head tilt. In *Proc. CHI*, pp. 4063–4068. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025521

[33] H. Tu, X.-D. Yang, F. Wang, F. Tian, and X. Ren. Mode switching techniques through pen and device profiles. In *Proc. APCHI*, pp. 169–176. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2350046.2350081

[34] J. von Willich, M. Schmitz, F. Müller, D. Schmitt, and M. Mühlhäuser. Podoportation: Foot-based locomotion in virtual reality. In *Proc. CHI*, pp. 1–14. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376626

[35] W. Xu, H.-N. Liang, A. He, and Z. Wang. Pointing and selection methods for text entry in augmented reality head mounted displays. In

Proc. ISMAR, pp. 279–288. IEEE Press, New York, NY, USA, 2019. doi: 10.1109/ISMAR.2019.00026

[36] W. Xu, H.-N. Liang, Y. Zhao, D. Yu, and D. Monteiro. *DMove: Directional Motion-Based Interaction for Augmented Reality Head-Mounted Displays*, pp. 1–14. Association for Computing Machinery, New York, NY, USA, 2019.

[37] W. Xu, H.-N. Liang, Y. Zhao, T. Zhang, D. Yu, and D. Monteiro. Ring-Text: Dwell-free and hands-free text entry for mobile head-mounted displays using head motions. *IEEE Transactions on Visualization and Computer Graphics*, 25(5):1991–2001, 2019. doi: 10.1109/TVCG. 2019.2898736

[38] Y. Yan, Y. Shi, C. Yu, and Y. Shi. HeadCross: Exploring head-based crossing selection on head-mounted displays. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(1), Mar. 2020. doi: 10.1145/3380983

[39] Y. Yan, C. Yu, X. Yi, and Y. Shi. HeadGesture: Hands-free input approach leveraging head movements for HMD devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(4), dec 2018. doi: 10.1145/3287076

[40] S. Yi, Z. Qin, E. Novak, Y. Yin, and Q. Li. GlassGesture: Exploring head gesture interface of smart glasses. In *Proc. IEEE INFOCOM*, pp. 1–9. IEEE Press, New York, NY, USA, 2016. doi: 10.1109/INFOCOM .2016.7524542

[41] C. Yu, Y. Gu, Z. Yang, X. Yi, H. Luo, and Y. Shi. Tap, Dwell or Gesture? Exploring head-based text entry techniques for HMDs. In *Proc. CHI*, pp. 4479–4488. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3025453.3025964

[42] D. Yu, H.-N. Liang, X. Lu, T. Zhang, and W. Xu. DepthMove: Leveraging head motions in the depth dimension to interact with virtual reality head-worn displays. In *Proc. ISMAR*, pp. 103–114. IEEE Press, New York, NY, USA, 2019. doi: 10.1109/ISMAR.2019.00-20

[43] D. Yu, X. Lu, R. Shi, H.-N. Liang, T. Dingler, E. Velloso, and J. Goncalves. Gaze-supported 3D object manipulation in virtual reality. In *Proc. CHI*. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445343