International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Novel Pointing Method Based on Bare-Hand Interactions Using the Palmar Surface In XR

Su Young Kim[1], Yu Jin Lee[2], Yoon Sang Kim*[3]

Submitted: 06/12/2023 **Revised**: 17/01/2024 **Accepted**: 27/01/2024

Abstract: Much of the current content (photos, videos, etc.) is provided through windows, which comprise a 2D graphical user interface. In extended reality (XR), it is common to render such content on a 3D plane (hereafter 3D window). XR, an emerging interactive environment, has been developed to support bare-handed input, and studies on interacting with 3D windows in this environment have primarily focused on pointing methods. However, conventional methods can cause various issues such as fatigue and discomfort owing to the use of mid-air gestures. In consideration of these problems, this paper proposes a novel pointing method. The proposed method augments a virtual pad on the palmar surface of one hand and uses the other hand to operate the pad for pointing. The pointing performance of the proposed method was evaluated by comparing it with three conventional representative pointing methods (Gaze&Gesture (GG), Handray&Gesture (HG), and virtual pad (VP)). From the analysis results, the proposed method performed better than VP but was inferior to GG and HG. However, qualitative analysis confirmed that the proposed method achieved accurate pointing, fast perceived operation speed, and low fatigue. It was also confirmed from the survey results on social acceptance that the proposed method is most preferred in public environments.

Keywords: *augmented reality, bare-hand interaction, extended reality, pointing*

1. Introduction

A window is composed of a 2D graphical user interface, and is defined as an open file or program [1] (hereafter 2D window). 2D windows have traditionally been used in many systems owing to their many advantages [2], such as familiarity, intuition, efficiency, and high accuracy. One of the typical input methods for interacting with 2D windows is the pointing method [3].

Recent technological advances have led to the emergence of a new interactive environment called extended reality (XR). XR is defined as an umbrella term that includes virtual reality (VR), augmented reality (AR), and mixed reality (MR) [4-6]. As mentioned in [7], much of current content is based on 2D windows, so research has been conducted to provide such content in XR as well. In XR, it is common to render this content on a 3D plane [2, 8] (hereafter 3D window).

Equipment to provide XR has been developed to head-mounted display (HMD) devices using bare-handed interactions [9-11] (hereafter XR HMD). Similar to 2D windows, research on interaction with 3D windows using bare-hands has primarily focused on pointing methods. However, most studies on pointing in XR have utilized mid-air gestures, which can cause various issues, such as user fatigue [11, 12], social awkwardness [11, 13], and discomfort due to the absence of physical haptic feedback [14].

This paper proposes a novel pointing method for XR that considers these issues. The proposed method can reduce the range of mid-air gestures by indirectly controlling the cursor through a virtual touch pad. It also provides physical haptic feedback by augmenting it on the palmar surface of a hand.

This paper is organized as follows. In Section 2, we introduce previous pointing studies based on bare-handed interactions in XR environments. Section 3 describes the proposed method in detail. Section 4 evaluates the performance of the proposed method by conducting a comparative experiment with conventional methods, and analyzes the experimental results. Finally, followed by conclusions in Section 5.

2. Related Works

Pointing-related studies in XR have primarily focused on two types: 1) a method using a ray projected in the direction of gaze from an HMD for pointing, and using gestures to handle events such as selection actions [15-17] (hereafter Gaze&Gesture), and 2) a method using a ray emitted from the hand for pointing, and using gestures to handle events [15, 17-19] (hereafter Handray&Gesture).

Mine (1995) introduced interaction methods to select distant objects [15]. The introduced interactions included

[1,2] BioComputing Lab, Department of Computer Science and Engineering, Korea University of Technology and Education (KOREATECH), Cheonan 31253, Republic of Korea
ORCID ID: [1] 0009-0002-9825-983X, [2] 0009-0005-9045-2303
[3] BioComputing Lab, Institute for Bioengineering Application Technology, Department of Computer Science and Engineering, KOREATECH, Cheonan 31253, Republic of Korea
ORCID ID: 0000-0002-0416-7938
** Corresponding Author Email: yoonsang@koreatech.ac.kr*

Gaze&Gesture, Handray&Gesture, and voice input. Poupyrev et al. (1996) proposed a pointing method based on Handray&Gesture [18]. The proposed method emitted a nonlinear ray from the hand for pointing. The results of the usability evaluation showed that the proposed method was able to naturally manipulate both near and far objects, but no performance comparison was conducted with other input methods. Pierce et al. (1997) proposed a pointing method that combined Gaze&Gesture and Handray&Gesture [19]. The proposed method emitted a ray from an HMD to between the thumb and index finger for pointing and handled events using gestures (the authors defined this method as a Head Crusher). The performance was positively evaluated by six users, but no performance comparison was conducted with other pointing methods. Brasier et al. (2020) proposed a pointing method using virtual pads augmented around the body [11]. Six types of virtual pads were implemented in the proposed method, considering two orientations (horizontal and vertical) and three augmented positions (wrist, waist, and thigh). The performance comparison experiment used six types of virtual pads and conventional methods (Gaze&Gesture, Handray&Gesture). From the comparison results, it is suggested that the proposed method is less fatiguing; the augmented position is recommended to be the wrist and waist; and the orientation should be the same as the screen orientation. Lystbæk et al. (2022) proposed a pointing method based on Gaze&Gesture [16]. The proposed method performed pointing as in Gaze&Gesture and then handled events when the gaze direction and the fingertip were aligned (the authors defined this method as Gaze&Finger). The authors further proposed a method, similar to Gaze&Finger, which used the tracked hand to compensate for pointing (they defined this method as Gaze&Hand). The performances of the two proposed methods were evaluated in comparison with two existing pointing methods (Gaze&Gesture, Handray&Gesture). The comparison results showed that the two proposed methods provide faster input than Handray&Gesture, and also showed that Gaze&Finger had relatively high accuracy. Wagner et al. (2023) proposed a pointing method that combined Gaze&Gesture and Handray&Gesture [17]. The proposed method performed pointing as in Gaze&Gesture and then handled events when the handray direction and the fingertip were aligned (the authors defined this method as Gaze&Finger). The proposed method was evaluated in comparison with four existing pointing methods (Head-Crusher, Gaze&Finger, Gaze&Gestrue, Handray&Gestrue). The comparison results showed that the proposed method had excellent overall performance and was preferred after Gaze&Gestrue.

In addition to the above-mentioned studies, various studies have been conducted on pointing in XR; however, there is insufficient research on reducing the range of mid-air gestures, such as [11]. Moreover, [11] did not provide physical haptic feedback, so an approach that considers both minimizing mid-air gestures and haptic feedback is required for pointing.

3. Proposed Pointing Method

A touch pad, also known as a track pad, is one of the most representative pointing devices and is widely used on laptops [20]. Owing to their intuitive usability, touch pads are used not only on laptops but also on a variety of controllers, such as remote controllers and VR controllers. The proposed method uses a virtual touch pad as an input device by augmenting it on the palmar surface (hereafter this virtual touch pad is referred to as a hand pad).

The hand pad in this paper was implemented using Microsoft's HoloLens 2, which is capable of tracking both the hands and fingers as shown in Fig. 1. Considering that most pointing devices, such as computer mouses, are designed for right-handed users, the hand pad is augmented on the left palmar surface. According to [9] and [21], using only the index finger was faster and had less errors than using all fingers when interacting with virtual objects. Therefore, the augmented hand pad interacts with only the right index finger. Fig. 2 shows the hand pad, which is a virtual input device, in the proposed method.

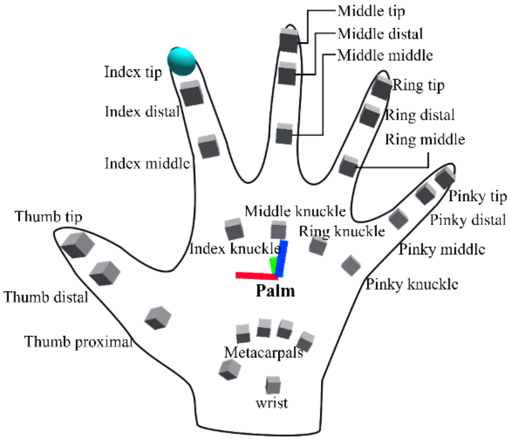


Fig. 1. Hand pose estimation of HoloLens 2.

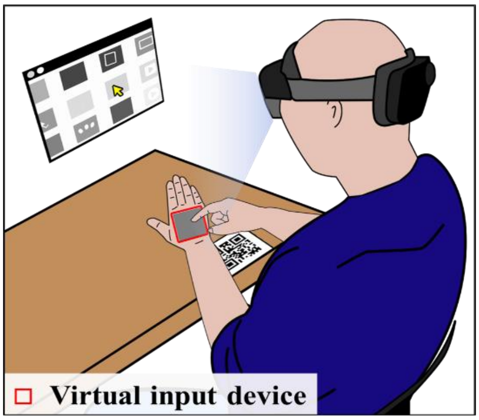


Fig. 2. Hand pad in the proposed method.

3.1.1. Personalized layout

Since virtual pads are commonly implemented as a rectangular plane [22-24], the hand pad utilizes joints corresponding to the corners of a rectangular shape formed by the metacarpals (the five bones in the palm), as shown in Fig. 3(a). The joints used are: the 2nd metacarpophalangeal (MCP) joint (hereafter a-joint), 5th MCP joint (hereafter b-joint), 1st carpometacarpal (CMC) joint (hereafter c-joint), and 5th CMC joint (hereafter d-joint). However, the palmar size varies from person to person, and the positions of the joints tracked by an XR HMD may not be accurate. To address these issues, the proposed method measures offsets by having users place their right index finger directly on each joint of the left hand. Fig. 3(b) shows the guided dialog used for this process.

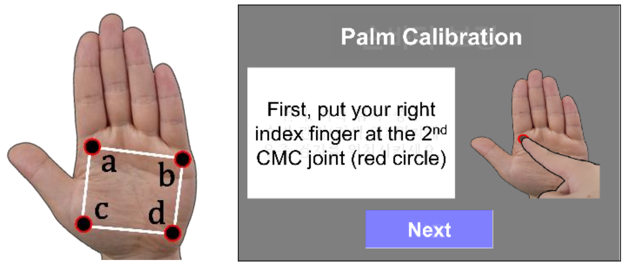


Fig. 3. Process of measuring offsets to calculate the rectangle in the palm: (a) positions to be input; (b) guided dialog.

Because the measured positions may not lie within a single plane, a fitting plane is calculated using ordinary least squares, as shown in equation (1) (the centroid of the palmar position is additionally measured and used to reduce distortion of the fitting plane, corresponding to e in (1)).

$$U = \begin{bmatrix} x_a & y_a & 1 \\ x_b & y_b & 1 \\ x_c & y_c & 1 \\ x_d & y_d & 1 \\ x_e & y_e & 1 \end{bmatrix}, V = \begin{bmatrix} z_a \\ z_b \\ z_c \\ z_d \\ z_e \end{bmatrix} \tag{1}$$

$$\therefore \begin{bmatrix} A \\ B \\ C \end{bmatrix} = pinv(U)V = (U^T U)^{-1} U^T V$$

where A, B, C are parameters of the plane equation satisfying $Ax + By + Cz + D = 0$.

Each corner of the rectangle (a-joint, b-joint, c-joint, and d-joint) is then projected onto the fitting plane to form a 2D coordinate system (hereafter, prime denotes the projected position). This 2D coordinate system uses the projected c-joint position as the origin and the following vectors as axes: a vector $\overrightarrow{c'a'}$ using the projected a-joint (hereafter y-axis), a normal $\overrightarrow{normal}$ to the fitting plane, and a cross product of $\overrightarrow{c'a'}$ and $\overrightarrow{normal}$ (hereafter x-axis). Fig. 4 shows the formed 2D coordinate system. The hand pad is generated using a bounding box based on the x-axis and y-axis to include the projected positions. Fig. 5 shows the generated hand pad.

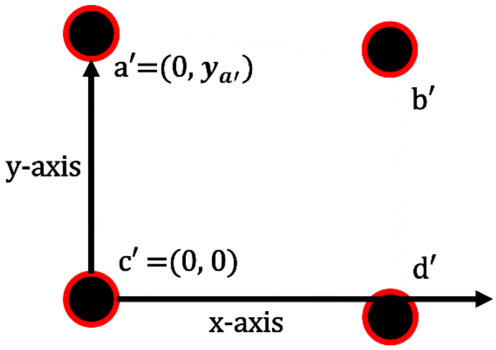


Fig. 4. Metacarpal-based rectangle projected into $\mathbb{R}^2$.

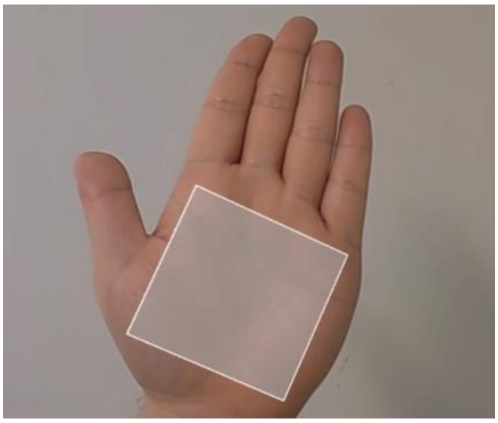


Fig. 5. Generated hand pad based on metacarpals.

3.1.2. Event handling

Commercial touch pads handle various events by combining two events (tap and slide), such as double tap (click) = tap + tap, drag = tap + slide, etc. The proposed method handles tap and slide events using the dwell time of the index finger touching the hand pad. The proposed method considers the hand pad and the finger to be touching if both of the following conditions are met: 1) when the hand pad is viewed in 2D, the position of the index fingertip is within the plane, and 2) this finger is within 1 cm in the normal direction from the plane, as in [25] (but only the inside direction is considered because the outside direction is blocked by the palmar surface).

A tap event is handled when the touch has a short dwell time. Hincapié-Ramos et al. (2015) set the dwell time to 150 ms for handling tap event in an AR HMD environment [26]. Liao et al. (2017) compared dwell times from 110 ms to 170 ms for handling tap even [27], and they confirm high accuracy at 150 ms and 170 ms. Both of these studies were conducted in environments where physical controllers were used, so it is possible to know the exact moment that the user inputs. However, because the XR environment interacts with virtual objects, it is difficult to determine the exact moment when dwell time occurs. Considering this, the proposed method uses a dwell time of 170 ms to handle the tap event. Fig. 6 shows an example of handling the tap event using the hand pad.

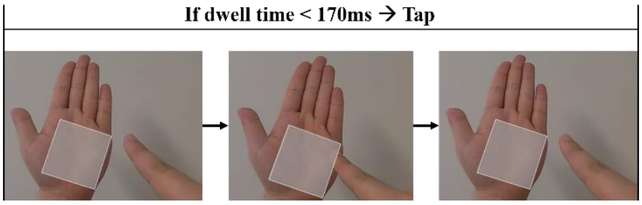


Fig. 6. Example of handling the tap event.

Touch pads are typically fixed to the physical environment and handle slide events by using the position of the finger being touched. However, the hand pad is augmented based on the hand and joint data being tracked, so handling the slide event on this pad based on simple position differences causes abrupt cursor movement. In the proposed method, the slide event is handled using a mean filter based on a queue to prevent this issue.

The detailed handling method is as follows. First, the difference in the touch position (hereafter △pos) between the previous and current frames is inserted into the queue, as shown in Fig. 7. If the time difference between the first input inserted into the queue and the current is less than 100 ms, the slide event is not handled. When the time difference is greater than 100 ms, the mean of all △pos values in the queue is used to handle the slide event and then dequeuing is operated, as shown in Fig. 8. When the slide event is handled in the proposed method, the cursor in the 3D window moves in proportion to the ratio of the mean △pos to the length of the hand pad (the ratio is applied on a 1/2 scale). For example, if the mean △pos corresponds to 1/2 of the hand pad width on the x-axis, the cursor moves 1/4 of the 3D window's width. If the mean △pos corresponds to 1/4 of the hand pad height on the y-axis, the cursor moves 1/8 of the 3D window's height.

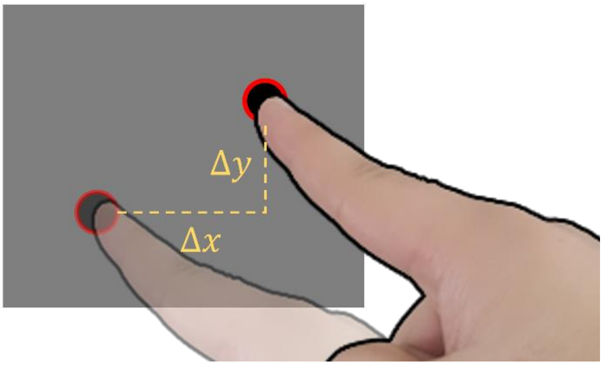


Fig. 7. △pos to insert into the queue.

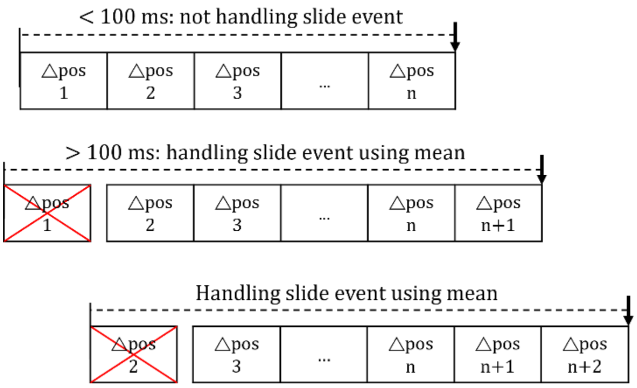


Fig. 8. Example of handling slide events based on mean filter.

4. Experiment and Results

4.1. Experimental environment

The performance of the proposed method was evaluated by comparing it with the conventional pointing methods. The XR HMD used in the experiment was HoloLens 2, which was used to implement the proposed method. The experimental procedure is shown in Fig. 9. The conventional pointing methods for comparison were two representative pointing methods (Gaze&Gesture and Handray&Gesture) and a pointing method using a virtual pad similar to [11] (hereafter Gaze&Gesture is referred to as GG, Handray&Gesture as HG, virtual pad as VP, and the proposed hand pad as HP).

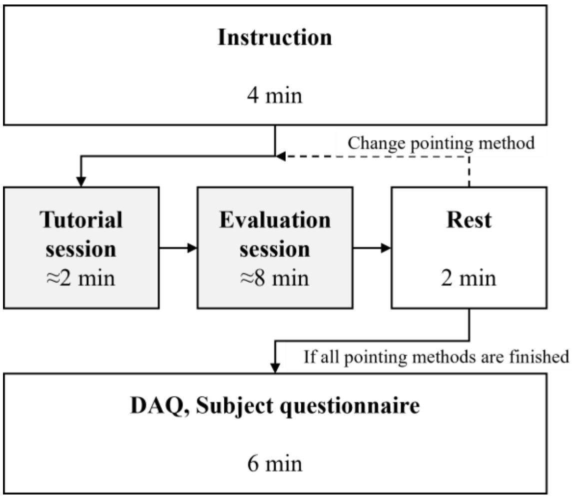


Fig. 9. Experimental procedure.

GG and HG were provided by the HoloLens 2. VP was implemented to have the same sizes of 183 mm × 103 mm as in [11], and for simple comparison, it was fixed at the lower field of view to always be sighted regardless of the user's gaze direction. In addition, to facilitate a comparison between VP and the proposed HP, tap event handling in VP was unified with HP.

Because the participants could adapt to the XR environment during the experiment, the order of the pointing methods was randomized. To compare the proposed method with the absolute coordinate virtual pad proposed in [11], the size of the 3D window, the pointing target, was set to 1.777:1 aspect ratio, which is the same as the virtual pad ratio in [11]. The 3D window was positioned 1 m from the participant to ensure sufficient distance and was sized (width × height) as (800 mm × 450 mm) for enough visibility (since the field of view of the HoloLens 2 is 43°× 29°). Fig. 10 shows the experimental environment.

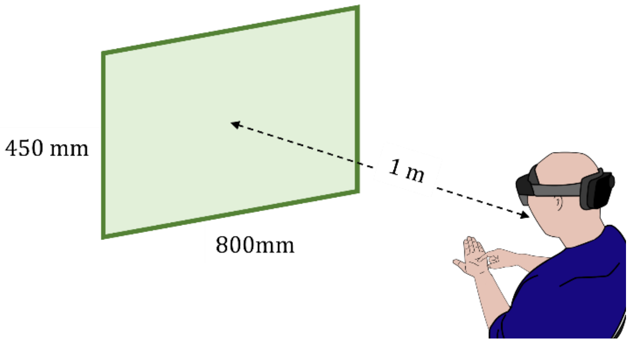


Fig. 10. Experimental environment.

A total of 20 participants was recruited through the intranet of the Korean National University of Technology and Education. Participants' eligibility was restricted to non-computer science students or first-year computer science students, who were expected to have less pointing experience in XR. The participants received experimental instructions for about 4 minutes and completed a consent form. After completing the consent form, participants were equipped with the XR HMD.

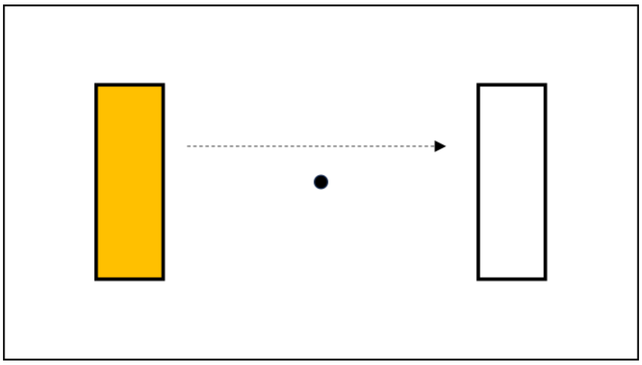
4.1.1. Tutorial sessions

Because the four pointing methods used in the experiment have distinct interaction mechanisms, the tutorial session was used to learn the current order of pointing methods. The learning for the pointing methods proceeded in the order of handling the slide event and then the tap event. Each event handling method was taught using two interactive circles displayed sequentially. For the slide event, the interaction was completed when the cursor touched the circle, and for the tap event, the interaction was completed when handling was performed while the cursor was inside the circle. When the participant completed interactions with a total of four circles, the session ended, and was proceeded to the evaluation session.

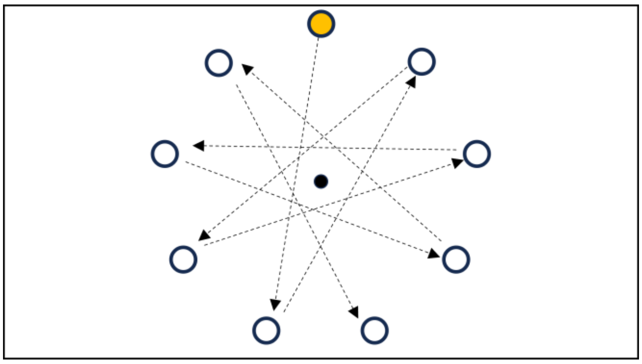
4.1.2. Evaluation sessions

The performance of pointing methods has typically been evaluated based on the ISO 9241-9. The evaluation session evaluated the performance of the pointing methods using two tasks (one- and multi-directions) defined in this standard document. Such tasks have primarily been designed using the index of difficulty (ID) from Fitts' law [28]. When the distance between the buttons is D and the width of a button is W, the ID is calculated as follows (unit: bits): $ID = \log_2(D/W + 1)$. Considering that previous pointing studies use the ID in the range of about 2 to 4 bits, we designed the tasks as shown in Fig. 11.

A detailed description of each task is as follows. The one-direction task is to tap the highlighted one among the buttons located on the left and right sides. In this experiment, the ID of this task was 2.807 bits. The participants were required to tap the highlighted button sequentially on the left and right as shown in Fig. 11(a), with a total of 10 repetitions (2 × 10 times). The multi-direction task is to tap the highlighted button among the circularly located buttons. In this experiment, the ID of this task was 3.821 bits, and it consisted of 9 buttons. Each button was highlighted in the order shown in Fig. 11(b), with a total of 2 repetitions (2 × 9 times).



(a)



(b)

Fig. 11. Two tasks in the evaluation session. Small black filled circles are the cursors, colored objects are highlighted: (a) One-direction task; (b) Multi-direction task

The performance of the pointing method using the tasks was evaluated using movement time (MT) and throughput (TP), which are metrics used in ISO 9241-9 to indicate the performance of the pointing method. MT represents the mean time taken to complete a task. In the experiment, MT was calculated as the mean of the time to tap the highlighted button in both one- and multi-direction tasks (20 trials and 18 trials). TP is a metric that indicates the efficiency of the task in Fitts' law and is calculated as follows: $TP = ID/MT$. Once the performance evaluation of the pointing method was completed, a 2 min rest was provided.

4.1.3. Questionnaires

The subjective performance of the pointing methods was evaluated using the device assessment questionnaire (DAQ) of ISO 9241-9 and a self-questionnaire to measure social acceptability. The DAQ is a questionnaire that evaluates factors such as force required, fatigue, and satisfaction with the input device on a 5-point Likert scale. The DAQ shown in Fig. 12(a) was used in the experiment.

The self-questionnaire was used to measure preferences for input methods when using the XR HMD in specific situations where social awkwardness might occur (hereafter this social acceptance self-questionnaire is referred to as SAQ). The SAQ, adapted from a subjective questionnaire that measures convenience in the public environment in [29], was created as shown in Fig. 12(b).

4.2. Experimental results

4.2.1. Movement time & Throughput

The evaluated MT is shown in Fig. 13, and the mean, standard deviation, and normality test (Shapiro Wilk test) results for each task are listed in Table 1. Since none of the pointing methods satisfied normality ($p<0.05$), the non-parametric Wilcoxon Signed-Rank test was used for the analysis.

In the one-direction task, HP took significantly longer than GG (difference(△)=6.481 s, W=5.0, p=1.907e-5) and HG (△=6.701 s, W=3.0, p=9.537e-6), and there was no significant difference with VP (W=91.0, p=0.622). In the multi-direction task, HP took significantly longer than GG (△=7.176 s, W=0.0, p=1.907e-6) and HG (△=5.232 s, W=24.0, p=0.001), but significantly faster than VP (△=-15.479 s, W=18.0, p=4.826e-6).

	Device Assessment	GG	HG	VP	HP
1	Force required for actuation was 1 2 3 4 5 too low … too high				
2	Smoothness during operation was 1 2 3 4 5 very rough … very smooth				
3	Mental effort required for operation was 1 2 3 4 5 too low … too high				
4	Physical effort required for operation was 1 2 3 4 5 too low … too high				
5	Accurate pointing was 1 2 3 4 5 easy … difficult				
6	Operation speed was 1 2 3 4 5 too fast … too slow				
7	Finger fatigue was 1 2 3 4 5 none … very high				
8	Wrist fatigue was 1 2 3 4 5 none … very high				
9	Arm fatigue was 1 2 3 4 5 none … very high				
10	Shoulder fatigue was 1 2 3 4 5 none … very high				
11	Neck fatigue was 1 2 3 4 5 none … very high				
12	General comfort was 1 2 3 4 5 very uncomfortable … very comfortable				
13	Overall, the input device was 1 2 3 4 5 very difficult … very high				

(a)

	Social Acceptance for mid-air gesture			
1	Imagine you're alone in a comfortable place, such as your home. Now, which of the four input methods (GG, HG, VP, HP) do you prefer when using XR HMD? Write down relative order of preference. Best … Worst			
2	Imagine you're in a room with close acquaintances. Some may be looking at you. Now, which of the four input methods do you prefer when using XR HMD? Write down relative order of preference. Best … Worst			
3	In the same situation as in 2, an acquaintance is mid-air gesturing through XR HMD. Imagine this acquaintance is pointing at you with bare-hand. Do you think you would feel uncomfortable in this situation? Write dow relative order of comfort. Comfortable … Uncomfortable			
4	Imagine you're in a public place, such as public transportation, with strangers. Some may be looking at you. Now, which of the four input methods do you prefer when using XR HMD? Write down relative order of preference. Best … Worst			
5	In the same situation as in 4, a stranger is mid-air gesturing through XR HMD. Imagine this stranger is pointing at you with bare-hand. Do you think you would feel uncomfortable in this situation? Write dow relative order of comfort. Comfortable … Uncomfortable			

(b)

Fig. 12. Questionnaires used in experiment: (a) DAQ; (b) SAQ.

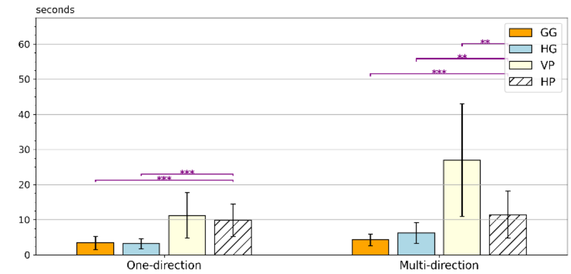


Fig. 13. Mean MT for all methods. Error bars indicate STD. *, **, and *** represent that p-value is less than 0.05, 0.01, and 0.001, respectively.

Table 1. Mean, STD, and normality results for MT.

Task	Method	M (s)	STD (s)	Shapiro Wilk	
				statistics	*p*
One	GG	3.393	1.929	0.633	6.557e-6
	HG	3.173	1.458	0.801	8.899e-4
	VP	11.228	6.459	0.813	0.001
	HP	9.874	4.631	0.775	3.778e-4
Multi	GG	4.257	1.704	0.771	3.308e-4
	HG	6.201	2.954	0.815	0.001
	VP	26.912	16.040	0.772	3.365e-4
	HP	11.432	6.666	0.696	3.514e-5

The evaluated TP is shown in Fig. 14, and the mean, standard deviation, and normality test results for each task are listed in Table 2. All pointing methods satisfied normality (p>0.05), so a paired t-test was used for the analysis.

In the one-direction task, HP was significantly lower than GG (△=-0.776 bps (bits per second), t=10.612, p=2.004e-9) and HG (△=-0.824 bps, t=9.978, p=5.450e-9), and there was no significant difference with VP (t=0.850, p=0.406). In the multi-direction task, HP was significantly lower than GG (△=-0.680 bps, t=9.992, p=5.328e-9) and HG (△=-0.477 bps, t=7.842, p=2.253e-7), but significantly higher than VP (△=0.196 bps, t=-5.242, p=4.639e-5).

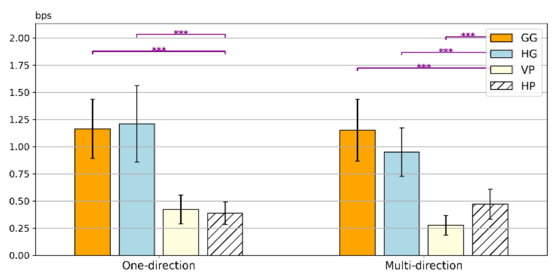


Fig. 14. Mean TP for all methods. Error bars indicate STD. *, **, and *** represent that p-value is less than 0.05, 0.01, and 0.001, respectively.

Table 2. Mean, STD, and normality results for TP.

Task	Method	M (bps)	STD (bps)	Shapiro Wilk	
				statistics	*p*
One	GG	1.163	0.272	0.960	0.536
	HG	1.211	0.353	0.976	0.879
	VP	0.422	0.135	0.977	0.884
	HP	0.387	0.104	0.909	0.061
Multi	GG	1.151	0.286	0.980	0.930
	HG	0.948	0.223	0.951	0.377
	VP	0.275	0.090	0.973	0.812
	HP	0.471	0.136	0.953	0.422

4.2.2. DAQ & SAQ

The four pointing methods across all the questions in the DAQ are shown in Fig. 15. As normality was not satisfied for all questions except HP in Q1 (p=0.053), HP in Q6 (p=0.078), and VP in Q8 (p=0.011), a non-parametric test was used for the analysis. The results of the comparison between HP and the other three pointing methods using the Wilcoxon Signed Rank test for each question are listed in Table 3.

Overall, it was confirmed that HP was more difficult to use than GG and HG but easier to use than VP (Q12, Q13). It was also confirmed that using a virtual pad, such as HP and VP, required more physical and mental effort than GG and HG, but it was perceived as more accurate and faster (Q3–Q6). In addition, HP caused less physical fatigue than GG and HG (Q7–Q11). Although VP caused less fatigue than HP across various fatigue aspects, most differences were not statistically significant.

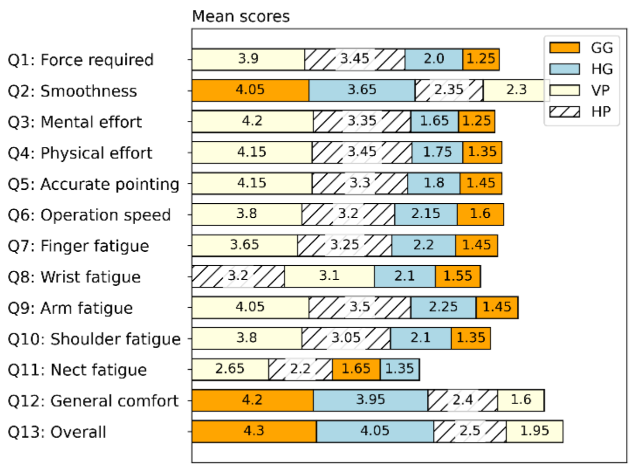


Fig. 15. Mean scores for the four pointing methods according to the DAQ (sorted left to right in descending order).

Table 3. Comparison of HP and Other three Pointing Methods in DAQ (blank values are not significant by p>0.05).

Question	GG		HG		VP	
	△	p	△	p	△	p
Q1	2.20	0.000	1.45	0.000	-0.45	0.048
Q2	-1.70	0.002	-1.30	0.003		0.922
Q3	2.10	0.000	1.70	0.000	-0.85	0.005
Q4	2.10	0.000	1.70	0.001	0.70	0.007
Q5	1.85	0.001	1.50	0.004	0.85	0.012
Q6	1.60	0.003	1.05	0.022		0.123
Q7	1.80	0.001	1.05	0.023	-0.40	0.035
Q8	1.65	0.001	1.10	0.014		0.589
Q9	2.05	0.000	1.25	0.006		0.100
Q10	1.70	0.001	0.95	0.020	-0.75	0.011
Q11	0.55	0.029	0.85	0.004		0.084
Q12	-1.80	0.000	-1.55	0.001	0.80	0.002
Q13	-1.80	0.000	-1.55	0.001	0.55	0.018

In the SAQ, a Friedman test was used to determine whether there was a preference order difference between the pointing methods (scored 1 for most preferred and 4 for least preferred). Significant differences in preference were confirmed for all questions (p_{Q_1}=1.072e-9, p_{Q_2}=3.038e-7, p_{Q_3}=0.004, p_{Q_4}=0.008, and p_{Q_5}=0.019). The percentages of the preferred methods for each question are shown in Fig. 16.

In a comfortable environment, GG and HG were generally preferred (Q1, Q2), but even in a comfortable environment, HP was preferred over HG when another user was pointing (Q3). In a public environment, it was confirmed that the proposed HP was the most preferred (Q4, Q5).

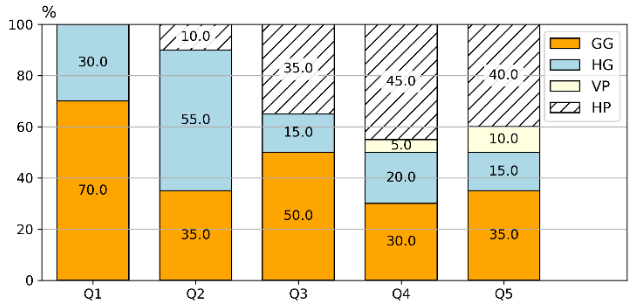


Fig. 16. The most preferred or comfortable method for each question in the SAQ.

5. Conclusion

This paper proposed a pointing method that used a small range of mid-air gestures. The proposed method virtually implemented a touch pad, a typical pointing device, and augmented it on the palmar surface of one hand. Pointing was performed with the other hand, and tap and slide events were handled through interactions with the virtual pad. The performance of the proposed method was evaluated using one- and multi-direction tasks of ISO 9241-9 and was compared with three conventional pointing methods (GG, HG, and VP). The analysis results confirmed that the proposed method performed better than VP but had lower performance than GG and HG. This may be attributed to the use of GG and HG in many studies and commercialized XR HMDs. However, the DAQ analysis results indicated that the proposed method was perceived to provide more accurate pointing and faster operation speed, and less fatigue than GG and HG. This means that the proposed method may reduce fatigue and be advantageous in complex manipulation scenarios. Moreover, in a public environment, the proposed method was confirmed to be the most preferred. Furthermore, considering the analysis results of SAQ, it suggests that the method proposed in this paper could contribute to the popularization of XR HMDs. This contribution means that it overcomes the major challenges of XR HMDs, such as privacy violations, and provides new usage possibilities.

Acknowledgements

This result was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE) (2021RIS-004). Also, this work was partially supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2023R1A2C2002838).

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Smith, D. C., Irby, C., Kimball, R., Verplank, B., & Harslem, E. (1982). Designing the Star user interface. Byte, 7, 242-282. [Online] Available: https://www.researchgate.net/publication/234781794_Designing_the_Star_user_interface_1982

[2] Ens, B., Hincapié-Ramos, J. D., & Irani, P. (2014). Ethereal planes: a design framework for 2D information space in 3D mixed reality environments. Proceedings of the 2nd ACM symposium on Spatial user interaction, 2-12. DOI: 10.1145/2659766.2659769

[3] Lin, J. L., Zheng, M. C., & Zhong, C. R. (2022). Research on the interactive behaviour of pointing devices: Is the Magic Mouse easy to use? 2022 IEEE International Conference on Consumer Electronics-Taiwan, 443-444. DOI: 10.1109/ICCE-Taiwan55306.2022.9869289

[4] Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. Telemanipulator and telepresence technologies, 2351, 282-292. DOI: 10.1117/12.197321

[5] Milgram, P., & Colquhoun, H. (1999). A taxonomy of real and virtual world display integration. Mixed reality: Merging real and virtual worlds, 1-26. DOI: 10.1007/978-3-642-87512-0_1

[6] Chuah, S. H. W. (2019). Wearable XR-technology: literature review, conceptual framework and future research directions. International journal of technology marketing, 13(3-4), 205-259. DOI: 10.1504/IJTMKT.2019.104586

[7] Lee, J. H., An, S. G., Kim, Y., & Bae, S. H. (2018). Projective windows: bringing windows in space to the fingertip. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1-8. DOI: 10.1145/3173574.3173792

[8] Feiner, S., MacIntyre, B., Haupt, M., & Solomon, E. (1993). Windows on the world: 2D windows for 3D augmented reality. Proceedings of the 6th annual ACM symposium on User interface software and technology, 145-155. DOI: 10.1145/168642.168657

[9] Kern, F., Niebling, F., & Latoschik, M. E. (2023). Text Input for Non-Stationary XR Workspaces: Investigating Tap and Word-Gesture Keyboards in Virtual and Augmented Reality. IEEE Transactions on Visualization and Computer Graphics, 29(5), 2658-2669. DOI: 10.1109/TVCG.2023.3247098

[10] Lee, T. H., & Lee, H. J. (2018). A new virtual keyboard with finger gesture recognition for AR/VR devices. Human-Computer Interaction. Interaction Technologies: 20th International Conference, 56-67. DOI: 10.1007/978-3-319-91250-9_5

[11] Brasier, E., Chapuis, O., Ferey, N., Vezien, J., & Appert, C. (2020). Arpads: Mid-air indirect input for augmented reality. 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 332-343. DOI: 10.1109/ISMAR50242.2020.00060

[12] Hincapié-Ramos, J. D., Guo, X., Moghadasian, P., & Irani, P. (2014). Consumed endurance: A metric to quantify arm fatigue of mid-air interactions. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1063-1072. DOI: 10.1145/2556288.2557130

[13] Zhang, Y., Kienzle, W., Ma, Y., Ng, S. S., Benko, H., & Harrison, C. (2019). ActiTouch: Robust touch detection for on-skin AR/VR interfaces. Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, 1151-1159. DOI: 10.1145/3332165.3347869

[14] Lindeman, R. W., Sibert, J. L., & Hahn, J. K. (1999). Towards usable VR: an empirical study of user interfaces for immersive virtual environments. Proceedings of the SIGCHI conference on Human factors in computing systems, 64-71. DOI: 10.1145/302979.302995

[15] Mine, M. R. (1995). Virtual environment interaction techniques. UNC Chapel Hill CS Dept. [Online] Available: https://www.researchgate.net/publication/2812583_Virtual_Environment_Interaction_Techniques

[16] Lystbæk, M. N., Rosenberg, P., Pfeuffer, K., Grønbæk, J. E., & Gellersen, H. (2022). Gaze-hand alignment: Combining eye gaze and mid-air pointing for interacting with menus in augmented reality. Proceedings of the ACM on Human-Computer Interaction, 6(ETRA), 1-18. DOI: 10.1145/3530886

[17] Wagner, U., Lystbæk, M. N., Manakhov, P., Grønbæk, J. E., Pfeuffer, K., & Gellersen, H. (2023). A Fitts' Law Study of Gaze-Hand Alignment for Selection in 3D User Interfaces. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 1-15. DOI: 10.1145/3544548.3581423

[18] Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). The go-go interaction technique: non-linear mapping for direct manipulation in VR. Proceedings of the 9th annual ACM symposium on User interface software and technology, 79-80. DOI: 10.1145/237091.237102

[19] Pierce, J. S., Forsberg, A. S., Conway, M. J., Hong, S., Zeleznik, R. C., & Mine, M. R. (1997). Image plane interaction techniques in 3D immersive environments. Proceedings of the 1997 symposium on Interactive 3D graphics, 39-43. DOI: 10.1145/253284.253303

[20] Akamatsu, M., & MacKenzie, I. S. (2002). Changes in applied force to a touchpad during pointing tasks. International Journal of Industrial Ergonomics, 29(3), 171-182. DOI: 10.1016/S0169-8141(01)00063-4

[21] Dudley, J. J., Benko, H., Wigdor, D., & Kristensson, P. O. (2019). Performance envelopes of virtual keyboard text input strategies in virtual reality. 2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 289-300. DOI: 10.1109/ISMAR.2019.00027

[22] Zhang, Z. (2003). Vision-based Interaction with Fingers and Papers. Proceedings International Symposium on the CREST Digital Archiving Project, 83-106. [Online] Available: https://www.microsoft.com/en-us/research/publication/vision-based-interaction-fingers-papers/

[23] Malik S., & Laszlo, J. (2004). Visual touchpad: a two-handed gestural input device. Proceedings of the 6th international conference on Multimodal interfaces, 289-296. DOI: 10.1145/1027933.1027980

[24] Edwin G., & Supriana, I. (2011). Hand detection for virtual touchpad. Proceedings of the 2011 International Conference on Electrical Engineering and Informatics, 1-5. DOI: 10.1109/ICEEI.2011.6021588

[25] Harrison, C., Benko, H., & Wilson, A. D. (2011). OmniTouch: wearable multitouch interaction everywhere. Proceedings of the 24th annual ACM symposium on User interface software and technology, 441-450. DOI: 10.1145/2047196.2047255

[26] Hincapié-Ramos, J. D., Ozacar, K., Irani, P. P., & Kitamura, Y. (2015). GyroWand: IMU-based raycasting for augmented reality head-mounted displays. Proceedings of the 3rd ACM Symposium on Spatial User Interaction, 89-98. DOI: 10.1145/2788940.2788947

[27] Liao, Y. C., Chen, Y. C., Chang, L., & Chen, B. Y. (2017). Dwell+ multi-level mode selection using vibrotactile cues. Proceedings of the 30th annual acm symposium on user interface software and technology, 5-16. DOI: 10.1145/3126594.3126627

[28] Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6), 381-391. DOI: 10.1037/h0055392

[29] Ahlström, D., Hasan, K., & Irani, P. (2014). Are you comfortable doing that? Acceptance studies of around-device gestures in and for public settings. Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services, 193-202. DOI: 10.1145/2628363.2628381