Int. J. Human–Computer Studies 192 (2024) 103346

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Object manipulation based on the head manipulation space in VR

Xiaolong Liu [a], Lili Wang [a,b,*], Wei Ke [c], Sio-Kei Im [c]

[a] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China
[b] Peng Cheng Laboratory, Shengzhen, China
[c] Faculty of Applied Sciences, Macao Polytechnic University, Macao Special Administrative Region of China

ARTICLE INFO

Keywords:
Virtual reality
Object manipulation
Hand-free
Head

ABSTRACT

Object manipulation is fundamental in virtual and augmented reality, where efficiency and accuracy are crucial. However, repetitive object manipulation tasks using the hands can lead to arm fatigue, and in some scenarios, hands may not be feasible for object manipulation. In this paper, we propose a novel approach for object manipulation based on head movement. Firstly, we introduce the concept of head manipulation space and conduct an experiment to collect head manipulation space data to determine the manipulable space. Then, we propose a new method for object manipulation based on head speed and inter-frame viewpoint quality to enhance the efficiency and accuracy of head manipulation. Finally, we design two user studies to evaluate the performance of our head-based object manipulation method. The results show that our method is feasible in terms of task completion efficiency and accuracy compared to state-of-the-art methods and greatly reduces user fatigue and motion sickness. Moreover, our method significantly improves usability and reduces task load. Our method lays a foundation for head-based object manipulation in virtual and augmented reality and provides a new manipulation method for scenarios where hands are not suitable for object manipulation.

1. Introduction

Object Manipulation (translation, rotation, and scaling) is an essential interaction task for users in both virtual and augmented reality. Efficiency and accuracy in object manipulation are crucial in virtual reality. Many researchers have studied the efficiency and accuracy of object manipulation in virtual reality, such as mid-air manipulation (Mendes et al., 2017) and PinNPivot manipulation (Gloumeau et al., 2020).

Objects Manipulation in virtual reality currently includes using controllers, gloves, gesture recognition (Rantamaa et al., 2023), and other technologies to simulate hand movements, as well as using hand tracking technology to track the user's hand movements directly. Hand tracking technology is widely used and can track the user's hand position and posture by using cameras or sensors and map this information to the virtual reality scene to manipulate objects with the hands. As a result, most interactions with VR applications are currently performed using a controller. This means that common interactions with VR systems are hindered and limited in scenarios where the user's hands are used to manipulate objects directly. For example, in a tourism application, users may drive a vehicle to explore a city and receive feedback through tangible interfaces. In this case, hands-free techniques are needed to manipulate objects within the tangible interface. In addition, there are currently two scenarios unsuitable for object manipulation with hands: (1) For some users with physical injuries, disabilities, or limited mobility, it may be very difficult to use their hands to manipulate. (2) In some special scenarios, users must perform other interactions with their hands, such as doctors performing surgery, holding surgical equipment, and cannot perform other manipulations. Therefore, researchers have recently proposed many hand-free methods (Monteiro et al., 2021), including head manipulation methods. However, there are two challenges for object manipulation based on head manipulation space: the first is that the range of head motion is relatively small, and how to perform large-scale object manipulation with a small range of head motion. The second is avoiding the frequent need to adjust the head posture when manipulating objects, which can cause physical tension and discomfort during the interaction. Manipulation gains (Liu et al., 2022) (translation gains, rotation gains, and scale gains), which allows the user to manipulate the object in VR with limited hand space. Manipulation gains based on head movements are also one of the solutions to this challenge.

To address these problems, we propose a hand-free object manipulation based on head manipulation space in VR to improve the efficiency and accuracy of manipulation. We introduce the concept of head manipulation space and conduct an experiment to collect head manipulation space data to determine the manipulable space. Then,

* Corresponding author at: State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China.
E-mail addresses: liuxiaolong186@buaa.edu.cn (X. Liu), wanglily@buaa.edu.cn (L. Wang).

https://doi.org/10.1016/j.ijhcs.2024.103346
Received 4 November 2023; Received in revised form 29 May 2024; Accepted 3 August 2024
Available online 22 August 2024

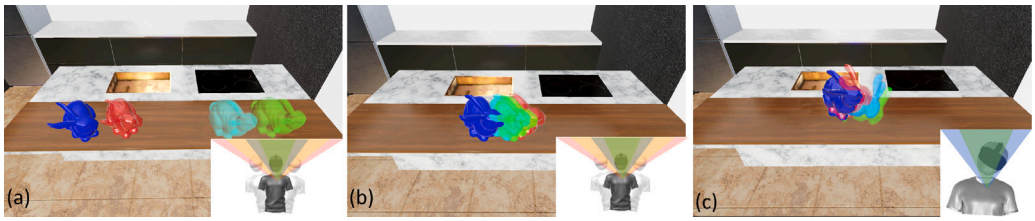

Fig. 1. The user is manipulating the bunny within the head manipulation space. Green indicates the target, blue indicates the original position of the component to be manipulated, red shows the result of manipulation with $gain = 1$, and cyan shows the result with our method. The lower right corner shows the head manipulation space. The translation and rotation manipulation spaces are different. For details, see Section 3.2 The user translates the bunny to the target in (a) and (b). In (a), the original bunny is far from the target, and the gain computed with our method is greater than 1, so we translate the component faster than the bunny with $gain = 1$. In (b), the original bunny is already very close to the target, and our adaptive gain is less than 1, so finer adjustments can be made to avoid the over manipulation problem caused by larger gains. In (c), the user rotates the object to overlap the target.

we propose a new method for object manipulation based on head speed and inter-frame viewpoint quality to enhance the efficiency and accuracy of head manipulation. Finally, we design two user studies to evaluate the performance of our head-based object manipulation method. The results show that our method is feasible in terms of task completion efficiency and accuracy compared to state-of-the-art methods and greatly reduces user fatigue and motion sickness. Moreover, our method significantly improves usability and reduces task load. Our method lays a foundation for head-based object manipulation in virtual and augmented reality and provides a new manipulation method for scenarios where hands are not suitable for object manipulation (see Fig. 1). In summary, our main contributions are as follows:

- We construct the comfortable zone for head movement;
- For the first time, we propose an object manipulation method with adaptive gain based on head movement;
- We design a user study to evaluate the efficiency and accuracy of our method.

The rest of the paper is organized as follows: in Section 2, the previous work on object manipulation, hand-free object manipulation, and viewpoint quality is reviewed; Then, we describe the head manipulation space in Section 3. user study and the results are discussed in Section 5 and Section 6; conclusions, limitations, and future work are discussed in Section 7.

2. Related work

Object manipulation is a critical component of virtual reality (VR) applications as it enables users to interact directly with objects in virtual reality. Specific techniques and methods are required to ensure accurate and efficient selection and manipulation of virtual objects in VR. Numerous studies have explored object manipulation in VR in recent years, providing valuable theoretical foundations. To get a more in-depth understanding of object manipulation methods in virtual reality (VR), we suggest that readers refer to the survey paper by Mendes et al. (2019).

2.1. Object manipulation

When manipulating objects in virtual environments (VE), the most intuitive approach is to use the user's hands for direct manipulation (Frees et al., 2007; Song et al., 2012; Mendes et al., 2017; Gloumeau et al., 2020). However, this can lead to arm fatigue and may not be suitable for all manipulation scenarios. Various object manipulation methods have been proposed in the literature to address these challenges. Poupyrev et al. (1998) introduced the Go-Go method, which involves growing the user's arm and using nonlinear mapping to reach and manipulate distant objects. Bowman and Hodges (1999) proposed a ray casting-based object manipulation method, where users intersect rays with objects to grab and manipulate them. To enhance the accuracy and efficiency of object manipulation, Aguerreche et al. (2009) proposed the 3-hand manipulation method, which allows 2–3 users to collaboratively manipulate an object by manipulating three misaligned manipulation points. Nguyen and Duval (2013) introduced the 3-Point++ tool technology, which contains the center of gravity (6DOF) and three manipulate points (3DOF) for coarse and fine manipulations. They (Nguyen et al., 2014) also proposed the 7-Handle manipulation technique to generate a widget with multiple points for manipulating the object. Gloumeau et al. (2020) introduced the PinNPivot manipulation technology, which uses pins to constrain the rotation of the object, improving the efficiency and accuracy of manipulation. These multiple point-based manipulation methods improve the efficiency and accuracy of direct manipulation of virtual objects by hand at close range, but they may not be suitable for manipulating objects far from the user. Some researchers have used the speed of user control to determine the factors of manipulation. Frees et al. (2007) proposed the PRISM method, which divides the user's state into two modes depending on the speed of the hand. Pierce and Pausch (2002) scaled the manipulation of the HOMER method for higher precision long- and short-range manipulation tasks without slowing down. Osawa (2008) proposed two adjustment methods for position and viewpoint adjustments. Kim et al. (2015) proposed a non-linear mapping method to improve the manipulation efficiency, determining the combination coefficient through the speed and acceleration of hand translation and rotation. Liu et al. (2022) introduced the manipulation gains (translation gains, rotation gains, and scale gains), which allows the user to manipulate the object in VR with limited hand space. Manipulation gains are defined as some adjustment factors that map the handle of motion to the motion of the manipulated object. When manipulating objects with handles, the user's perspective is not affected. When using an HMD (Head Mounted Display) for object manipulation, the user's view will change as their head moves because the HMD display is fixed on the user's head. However, these methods require users to pay attention to the speed of their hand motion, leading to a high task load that reduces the efficiency and accuracy of manipulation. Therefore, most current virtual reality object manipulation methods in VR are performed through hand. This means that common interactions with VR systems can be hindered and restricted when the user's hands are used to directly manipulate objects.

2.2. Object manipulation based on eye and head

Currently, most of the interaction with VR applications is mainly achieved through controllers, which may limit and hinder the conventional interaction with VR systems in scenarios where users need to directly manipulate objects. For instance, a surgeon in VR should be able to place their hands on instruments and perform secondary tasks without causing any destructive events to the surgical task at hand. Therefore, some hand-free object manipulation interaction methods have been proposed recently. Liu et al. (2020a) introduced OrthoGaze, a novel interface that allows users to manipulate the 3D position

of virtual objects using only eye or head gaze. The method has the advantage of being hands-free and not requiring external hardware. However, the planes of OrthoGaze may occlude other objects in the environment, thus affecting the user experience. Liu et al. (2020b) designed and explored the usability of three different methods – RotBar, RotPlane, and RotBall – for 3D object rotation based on eye gaze. The experimental results demonstrated that RotBar and RotPlane were faster and more accurate in executing single-axis rotations, while RotBall outperformed the other two methods significantly in executing multi-axis rotations. Kim et al. (2017) proposed a hands-free natural user interface (NUI) for VR head-mounted displays (HMDs). A set of 7 kinds of commands required for 3D interaction with VR HMD is defined, among which the proposed seven control commands and their corresponding matching with facial gestures and head movements. Park et al. (2021) proposed a new approach to hands-free human–computer interaction. The proposed method provides coarse-to-fine interaction, which can support more effective and intuitive human–computer interaction. Coarse Interaction uses eye gaze to search and preview objects and UI. Fine-grained interactions can support the final selection of objects or UIs and precise 3D manipulation of objects using head gestures.

However, using eye gaze and head pose for human–computer interaction can be demanding in terms of time and effort. When manipulating multiple objects, the frequent need to adjust the head pose can cause physical strain and discomfort during the interaction. For instance, users may need to move their heads to focus on different objects or adjust their viewing angle, which can be tiring and time-consuming. This can be particularly challenging for individuals with limited mobility or neck-related issues. Moreover, prolonged use of eye gaze and head pose as the primary means of interaction may lead to eye strain and headaches. Therefore, it is important to consider the ergonomic implications of these interaction methods and explore alternative solutions that can mitigate the physical demands on users.

3. Head movement comfortable zone

In this section, we introduce the concept of head manipulation space (HMS) in Section 3.1. Then the construction of HMS is given in Section 3.2.

3.1. Definition

Head movements as a mode of interaction offer unique advantages in the field of human–computer interaction (HCI). First, since head movements are part of everyday human behavior, such as nodding or shaking the head, they provide an intuitive way of interaction. Users do not need to learn new gestures or movements because head movements directly map their interaction intent. This natural and intuitive feature makes head movements a necessary interaction in some settings. For users with motor impairments or limited hand function, head movements provide a viable interaction alternative that allows them to use technology without barriers. Additionally, head movements allow the user to interact with the device while performing other tasks, such as walking or holding objects, thus improving multitasking capabilities. Head movement can be used as a space-saving interaction in space-constrained environments, such as compact workspaces. Compared to other interaction methods that require more space, head movements can be performed in a limited space without additional equipment or space. In addition, prolonged use of gesture interaction may lead to muscle fatigue. Head movement, as a low-intensity interaction method, reduces the risk of muscle fatigue and enables users to interact with the device comfortably and efficiently for a longer period of time. The size and stability of the range of motion of the head are very important for daily life and physical performance. The range of motion of the head refers to the range of motion of the head in various directions. This range involves movements in various head directions, including flexion, extension, lateral flexion, and rotation to the left and right. In VR, HMS means that the user can manipulate objects through head movement in the head's range of motion, allowing users to interact with and manipulate objects more naturally.

The HMS of the head can be divided into two parts: the head movement space centered on the neck and the head movement space centered on the body's central axis. The HMS centered on the neck refers to the various postures and movements that the head can make based on the cervical spine, including pitch, roll, and yaw, as shown in Fig. 2(a). The space is mainly used for rotation and scale manipulation. The HMS centered on the body's central axis refers to the various postures and movements that the head can make around the body's central axis, including rotation, supine, prone, etc., as shown in Fig. 2(b). This part of the space is mainly used for translation manipulation. These two parts of space together enable the head to manipulate objects in VR.

3.2. Construction

In HMS, frequent head movements cause physical stress during the interaction (Park et al., 2021). It is very important for constructing a comfortable HMS. In this section, we design the experiment to sample the position of the head in HMS, which gets the comfortable HMS. We constructed the HMS of the user in sitting and standing postures, respectively.

Participants. We have recruited 12 participants, through social platforms (6 males and 6 females), between 20 and 30 years old. Each participant spent 55–60 min, which rewarded 100 yuan. Seven of our participants had used HMD VR applications before. Participants had normal and corrected vision, and none reported vision or balance disorders.

Hardware and software setup. We used an HTC Focus3 VR HMD system with two handheld controllers. The HMD was connected to its own workstation with a 3.6 GHz Intel(R) Core(TM) i7-9900KF CPU, 16 GB of RAM, and an NVIDIA GeForce GTX 3080 graphics card. To get the (Heart Rate Variability) (HRV) data, we use the polar H10. The tracked physical space hosting the VR applications is 4 m × 4 m. We used Unity 2019.1 to implement our VR manipulation tasks. The virtual environments were rendered at 90fps for each eye.

Manipulation implementation. We map the head motion to the motion of the manipulated virtual object to manipulate. When the user keeps pressing and holding the "Translate" button on the handheld controller, the head's pan is mapped 1:1 into the virtual space. The rotation of the head is mapped 1:1 into virtual space when the user keeps the "Rotation" button on the handheld controller pressed. When the user keeps pressing and holding the "Scale" button on the handheld controller, the yaw angle of the head is scaled 1:1 to the object.

Tasks. The user was asked to complete two tasks in both standing and sitting postures in a bunny scene. Participants took part in 4 sets of experiments ($2poses \times 2HMS$), and they did not see the visualization of HMS. There was a break of at least 5 min between each set of experiments. In the first task, users were asked to manipulate an object using maximum head movements in both standing and sitting postures for a duration of 10 min. In the second task, the user was asked to manipulate an object using comfortable head movements in both standing and sitting postures, also for a duration of 10 min. In both tasks, we sequentially fixed the target location of the bunny, and the user's task was to manipulate the object to reach the target location. We informed the user that our purpose was to sample the head position.

Procedure. The first step for participants is to read and sign a consent form. The researcher will then first explain how to place the Polar H10. The researcher will then ask the participant to move to a room to wear the Polar H10. For privacy, the participant will move to an empty private room. Once equipped with the Polar H10, participants returned to the lab, where the researcher checked to see if the Polar H10 was being worn correctly. After verification, participants were

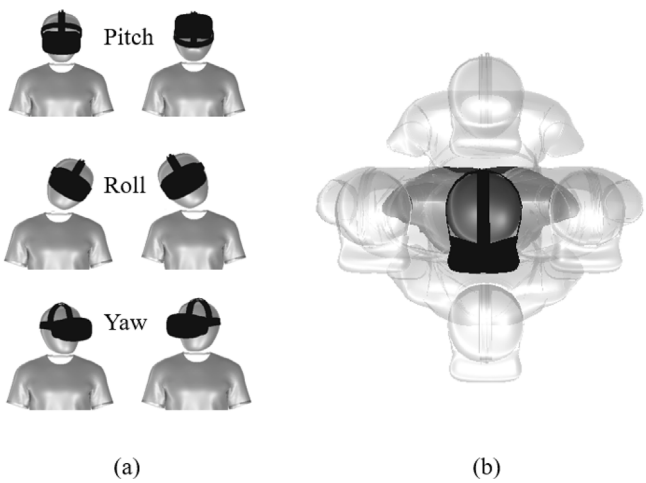

Fig. 2. Head manipulation space. (a) The HMS is centered on the neck. (b) The HMS is centered on the body's central axis.

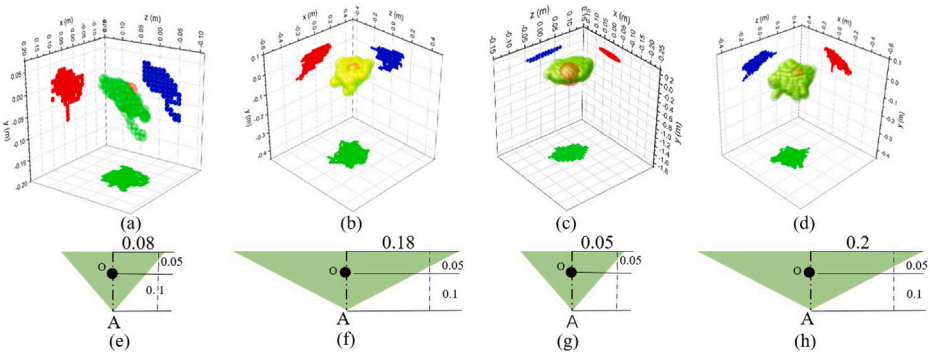

Fig. 3. Head manipulation space. (a) is the head point cloud visualization centered on the when sitting, and (e) is the HMS obtained from (a). (b) is the head point cloud visualization centered on the body's central axis when sitting, and (f) is the HMS obtained from (b). (c) is the head point cloud visualization centered on the when standing, and (g) is the HMS obtained from (c). (d) is the head point cloud visualization centered on the body's central axis when standing, and (h) is the HMS obtained from (d).

asked to complete the Checklist Individual Strength (CIS) questionnaire (Vercoulen et al., 1994). Before starting the formal experiment, the user should practice for 1 min to get familiar with the experiment. After the user starts to enter the bunny scene, he/she first needs to start formally through the "Start" button on the handle, and we recorded the position of the head in the first frame after the start of the "Start" button, which is set as the reference points.

Results. The head position data of the 12 users are first processed according to the relative positioning method, where the first frame of the head position data recorded by the user through the "start" button is used as the reference point. We removed some data with large differences using the percentile method, removing a total of 23. The processed head position data is then visualized in the 3D coordinate system, as shown in the first row of Fig. 3, which only shows the comfortable HMS. We use a cone as the basic space of the HMS. The basic parameters of a cone include base radius (r) and height (h). We each project the processed head position data in (a)(b)(c)(d) of Fig. 3 onto three coordinate planes. With the three projection planes we obtain the base radius (r) and height (h) of each HMS, ensuring that the positions of the collected heads are in that HMS, shown as the second row of Fig. 3. The surface equations corresponding to the cones in (e) (f) (g) (h) are as follows:

$$f(x,z) = \sqrt{x^2+z^2}cot30.5° - 0.1 \tag{1}$$

$$f(x,z) = \sqrt{x^2+z^2}cot68.75° - 0.1 \tag{2}$$

$$f(x,z) = \sqrt{x^2+z^2}cot19.08° - 0.1 \tag{3}$$

$$f(x,z) = \sqrt{x^2+z^2}cot71.6° - 0.1 \tag{4}$$

We visualize these four HMS in the user's field of view, as shown in Fig. 4. The small red ball is a visualization of the head position in HMS. Users can know whether their head is in HMS through this visualization. We performed an analysis of HRV parameters extracted from the same period to determine the physiological stress of different HMS. The analysis compared the values obtained in each parameter with the mean values of the same parameters shown in the healthy population. Means for the healthy population are from Voss et al. (2015), and trends for each parameter under stress are from the (Castaldo et al., 2015). Narciso et al. (2022) listed the HRV parameters used, a brief description of each parameter, their stress trends, and their mean values in a healthy population. For HRV, we evaluated the experimental results with non-parametric statistical tests. This is due to the small study sample size, as HRV responses vary across individuals, making the data more prone to outliers. Friedman tests were run to determine if there were differences in HRV parameters between $HMSs$. The descriptive statistics in Table 1 show that the AVNN (Average interval between normal heart beats in milliseconds) and pNN50 (Percent of successive differences of intervals between normal heart beats > 50 ms) values drop from the baseline environment to Comfortable HMS (cmy-HMS), and they drop again from cmy-HMS to max-HMS. LF/HF shows the opposite pattern, rising from the baseline environment to cmy-HMS and then from cmy-HMS to max-HMS. This indicates an increase in physiological stress from baseline to cmy-HMS and from cmy-HMS to max-HMS. These values point to the same conclusion

Fig. 4. Visualization of HMS. The small red ball is a visualization of the position of the head in the HMS.

Table 1
Median values of HRV parameters and corresponding Friedman tests results (N = 12).

Human posture	HRV parameter	Baseline (Mdn)	Max HSM	cmy HSM	$\chi^2(2)$	p
$Sitting^1$	$AVNN$	781.81	778.33	780.58	3.5	0.016
	$SDNN$	45.05	48.98	45.9	6.4	0.09
	$RMSSD$	28.60	26.47	28.31	7.5	0.02
	$pNN50$	8.26	5.38	5.38	4.26	0.12
	LF/HF	2.3	2.39	2.12	8	0.022
$Standing^1$	$AVNN$	763.02	739.81	757.97	5.9	0.009
	$SDNN$	48.17	52.03	41.89	2.4	0.23
	$RMSSD$	26.11	22.99	25.64	6.9	0.0052
	$pNN50$	2.58	2.07	2.47	6.26	0.03
	LF/HF	2.32	2.67	2.39	9.3	0.0078
$Sitting^2$	$AVNN$	781.81	669.51	760.54	7.1	0.02
	$SDNN$	45.05	37.07	46.72	3.4	0.10
	$RMSSD$	28.60	20.49	24.42	3.5	0.042
	$pNN50$	8.26	2.68	2.72	2.26	0.132
	LF/HF	2.3	3.45	2.61	9	0.034
$Standing^2$	$AVNN$	763.02	662.08	717.97	10	0.0001
	$SDNN$	48.17	32.51	35.96	4.4	0.0876
	$RMSSD$	26.11	19.57	21.78	7.9	0.001
	$pNN50$	2.58	1.63	1.67	5.86	0.038
	LF/HF	2.32	5.77	5.67	1.83	0.18

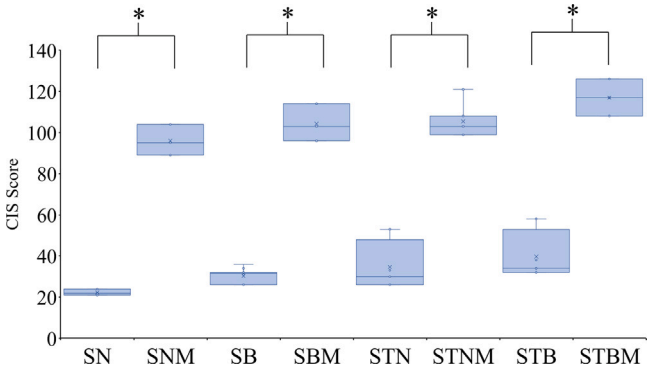

Fig. 5. The CIS Score.

that participants were not under physiological stress in the baseline environment, were under higher physiological stress in the max-HMS, and were closer to baseline in the cmy-HMS. And in cmy-HMS, according to the order of physiological pressure from small to large, it is $sitting^1 < sitting^2 < Standing^1 < Standing^2$.

We used a Wilcoxon signed rank to determine whether there was a significance between cmy-HMS and max-HMS of fatigue. Fig. 5 shows the results of the CIS. The fatigue of SN (cmy-HMS centered around the neck in the sitting position of the user) was significantly lower than that of SNM (max-HMS centered around the neck in the sitting position of the user) (z=−3.095, p=0.002). The fatigue of SB (cmy-HMS centered around the body in the sitting position of the user) was significantly lower than that of SBM(max-HMS centered around the body in the sitting position of the user) (z=−3.0389, p=0.00237). The fatigue of STN (cmy-HMS centered around the neck in the standing position of the user) was significantly lower than that of STNM (max-HMS centered around the neck in the standing position of the user) (z=−3.034, p=0.00210). The fatigue of STB (cmy-HMS centered around the body in the standing position of the user) was significantly lower than that of STBM (max-HMS centered around the body in the standing position of the user) (z=−3.0484, p=0.00230). In addition, the cmy-HMS of sitting is less fatigued than the cmy-HMS of standing. We asked participants the question, "Is it suitable for object manipulation in standing?" after the experiment. Only one participant felt it could be considered, but he felt that manipulating in a standing state was more fatiguing than sitting. Participants' stress responses, both subjective and objective, were consistent. And the pressure of the cmy-HMS we constructed is obviously smaller than that of max-HMS, and the cmy-HMS under sitting is lower than that under standing. The four cmy-HMS we constructed will be collectively called HMS in the following sections.

4. Method

In this section, we first gave the Maximum gains. Based on the Maximum gains, two strategies for adaptive computation gain: (1) head velocity based gains computation (Section 4.3.1); (2) view-quality based gains computation (Section 4.3.2). At last, We divide the adaptive gain calculation method into two phases by detecting the velocity of the head (Section 4.3.3).

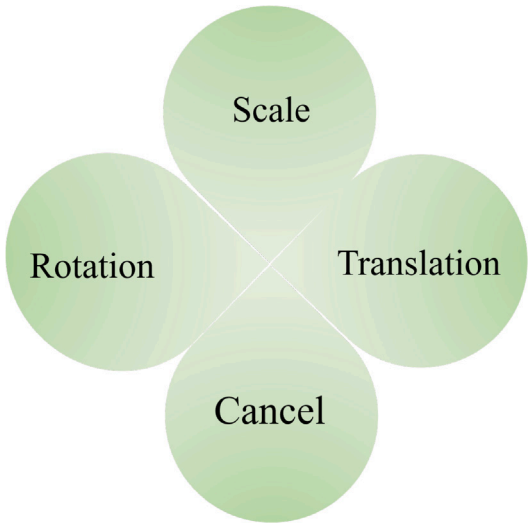

Fig. 6. The UI with four options: Rotation, Translation, Scale and Cancel.

4.1. Manipulation design based on head

In this section, we design methods for the translation, rotation, and scaling of objects in HMS. For translation, we manipulate the object's front, rear, left, and right translation through the HMS centered on the body's central axis; we translate up and down through the vertical position difference of the head pitch movement centered on the neck. We perform a rotation manipulation with a head rotation movement centered on the neck. For scale, when the user's head is upward, the object is zoomed in, and vice versa. We designed the UI (Fig. 6) with four options. And only when the user selects a type, the UI will be generated in the user's view. Users can select the corresponding option by looking at the corresponding direction. We define a selection confirmation time (set to 1 s).

We cancel the user to perform the corresponding type of manipulation by blinking twice. After the user completes the corresponding manipulation type, the manipulation type is performed by canceling the movement of the user's head by blinking twice.

4.2. Determine maximum gains

In Section 3.2, the HMS we obtained is relatively small, but the range of objects manipulated in the virtual reality environment can be very large. Using the HMS to manipulate objects can lead to low manipulation efficiency, accuracy, and a large task load. We need to determine the maximum matching for object manipulation within a limited range of HMS to ensure object manipulation. This section mainly determines the maximum matching for object manipulation within a limited HMS.

Manipulation Gains based on Head Movement. Manipulation gains (Liu et al., 2022) allow the user to manipulate objects in VR within a limited hand space. Based on the hand object manipulation gain, we first introduce the head manipulation gain. This paper defines the manipulation gain as some adjustment factor that maps the HMD motion to the manipulated object motion. When manipulating an object using an HMD, the user's viewpoint changes with the movement of the head.

Viewpoint Space. Viewpoint Space is the space the user can see in VR. We define Viewpoint Space based on two factors: field of view (FOV) and interaction distance (d), as shown in Fig. 7(a). We specify d to be 10, and the FOV is 120°.

The HMS (Fig. 2(f)) is used to translate objects, and the max head distance is 0.18. The HMS (Fig. 2(g)) is used to rotate and scale objects, and the max angle is 30.5°. In Fig. 7(b), we calculate the distance of OB is 17.32. In addition, when the user is manipulating the object, the user's viewpoint space is moving, and the left and right movement distance is the largest, so we only need to calculate the distance OB' of moving the head left and right to calculate the translation gain. BO' includes the OB and the head distance; OB' is 17.5. The translation gain is 97.2 (17.5/0.18), and the rotation gain is 360/30.5=11.8. For Scale, the scale gain is 15.

4.3. Compute gains adaptively

For Max gains, users cannot perform fine manipulation. So we propose the following two strategies for adaptive computation gain: (1)head velocity based gains computation (Section 4.3.1); (2) view-quality based gains computation (Section 4.3.2)

4.3.1. Head velocity based gains computation

The commonly used method to manipulate virtual objects is by tracking the user's hand motion. Among these hand-based methods (Frees et al., 2007; Kim et al., 2015; Wilkes and Bowman, 2008), one type relies on calculating a control/display ratio based on the speed of the user's hand motion and uses this ratio to adjust the user's hand motion to manipulate objects in the virtual environment. We first use the speed of the user's head motion to calculate manipulation gains.

When the user performs head manipulation in the HMS, there are maximum and minimum speeds. In this section, we collect the moving speed and rotation speed of the object manipulated by the participant's head. From these speeds, we obtain the maximum, minimum, and head speeds when the gain is 1. We developed a data collection application using Unity with Focus 3 to collect the head tracking data from the Focus 3 headset.

Participants. We have recruited 12 participants through social platforms (6 males and 6 females), between 20 and 32 years old. Five of our participants had used HMD VR applications before. Each participant spent 45–60 min, which rewarded 100 yuan. Participants had a normal and corrected vision, and none reported vision or balance disorders.

Tasks. Participants were asked to use head movements to manipulate the bunny to the target position, where the gains were 1:1. Participants took part in an experiment, and they can see the visualization of HMS. In the fine manipulation phase, participants were asked to adjust the gains and record the fit minimum gains through the "record" button on the controller. When the three precision conditions are all met, the task ends automatically.

Procedure. The first step for participants is to read and sign a consent form. Before starting the formal experiment, the user should practice for 1 min to get familiar with the experiment. After the user starts to enter the bunny scene, he/she first needs to start formally through the "Start" button on the handle, and we record the tracking data of the head in every frame. When participants translate and scale objects, we record the head movement speed. When we participants rotate objects, we record the head rotation speed. We store the three speeds in three different files.

Results. We get the maximum and minimum values in three files from the participants. To analyze the collected data, we first discard the outliers that deviated more than three standard deviations from the mean value (mean ±3std.). We average the maximum and minimum values collected from the participants as the maximum and minimum head movement and rotation speeds. The maximum head movement speed when translating is 118.32 (cm/s). The minimum head movement speed when translating is 0.29 (cm/s). When the translation gain is 1, the average of head movement speed is 0.345 (cm/s). And the average minimum translation gain is 0.69. The maximum head rotation speed when rotating is 5.36 (rad/s). The minimum head movement speed when rotating is 0.02 (rad/s). When the rotation gain is 1, the average of head rotation speed is 1.2 (rad/s). And the average minimum

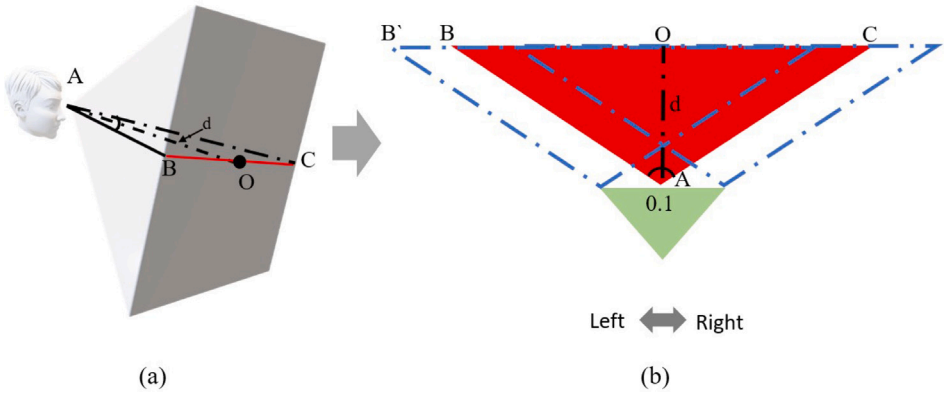

Fig. 7. Viewing space.

rotation gain is 0.45. The maximum head rotation speed when scaling is 5.36 (rad/s). The minimum head movement speed when scaling is 0.02 (rad/s). When the scale gain is 1, the average of head rotation speed is 0.9 (rad/s). And the average minimum scale gain is 0.0053.

In Section 4.2, we get the max translation gain, rotation gain, and scale gain. When the head movement speed is high, the gain is large; when the head movement speed is small, the gain is small. So for the translation gain function, we assume that the function has three points (0.16, 0.089), (0.345, 1), and (118.32, 97.2), and we get the translation gain function Eq. (5). For the rotation gain function, we assume that the function has (0.02, 0.45) (1.2, 1), and (5.36, 17.5) three points, and we get the translation gain function Eq. (6). For the scale gain function, we assume that the function has (0.312, 0.175) (1.43, 1), and (4.18, 15) three points, and we get the translation gain function Eq. (7).

$$\begin{cases} g_t = 16.479 \ln v_t + 18.538 (v_t \in [0.345, 118.32]) \\ g_t = 28.518 v_t^{3.1484} (v_t \in [0.16, 0.345]) \end{cases} \tag{5}$$

$$\begin{cases} g_r = 10.904 \ln v_r - 0.9881 (v_r \in [1.2, 5.36]) \\ g_r = 0.723 v_r^{1.7792} (v_r \in [0.21, 1.2]) \end{cases} \tag{6}$$

$$\begin{cases} g_s = 13.005 \ln v_s - 3.6683 (v_s \in [1.43, 4.18]) \\ g_s = 0.664 v_s^{1.1449} (v_s \in [0.312, 1.43]) \end{cases} \tag{7}$$

4.3.2. View-quality based gains computation

For head velocity based gains computation, the efficiency and accuracy of manipulation are affected due to the imprecise perception of the user's head motion speed. We propose a gains computation method for inter-frame view quality.

View quality. The view quality is the ratio of the area object to image size. Inter-frame view quality is pixel difference variance, calculated by Algorithm 1.

In this algorithm, we first calculate the area of the manipulated object in the NUM frame buffer and store it in the Lp list (lines 1–7). We set NUM as 5. We use V to render image I_o with a single object o (line 2). Then we calculate the area of the object A_o in I_o (line 3) and the area of I_o(line 4). We add the A_o to the list Lp (line 5). the k factor is computed by the line 8. At last, we get the Inter-frame view quality VQ_{if} (lines 9–10).

Participants. We have recruited 12 participants through social platforms (6 males and 6 females), between 20 and 32 years old. Five of our participants had used HMD VR applications before. Each participant spent 50–60 min, which rewarded 100 yuan. Participants had normal and corrected vision, and none reported vision or balance disorders.

Algorithm 1 Inter-frame view quality

Input: object o, viewpoint V
Output: inter-frame view quality VQ_{if}
1: **While** Frame F< NUM
2: I_o = RenderObject(v, o);
3: A_o=Area(I_o, O);
4: A=Area(I_o);
5: Lp=List.add(A_o);
6: F++;
7: **EndWhile**
8: k=A_o/A;
9: VQ_{if}=getDifferenceVariance(Lp);
10: VQ_{if}=VQ_{if}/k;
11: **return** VQ_{if};

Tasks. Participants were asked to use head movements to manipulate the bunny to the target position, where the gains were 1:1. Participants took part in an experiment, and they can see the visualization of HMS. In the fine manipulation phase, participants were asked to adjust the gains and record the fit minimum gains through the "record" button on the controller. When the three precision conditions are all met, the task ends automatically.

Procedure. The first step for participants is to read and sign a consent form. Before starting the formal experiment, the user should practice for 1 min to get familiar with the experiment. After the user starts to enter the bunny scene, he/she first needs to start formally through the "Start" button on the handle, and we record the VQ_{if} in every frame. When participants translate and scale objects, we record the head movement speed. We recorded VQ_{if} in three different files as participants translated, rotated, and scaled objects.

Results. We get the maximum and minimum VQ_{if} in three files from the participants. To analyze the collected data, we first discarded the outliers that deviated more than three standard deviations from the mean value (mean ±3std.). We average the maximum and minimum values collected from the participants as the maximum and minimum VQ_{if}. The maximum head VQ_{if} when translating is 118.32 (cm/s). When translating, the minimum VQ_{if} is 0.29 (cm/s). When the translation gain is 1, the average of VQ_{if} is 0.345 (cm/s). And the average minimum translation gain is 0.69. When rotating, the maximum VQ_{if} is 5.36 (rad/s). When rotating, the minimum VQ_{if} is 0.02 (rad/s). When the rotation gain is 1, the average of head VQ_{if} is 1.2 (rad/s). And the average minimum rotation gain is 0.45. When scaling, the maximum VQ_{if} is 5.36 (rad/s). When scaling, the minimum VQ_{if} is 0.02 (rad/s).

When the scale gain is 1, the average of VQ_{if} is 0.9 (rad/s). And the average minimum scale gain is 0.0053.

In Section 4.2, we get the max translation gain, rotation gain, and scale gain. When the VQ_{if} is high, the gain is large; when the VQ_{if} is small, the gain is small. So for the translation gain function, we assume that the function has three points (0.06, 0.069), (0.163, 1), and (123.3, 97.2), and we get the translation gain function Eq. (8). For the rotation gain function, we assume that the function has (0.04, 0.45) (0.163, 1), and (66.6, 17.5) three points, and we get the translation gain function Eq. (9). For the scale gain function, we assume that the function has (0.02, 0.175) (1.43, 1), and (36.34, 15) three points, and we get the translation gain function Eq. (10).

$$\begin{cases} g_t = 14.513 \ln q_t + 27.326 (q_t \in [0.163, 123.3]) \\ g_t = 128.11 q_t^{2.6752} (q_t \in [0.06, 0.163]) \end{cases} \tag{8}$$

$$\begin{cases} g_r = 3.2766 \ln q_r + 3.7426 (q_r \in [0.433, 5.36]) \\ g_r = 1.3239 q_r^{0.3352} (q_r \in [0.04, 0.433]) \end{cases} \tag{9}$$

$$\begin{cases} g_s = 4.3273 \ln q_s - 0.5478 (q_s \in [1.43, 36.34]) \\ g_s = 0.8641 q_s^{0.4082} (q_s \in [0.02, 1.43]) \end{cases} \tag{10}$$

4.3.3. Two phrases gains computation

For the gain computation based on head velocity (Section 4.3.1), the efficiency and accuracy of the manipulation will be affected due to the inaccuracy of the user's perception of the head movement velocity during the fine manipulation stage. For the view-quality based gains computation, it is necessary to ensure that the area of the Object in the user's view is not 0, which makes the method unnatural to use. This method is more suitable for use in the stage of fine manipulation. Therefore, we divide the two methods into Phase 1 and Phase 2 for use.

Phase 1. For Phase 1, we used the gain computation based on head velocity (Section 4.3.1). For the gain computation based on head velocity, the efficiency and accuracy of the manipulation will be affected due to the inaccuracy of the user's perception of the head movement velocity during the fine manipulation stage.

Phase 2. For phase 2, we used the gain computation based on view-quality. For the view-quality based gains computation, it is necessary to ensure that the area of the Object in the user's view is not 0, which makes the method unnatural to use. This method is more suitable for use in the stage of fine manipulation.

Detect. For Phase 1 it is mainly coarsely manipulated and for Phase 2 it is finely manipulated. The two-phase segmentation is mainly performed by detecting the velocity of the head movement. For the translation gain function Eq. (11), when the head velocity is detected to be less than 0.8 (cm/s). For the rotation gain function Eq. (12), when the head rotation velocity is detected to be less than 1.35 (rad/s). For the scale gain Eq. (13) function, when the head rotation velocity is detected to be less than 1.64 (rad/s).

$$\begin{cases} g_t = 16.479 \ln v_t + 18.538 (v_t \in [0.8, 118.32]) \\ g_t = 128.11 q_t^{2.6752} (q_t \in [0.06, 0.163]) \end{cases} \tag{11}$$

$$\begin{cases} g_r = 10.904 \ln v_r - 0.9881 (v_r \in [1.35, 5.36]) \\ g_r = 1.3239 q_r^{0.3352} (q_r \in [0.04, 0.433]) \end{cases} \tag{12}$$

$$\begin{cases} g_s = 13.005 \ln v_s - 3.6683 (v_s \in [1.64, 4.18]) \\ g_s = 0.8641 q_s^{0.4082} (q_s \in [0.02, 1.43]) \end{cases} \tag{13}$$

5. User study 1

In this study, our goal is to evaluate and compare manipulation techniques for head manipulation space (HMS). We aim to gain a better understanding of how object manipulation methods based on HMS can facilitate more convenient and efficient 3D object manipulation in VR. The study primarily focuses on the main workspace, where all the objects of interest are located in front of the user. Most of the work in VR is likely to occur in this area, so users do not need to return or move around in the virtual environment frequently (see Fig. 8).

5.1. Participants, hardware and software setup

We recruited 24 participants, consisting of 16 males and 8 females, aged between 20 and 32 years. Seventeen of the participants had prior experience with VR. Participants had normal vision and corrected visual acuity, and none reported any visual or balance disorders. There were four control conditions and one experimental condition. Control condition CC_1 utilized the traditional method of object manipulation using handheld controllers, with a manipulation gain of 1 (Liu et al., 2022). The purpose of CC_1, which is a method of object manipulation using the hand, compared to the method of object manipulation by hand, is to verify where the advantages of our method lie compared to the method of object manipulation by hand. Control condition CC_2 employed head motion for object manipulation, with a manipulation gain of 1. Experiment condition EC_1 involved a manipulation method based on head velocity, with adaptive manipulation gain calculated based on head velocity. Experiment condition EC_2 featured a manipulation method with adaptive manipulation gain based on inter-frame viewpoint quality. Experiment condition EC_3 had adaptive manipulation gain calculated using a two-stage gain.

We used the HTC Focus 3 VR HMD system, which consists of a headset with two handheld controllers. The HMD was connected to its own workstation, equipped with a 3.6 GHz Intel(R) Core(TM) i7-9900KF CPU, 16 GB RAM, and an NVIDIA GeForce GTX 3080 graphics card. For capturing HRV data, we utilized the Polar H10. The physical tracking space hosting the VR application was 4m×4 m. We implemented our VR manipulation tasks using Unity 2019.1. The virtual environment was rendered at a rate of 90fps for each eye. We employed the HTC Focus 3's Eye Tracker for eye tracking, a dual-camera setup with infrared illumination. It can track your gaze direction and origin while measuring your pupil size and blink rate.

Hypotheses. Our method aims to enable users to manipulate objects effectively toward their targets. Therefore, we propose the following hypotheses:

H1: Users can manipulate objects toward their targets more quickly using EC_{1-3} compared to CC_2.

H2: The user task load is lower for EC_{1-3}, compared to CC_2.

5.2. Task

Participants were instructed to manipulate objects to their designated target positions as quickly and accurately as possible during the task. There was one scene in the task, and the targets within the scene were fixed. The size of the manipulated objects was randomly generated, ranging from approximately 0.4 to 2 times the size of the target. The initial positions of the objects and the users were randomly placed within the scene. The task automatically ended when the movement error of the manipulated object was less than 15 mm, the rotation error was less than 3.5°, and the scaling error was less than 0.01. We recorded

Fig. 8. Bunny scene.

the completion time of the task. In the Bunny scene, participants were required to manipulate a rabbit model to the target position (Fig. 8). The Bunny scene had dimensions of 20 m × 20 m, and the target size was 0.9 m × 0.8 m × 0.8 m.

5.3. Procedure

In the experiment, each condition was repeated five times, resulting in a total of 600 trials ($24 participants \times 5 techniques \times 5 repetitions$). The entire experiment lasted approximately 60 minutes. Participants first completed a questionnaire to collect their demographic information. They were then introduced to the experimental tasks and VR equipment and instructed to perform the trials as quickly and accurately as possible. Next, participants were asked to put on the VR headset and engage in a 5-minute VR experience to familiarize themselves with the environment. The VR experiment consisted of five sessions corresponding to the five manipulation techniques. Each session started with ten warm-up trials to acquaint participants with the method, followed by formal testing trials. After each session, we administered the $NASA{-}TLX$ questionnaire (Hart, 2006; Hart and Staveland, 1988) to gather subjective feedback.

5.4. Evaluation metrics

The task performance is measured using objective metrics: (1) Task completion time, measured in seconds, represents the time taken from participants pointing to the object until the manipulation task is completed. (2) Movement time, measured in seconds, represents the time spent on manipulating the task by moving from pointing to the object until completion. (3) Rotation time, measured in seconds, represents the time spent on manipulating the task by rotating from pointing to the object until completion. (4) Scaling time, measured in seconds, represents the time spent on manipulating the task by scaling from pointing to the object until completion. (5) HRV feature values (Narciso et al., 2022) . We assess the user's task workload using the standardized NASA TLX questionnaire and the user's simulator sickness using the standardized SSQ questionnaire (Kennedy et al., 1993).

5.5. Statistics

To analyze the collected data, we first discarded outliers that deviated from the mean by more than three standard deviations (mean ±3 std.) in each condition (20 trials, 1.3%). Additionally, the Shapiro–Wilk test indicated that the data did not follow a normal distribution. Therefore, all data were preprocessed using the aligned rank transform (ART) (Wobbrock et al., 2011). Next, we performed repeated measures analysis of variance (RM-ANOVA) (Gelman, 2005) and paired comparisons with Bonferroni adjustment for each measurement. The sphericity assumption was assessed using Mauchly's test (Mauchly, 1940). In case of violation of the assumption, the Greenhouse–Geisser correction was applied. Subsequently, an overall analysis of variance was conducted to investigate if the null hypothesis, indicating no statistically significant differences between conditions, could be rejected. Post hoc tests were performed with Bonferroni correction to examine differences between the five pairs when the null hypothesis was rejected. The effect sizes were quantified using Cohen's d (Cohen, 2013).

Table 2
The completion time, in s.

Condition	Avg± std. dev.	$(CC_i\text{-}EC) / CC_i$	p	Cohen's d	Effect size
EC_3	53.30 ± 12.10				
CC_1	81.27 ± 23.90	34.4%	< 0.001*	1.48	Very Large
CC_2	79.83 ± 24.91	33.2%	< 0.001*	1.35	Very Large
EC_1	74.90 ± 23.30	28.8%	< 0.001*	1.16	Large
EC_2	68.67 ± 10.96	18.3%	< 0.001*	1.33	Very Large

Table 3
The translation time, in s.

Condition	Avg± std. dev.	$(CC_i\text{-}EC) / CC_i$	p	Cohen's d	Effect size
EC_3	53.30 ± 12.10				
CC_1	81.27 ± 23.90	34.4%	< 0.001*	1.48	Very Large
CC_2	79.83 ± 24.91	33.2%	< 0.001*	1.35	Very Large
EC_1	74.90 ± 23.30	28.8%	< 0.001*	1.16	Large
EC_2	68.67 ± 10.96	18.3%	< 0.001*	1.33	Very Large

5.6. Results

Table 2 gives the task completion time. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.329,19.935} = 132.15, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly shorter than for CC_1, CC_2, EC_1 and EC_2. EC_3 method significantly improves the task time performance, and the effect size is "Huge".

Table 3 gives translation time. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.58,23.705} = 108.929, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly shorter than for CC_1, CC_2, EC_1 and EC_2. EC_3 method significantly improves the task time performance, and the effect size ranges from "Very Large" to "Huge".

Table 4 gives the rotation time. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.817,27.257} =$

Table 4
Rotation time, in s.

Condition	Avg± std. dev.	$(CC_i\text{-}EC)$ / CC_i	p	Cohen's d	Effect size
EC_3	53.30 ± 12.10				
CC_1	81.27 ± 23.90	34.4%	$< 0.001^*$	1.48	Very Large
CC_2	79.83 ± 24.91	33.2%	$< 0.001^*$	1.35	Very Large
EC_1	74.90 ± 23.30	28.8%	$< 0.001^*$	1.16	Large
EC_2	68.67 ± 10.96	18.3%	$< 0.001^*$	1.33	Very Large

Table 5
The scale time, in s.

Condition	Avg ± std. dev.	$(CC_i\text{-}EC)$ / CC_i	p	Cohen's d	Effect size
EC_3	53.30 ± 12.10				
CC_1	81.27 ± 23.90	34.4%	$< 0.001^*$	1.48	Very Large
CC_2	79.83 ± 24.91	33.2%	$< 0.001^*$	1.35	Very Large
EC_1	74.90 ± 23.30	28.8%	$< 0.001^*$	1.16	Large
EC_2	68.67 ± 10.96	18.3%	$< 0.001^*$	1.33	Very Large

8.122, $P = 0.002$). Post-hoc analysis reveals that EC_3 was significantly shorter than for CC_1 and CC_2. EC_3 method significantly improves the task time performance than CC_1 and CC_2, and the effect size ranges from "Large" to "Very Large". EC_3 does not offer significant improvements over EC_1 and EC_2 in terms of rotational task performance. Table 5 gives the scale time. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.623,24.352} = 142.562$, $P <$ 0.001). Post-hoc analysis reveals that EC_3 was significantly shorter than for CC_1, CC_2, EC_1 and EC_2. EC_3 method significantly improves the task time performance than CC_1 and CC_2, and the effect size is "Huge". EC_3 does not offer significant improvements over EC_1 and EC_2 in terms of scale task performance.

Table 6 gives the HRV. Statistical significance is indicated by an asterisk. For $AVNN$, the sphericity assumption is verified: $p = 0.053$. The overall ANOVA reveals significant differences between the five conditions: ($F_{4,24} = 78.873, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2 and EC_1. For $SDNN$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.635,14.711} = 7.433$, $P = 0.008$). Post-hoc analysis reveals that EC_3 was significantly smaller than for CC_2 and EC_1. For $RMSSD$, the sphericity assumption is violated: $p = 0.049$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.028,18.254} = 16.982, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2, and EC_1. For $pNN50$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.098,18.833} = 18.704, P <$ 0.001). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2, and EC_1. For LF/HF (A ratio of Low Frequency to High Frequency), the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.397,12.571} =$ 22.595, $P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2, EC_1 and EC_2. The EC_3 method significantly reduces psychological stress compared to CC_1, CC_2, and EC_1.

Fig. 9 gives the task load. Statistical significance is indicated by an asterisk. The positive fraction performance is replaced by its complement, so smaller values are more favorable. We tested six aspects of the spherical assumption in NASA-TLX and applied Greenhouse–Geisser correction when necessary. For $Mental$, the sphericity assumption is violated: $p = 0.008$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{4,36} = 78.873, P < 0.001$). For $Physical$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.102,18.914} = 66.954, P < 0.001$). For $Temporal$, the sphericity assumption is violated: $p = 0.0021$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.106,18.954} =$ 120.953, $P < 0.001$). For $Performance$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.339,21.055} = 60.650$, $P < 0.001$). For $Effort$, the sphericity assumption is violated: $p = 0.003$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.068,18.610} = 54.851, P < 0.001$). For $Frustration$, the sphericity assumption is violated: $p = 0.002$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.809,16.283} =$ 45.407, $P < 0.001$). For $Overall$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.482,22.342} = 184.826, P < 0.001$). Compared to CC_2, EC_1, EC_2, and EC_3 show significant improvements in all six aspects. Compared to CC_1, EC_3 significantly improves four aspects, except for Metal and Physical. Compared to EC_1 and EC_2, EC_3 significantly improves effort and frustration. Overall, EC_3 significantly improves total score compared to CC_1, CC_2, EC_1, and EC_2.

We used the standard simulator sickness questionnaire (SSQ) (Kennedy et al., 1993) (Table 7) to measure simulator sickness. The SSQ was managed before and after the experiment for each task and each condition. The SSQ scores are not normally distributed, and the differences before and after the Wilcoxon signed-rank test were used. These differences are not statistically significant with CC_1, CC_2, EC_1, EC_2, and EC_3. but the difference is statistically significant with CC_2.

6. User study 2

This user study aims to assess and compare the operational techniques of Head-Mounted Space (HMS) to understand how to perform 3D object manipulation more conveniently and efficiently in virtual reality. In this study, users are not required to return or physically move within the virtual environment frequently. The hardware equipment remains consistent with User Study 1.

6.1. Participants

We recruited 24 participants for the study, consisting of 18 males and 6 females, with ages ranging from 20 to 32 years. Among the participants, 15 had previous experience with virtual reality (VR). All participants had normal vision and corrected vision, and no one reported any visual or balance disorders.

Hypotheses. Our method aims to enable users to manipulate objects effectively toward their targets. Therefore, we propose the following hypotheses:

H1: Users can manipulate objects more accurately toward their targets using EC_{1-3} compared to CC_{1-2}.

H2: Users can manipulate objects toward their targets more quickly using EC_{1-3} compared to CC_2.

H3: The user task load is lower for EC_{1-3}, compared to CC_2.

6.2. Task

During the task, participants were instructed to manipulate four objects to their designated target positions as quickly and accurately as possible. The task involved one scene with four fixed targets. The size of the manipulated objects was randomly generated, approximately ranging from 0.4 to 2 times the size of the targets. The positions of the

Table 6
Median values of HRV parameters and corresponding Friedman tests results of user study 1.

HRV parameter	Baseline (Mdn)	CC_1	CC_2	EC_1	EC_2	EC_3	p_1	p_2	p_3	p_4
$AVNN$	855.3	743.2	783.5	813.5	837.3	841.2	$< 0.01^*$	$< 0.01^*$	0.02^*	0.99
$SDNN$	50.2	58.2	61.8	57.0	54.2	52.2	0.466	$< 0.01^*$	$< 0.01^*$	0.09
$RMSSD$	47.5	29.49	30.81	31.47	38.9	40.5	0.017	0.037^*	0.01^*	0.98
$pNN50$	9.32	5.62	6.87	7.486	7.95	8.85	0.01^*	0.001^*	0.177	0.97
LF/HF	2.45	2.94	2.64	2.581	2.655	2.414	0.01^*	0.001^*	0.01^*	0.06

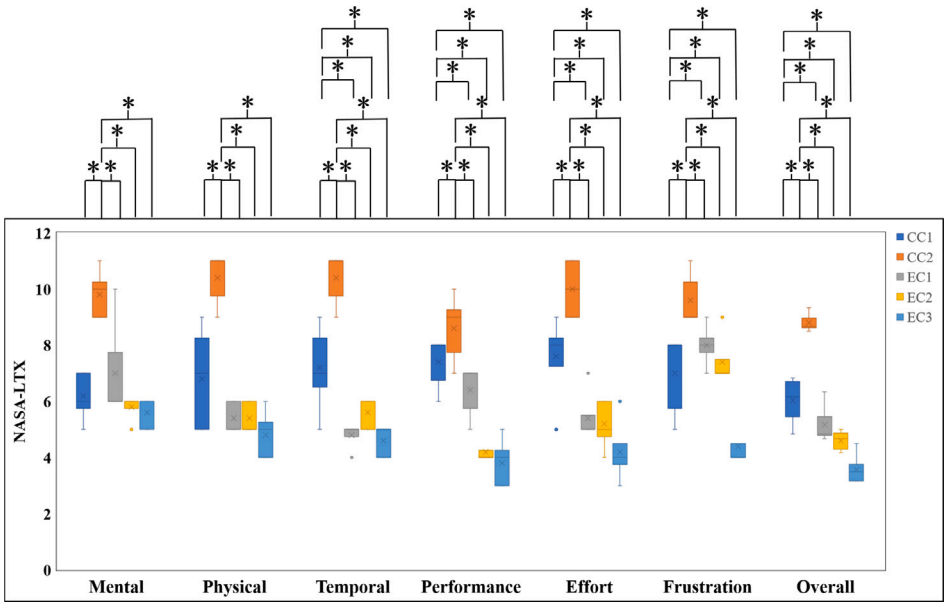

Fig. 9. NASA-TLX scores for individual questions. Significant differences are denoted with an asterisk and line.

Table 7
Simulator Sickness Questionnaire data (User study 1).

Condition	PreAvg± std. dev.	PostAvg± std. dev.	p
CC_1	1.53 ± 0.74	1.60 ± 0.81	0.23
CC_2	1.74 ± 0.67	5.42 ± 0.74	0.001^*
EC_1	1.87 ± 1.02	1.99 ± 0.71	0.45
EC_2	1.73 ± 0.66	2.03 ± 0.87	0.29
EC_3	1.89 ± 0.85	1.97 ± 0.93	0.33

objects and the user were initially placed at random locations within the scene and marked as incomplete. When the positional error of the manipulated object was less than 0.1 cm (target color changed to red), the rotational error was less than 1° (target color changed to blue), and the scaling error was less than 0.005 (target color changed to cyan), the task automatically ended the manipulation of that object. The completion time of each target task and the corresponding accuracy were recorded. Each object was allowed a maximum manipulation time of 2.5 min, and if exceeded, the manipulation of the current object would automatically end. The accuracy of the object manipulation was recorded, and the object was marked as incomplete.

In this scene, participants were required to manipulate sofa (Fig. 10(a)), TV (Fig. 10(b)), bunny (Fig. 10(c)), and chair (Fig. 10(d)) to their respective transparent green target positions (as shown in Fig. 10). The scene size was 10 m × 10 m. The dimensions of the sofa were1.0 m × 0.5 m × 0.2 m, the TV measured 0.5 m × 0.8 m × 0.1 m, the rabbit had dimensions of 0.9 m × 0.8 m × 0.8 m, and the chair measured 1.0 m × 0.7 m × 0.7 m.

6.3. Procedure

In the experiment, each condition was repeated 5 times, resulting in a total of 600 trials ($24 participants \times 5 techniques \times 5 repetitions$). The entire experiment lasted approximately 60 minutes. Participants first completed a questionnaire to collect their demographic information. They were then introduced to the experimental task and VR equipment and instructed to complete the trials as quickly and accurately as possible. Next, participants were instructed to put on the VR headset and engage in a 5-minute VR experience. The VR experiment consisted of five sessions corresponding to the five operational techniques. Each session started with a warm-up phase consisting of ten practice trials to familiarize participants with the techniques, followed by the formal testing trials. After each session, we collected user feedback using the NASA-TLX questionnaire.

6.4. Metrics

The task performance was measured using objective metrics: (1) Task completion time, in seconds, which measured the time from when the participant pointed to the object until the manipulation was completed (when the participant placed the object within the accuracy range). If the participant ran out of time, the time recorded was 2.5 min; (2) Position error, in millimeters, indicating the distance between the center of the manipulated object and the center of the target position at the end of manipulation; (3) Rotation error, in degrees, representing the angular difference between the local coordinate system of the manipulated object and the target coordinate system at the end of manipulation. If the angle difference of the three coordinate axes is α, β, γ, the rotation error is $\sqrt{\alpha^2 + \beta^2 + \gamma^2}$; (4) Scaling error, measured in times, indicating the absolute difference between the diagonal length of the manipulated object's bounding box (BBO) and the diagonal length of the target BBO, divided by the diagonal length of the target BBO; (5) HRV feature values. The user workload was measured using the standard NASA TLX questionnaire.

The statistical methods employed in this study were the same as those used in User Study 1.

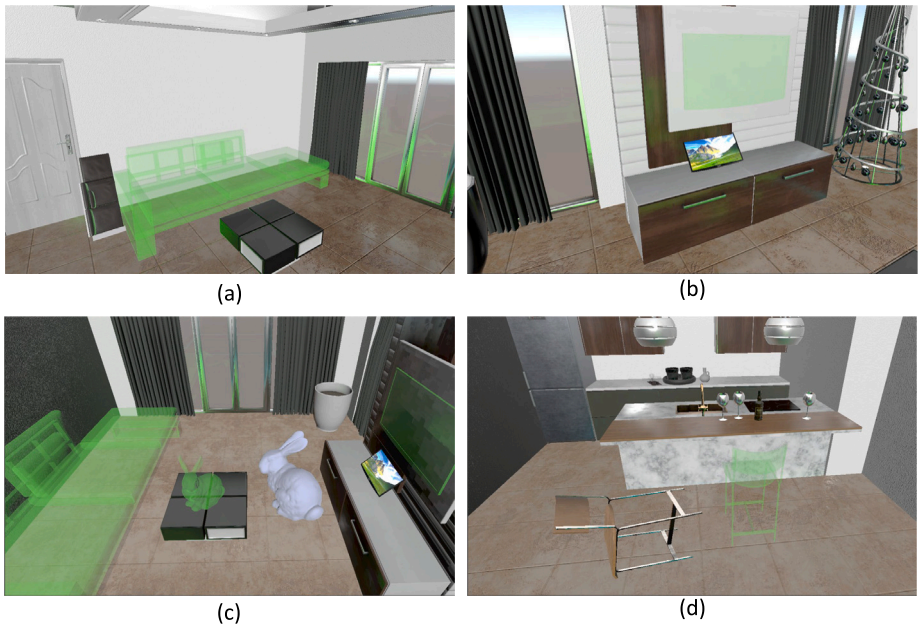

Fig. 10. (a) sofa; (b) TV; (c) bunny; (d) chair.

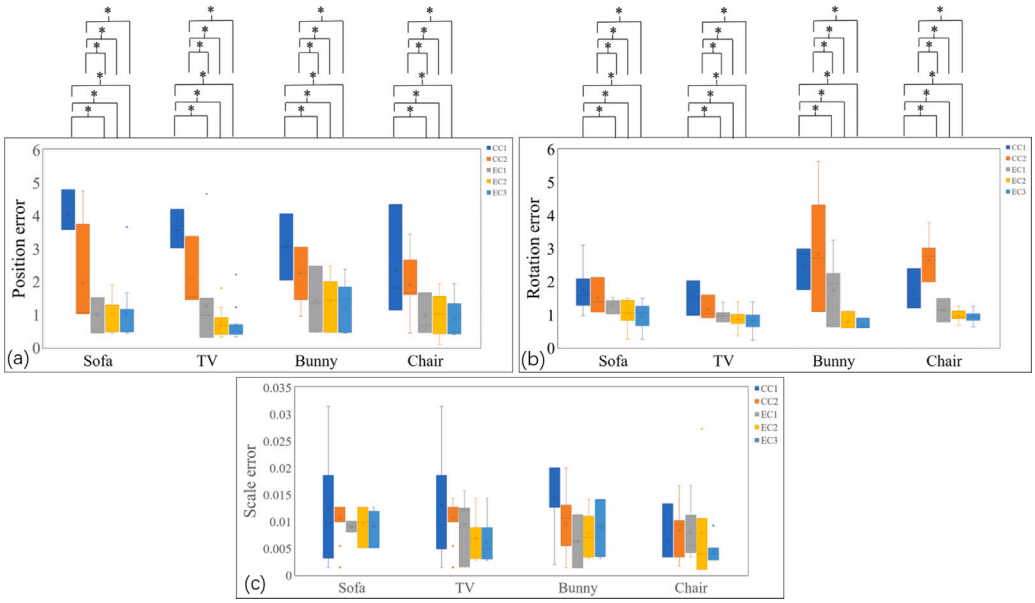

Fig. 11. (a) is the position error. (b) is the rotation error. (c) is the scale error. Significant differences are denoted with an asterisk and line.

6.5. Results

Fig. 11(a) gives the position error. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p = 0.023$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{3.349,130.240} = 47.64, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly smaller than for CC_1 and CC_2. Fig. 11(b) gives the rotation error. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p = 0.023$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.074,80.878} = 32.166, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly smaller than for CC_1 and CC_2. Fig. 11(c) gives the scale error. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p = 0.023$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{3.026,117.997} = 3.837, P = 0.057$). Post-hoc analysis reveals that EC_3 was not significantly smaller than for CC_1 and CC_2.

Table 8
The percentage of participants who successfully completed the task.

	CC_1	CC_2	EC_1	EC_2	EC_3
Sofa	4%	0%	12.54%	16.75%	33.3
TV	8%	4%	16.75%	29.2%	42.6
Bunny	4%	0%	12.54%	25%	37.5
Chiar	8%	4%	20.8%	25%	42.6

Fig. 12 gives the Time. Statistical significance is indicated by an asterisk. The sphericity assumption is violated: $p = 0.023$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.074,30.728} = 39.266, P <$

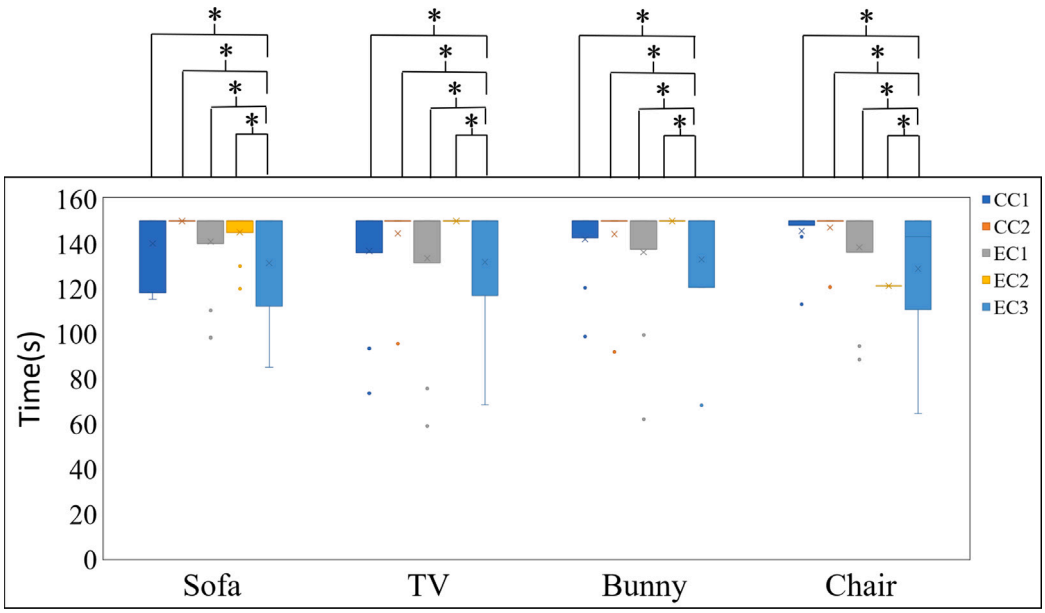

Fig. 12. The complete time, in seconds.

Table 9
Median values of HRV parameters and corresponding Friedman tests results of user study 2.

HRV parameter	Baseline (Mdn)	CC_1	CC_2	EC_1	EC_2	EC_3	p_1	p_2	p_3	p_4
$AVNN$	824.2	730.8	710.9	817.8	797.6	821.1	< 0.01*	< 0.01*	0.02*	0.99
$SDNN$	53.2	57.4	59.8	53.6	54.6	53.3	0.046*	< 0.01*	0.83	0.06
$RMSSD$	47.5	27.19	25.2	31.42	27.8	31.3	0.034*	0.01*	0.056	0.07
$pNN50$	9.32	4.77	3.36	9.14	5.54	9.19	< 0.01*	< 0.01*	0.117	0.97
LF/HF	2.03	3.55	3.96	2.33	2.655	2.17	< 0.01*	< 0.01*	0.04*	0.03*

0.001). Post-hoc analysis reveals that EC_3 was significantly smaller than for CC_1 and CC_2. Table 8 shows the percentage of participants who could place the object within the time limit. On average, close to half of the participants could complete the task using EC_3, while significantly fewer individuals could complete the task using other techniques.

Table 9 gives the HRV. Statistical significance is indicated by an asterisk. For $AVNN$, the sphericity assumption is verified: $p = 0.053$. The overall ANOVA reveals significant differences between the five conditions: ($F_{4,24} = 78.873, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2 and EC_1. For $SDNN$ (Standard deviation of N-N intervals in ms), the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.635,14.711} = 7.433, P = 0.008$). Post-hoc analysis reveals that EC_3 was significantly smaller than for CC_2 and EC_1. For $RMSSD$ (Root mean square of successive differences of normal-to-normal intervals in ms), the sphericity assumption is violated: $p = 0.049$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.028,18.254} = 16.982, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2, and EC_1. For $pNN50$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.098,18.833} = 18.704, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2, and EC_1. For LF/HF, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.397,12.571} = 22.595, P < 0.001$). Post-hoc analysis reveals that EC_3 was significantly higher than for CC_1, CC_2, EC_1 and EC_2. The EC_3 method significantly reduces psychological stress compared to CC_1, CC_2, and EC_1.

Fig. 13 gives the task load. Statistical significance is indicated by an asterisk. The positive fraction performance is replaced by its complement, so smaller values are more favorable. We tested six aspects of the spherical assumption in NASA-TLX and applied Greenhouse–Geisser correction when necessary. For $Mental$, the sphericity assumption is violated: $p = 0.008$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{4,36} = 78.873, P < 0.001$). For $Physical$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.102,18.914} = 66.954, P < 0.001$). For $Temporal$, the sphericity assumption is violated: $p = 0.0021$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.106,18.954} = 120.953, P < 0.001$). For $Performance$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.339,21.055} = 60.650, P < 0.001$). For $Effort$, the sphericity assumption is violated: $p = 0.003$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.068,18.610} = 54.851, P < 0.001$). For $Frustration$, the sphericity assumption is violated: $p = 0.002$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{1.809,16.283} = 45.407, P < 0.001$). For $Overall$, the sphericity assumption is violated: $p < 0.001$. After applying the Greenhouse–Geisser correction, the overall ANOVA reveals significant differences between the five conditions: ($F_{2.482,22.342} = 184.826, P < 0.001$). Compared to CC_2, EC_1, EC_2, and EC_3 show significant improvements in all six aspects. Compared to CC_1, EC_3 significantly improves four aspects, except for Metal and Physical. Compared to EC_1 and EC_2, EC_3 significantly improves

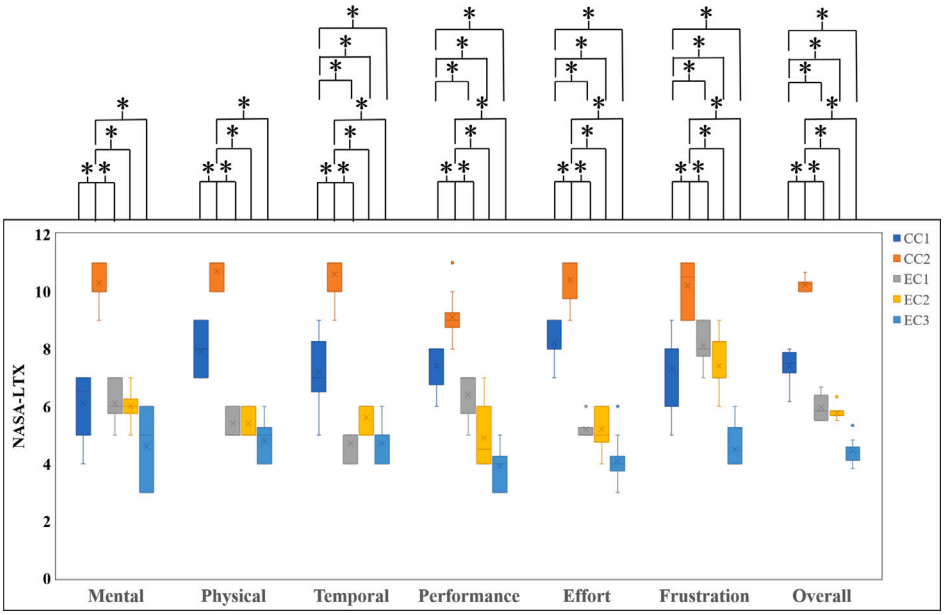

Fig. 13. NASA-TLX scores for individual questions. Significant differences are denoted with an asterisk and line.

Table 10
Simulator Sickness Questionnaire data (User study 2).

Condition	PreAvg± std. dev.	PostAvg± std. dev.	p
CC_1	1.87 ± 0.94	2.12 ± 1.12	0.29
CC_2	1.92 ± 0.82	4.98 ± 1.26	0.001*
EC_1	1.83 ± 0.99	2.09 ± 1.12	0.23
EC_2	1.84 ± 0.86	2.12 ± 1.07	0.28
EC_3	1.83 ± 0.89	2.07 ± 1.13	0.46

effort and frustration. Overall, EC_3 significantly improves total score compared to CC_1, CC_2, EC_1, and EC_2.

We used the standard simulator sickness questionnaire (SSQ) (Kennedy et al., 1993) (Table 10) to measure simulator sickness. The SSQ was managed before and after the experiment for each task and each condition. The SSQ scores are not normally distributed, and the differences before and after the Wilcoxon signed-rank test were used. These differences are not statistically significant with CC_1, CC_2, EC_1, EC_2, and EC_3. but the difference is statistically significant with CC_2.

6.6. Discussion

The results from Tables 2, 3, 4 and Fig. 12 support **H1**, indicating that the effectiveness of our method can be attributed to the following reasons: (1) Our method achieves manipulation gains greater than 1 in each frame when the object is far from the target in the user's view. This allows users to manipulate the object to the vicinity of the target quickly. As the object approaches the target in the user's view, our method generates gains less than 1, reducing the likelihood of excessive manipulation. (2) The manipulation gain is automatically calculated based on the view, eliminating the need for users to consider the speed of head movement, which is difficult to control.

In general, EC_3 is more efficient than EC1 and EC_2 due to the following reasons: In fine manipulation, controlling the speed of head movement is challenging. Based on inter-frame gaze quality, it is necessary to ensure the presence of the manipulated object in the current frame. However, the object may not be present in the current frame in coarse manipulation. In fine manipulation, the object remains consistently present in the current frame, and the inter-frame gaze quality is stable, making it suitable for fine manipulation. While EC_1 is inefficient in controlling the speed of head movement during fine manipulation, EC_3 utilizes inter-frame gaze quality, making it suitable for fine manipulation. Therefore, EC_3 is faster than EC_2. During coarse manipulation, EC_2 tends to lose the manipulated object in the current frame, resulting in a manipulation gain of 0. However, EC_3, based on a velocity-based method, is not affected by this issue, making it more efficient than EC_1. Additionally, from Tables 3, 4, and 5, it can be observed that EC3 outperforms EC_1 and EC_2 in terms of efficiency, primarily due to significantly lower translation and rotation manipulation times, while the scaling manipulation time does not decrease significantly.

The results from Table 1, Fig. 9, Table 9, and Fig. 13 support H2, indicating that the task load of our method is significantly reduced compared to CC_1 and CC_2, possibly due to the following reasons: CC_1 and CC_2 maintain a constant gain of 1, requiring a long time for users to manipulate the object to the target and causing fatigue. Particularly, CC_2, with a smaller range of head movement compared to hands, involves more repetitions, leading to increased fatigue and psychological pressure. Overall, the task load of EC_3 is smaller than that of EC_2 and EC_3. From Figs. 9 and 13, it can be observed that compared to EC_1 and EC_2, EC_3 significantly improves effort and frustration. This improvement may be attributed to EC_3 addressing the difficulties in controlling head movement during fine manipulation in EC_1 and the issue of losing the manipulated object in the current frame during coarse manipulation in EC_2, thereby enhancing the feasibility of the method.

7. Conclusion, limitation, future

We proposed a novel method for object manipulation in virtual reality (VR) based on head movements to improve efficiency and accuracy. We introduced the concept of head manipulation space and conducted experiments to collect data on head manipulation space to determine the manipulable region. We proposed a new method based on head velocity and inter-frame gaze quality to enhance the efficiency and accuracy of head manipulation. Finally, we designed two user studies to evaluate the performance of our head-based object manipulation method. The results demonstrated that our method is feasible in terms of task completion efficiency and accuracy compared to state-of-the-art methods, significantly reducing user fatigue and motion sickness. Furthermore, our method significantly improved usability and reduced task load. Our approach lays the foundation for head-based object

manipulation in virtual and augmented reality and provides a new manipulation method for scenarios where hand-based object manipulation is not suitable.

One limitation of our method is that it requires a 3D user interface (3DUI) for manipulating the types of manipulation and switching between manipulated objects, which significantly impacts the efficiency of our method. Another limitation is that our method visualizes the head manipulation space primarily in the central field of view, affecting the comfort aspect of our method. Another limitation is using a handle to start the task. For future work, it may be necessary to consider an approach that utilizes an HMTS (Head Motion Tracking System) to facilitate the entire process, thereby simplifying the user experience and enhancing the autonomy of the system. For example, interaction protocols based on head movements could be developed to enable users to initiate and control tasks with simple head movements. Future work can focus on removing the need for a 3DUI and allowing users to switch manipulation types and objects using eye gaze. Another avenue for future work is selecting a suitable visualization method for head manipulation space. In our user studies, we did not investigate the feasibility of head movement object manipulation in older users. Future research could explore the feasibility of head movement object manipulation in older users.Future work could consider emphasizing the potential ergonomic impact of prolonged use of head-based manipulation, which may raise concerns about neck stress or related issues.

CRediT authorship contribution statement

Xiaolong Liu: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Lili Wang:** Resources, Validation, Writing – review & editing. **Wei Ke:** Resources. **Sio-Kei Im:** Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work is supported by the National Natural Science Foundation of China through Projects 61932003 and 62372026, Beijing Science and Technology Plan Project Z221100007722004, and the National Key R&D plan 2019YFC1521102.

References

Aguerreche, L., Duval, T., Lécuyer, A., 2009. 3-hand manipulation of virtual objects. In: Eurographics.

Bowman, D.A., Hodges, L.F., 1999. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. ACM.

Castaldo, R., Melillo, P., Pecchia, L., 2015. Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review. In: Lacković, I., Vasic, D. (Eds.), 6th European Conference of the International Federation for Medical and Biological Engineering. Springer International Publishing, Cham, pp. 1–4.

Cohen, J., 2013. Statistical Power Analysis for the Behavioral Sciences. Academic Press.

Frees, S., Kessler, G.D., Kay, E., 2007. PRISM interaction for enhancing control in immersive virtual EnvironmentsPriss. ACM Trans. Comput.-Hum. Interact. (ISSN: 1073-0516) 14 (1), http://dx.doi.org/10.1145/1229855.1229857, 2–es, URL http://dx.doi.org/10.1145/1229855.1229857.

Gelman, A., 2005. Analysis of variance. Qual. Control Appl. Statist 20 (1), 295–300.

Gloumeau, P.C., Stuerzlinger, W., Han, J.H., 2020. PinNPivot: Object manipulation using pins in immersive virtual environments. IEEE Trans. Vis. Comput. Graphics PP (99), 1.

Hart, S., 2006. Nasa-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 50, pp. 904–908.

Hart, S., Staveland, L., 1988. Development of NASA-TLX (task load index): Results of empirical and theoretical research. Adv. Psychol. 52, 139–183.

Kennedy, R., Lane, N.E., Berbaum, K., Mg, L., 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol 3, 203–220.

Kim, J., Cha, J., Lee, H., Kim, S., 2017. Hand-free natural user interface for VR HMD with IR based facial gesture tracking sensor. In: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. VRST '17, Association for Computing Machinery, New York, NY, USA, ISBN: 9781450355483, http://dx.doi.org/10.1145/3139131.3143420, URL http://dx.doi.org/10.1145/3139131.3143420.

Kim, H., Lee, G.A., Billinghurst, M., 2015. A non-linear mapping technique for bare-hand interaction in large virtual environments. In: Australasian Computer-Human Interaction Conference.

Liu, C., Orlosky, J., Plopski, A., 2020a. Eye gaze-based object rotation for head-mounted displays. In: Proceedings of the 2020 ACM Symposium on Spatial User Interaction. SUI '20, Association for Computing Machinery, New York, NY, USA, ISBN: 9781450379434, http://dx.doi.org/10.1145/3385959.3418444, URL http://dx.doi.org/10.1145/3385959.3418444.

Liu, C., Plopski, A., Orlosky, J., 2020b. OrthoGaze: Gaze-based three-dimensional object manipulation using orthogonal planes. Comput. Graph..

Liu, X., Wang, L., Luan, S., Shi, X., Liu, X., 2022. Distant object manipulation with adaptive gains in virtual reality. In: 2022 IEEE International Symposium on Mixed and Augmented Reality. ISMAR, pp. 739–747. http://dx.doi.org/10.1109/ISMAR55827.2022.00092.

Mauchly, J.W., 1940. Significance Test for Sphericity of a Normal n-Variate Distribution. Ann. Math. Stat. 11 (2), 204–209. http://dx.doi.org/10.1214/aoms/1177731915, URL http://dx.doi.org/10.1214/aoms/1177731915.

Mendes, D., Caputo, F.M., Giachetti, A., Ferreira, A., Jorge, J., 2019. A survey on 3D virtual object manipulation: From the desktop to immersive virtual environments. In: Computer Graphics Forum.

Mendes, D., Sousa, M., Lorena, R., Ferreira, A., Jorge, J., 2017. Using custom transformation axes for mid-air manipulation of 3D virtual objects. In: Acm Symposium. pp. 1–8.

Monteiro, P., Gonçalves, G., Coelho, H., Melo, M., Bessa, M., 2021. Hands-free interaction in immersive virtual reality: A systematic review. IEEE Trans. Vis. Comput. Graphics 27 (5), 2702–2713. http://dx.doi.org/10.1109/TVCG.2021.3067687.

Narciso, D., Melo, M., Rodrigues, S., Cunha, J.P., Vasconcelos-Raposo, J., Bessa, M.E., 2022. Using heart rate variability for comparing the effectiveness of virtual vs real training environments for firefighters. IEEE Trans. Vis. Comput. Graphics 1. http://dx.doi.org/10.1109/TVCG.2022.3156734.

Nguyen, T.T.H., Duval, T., 2013. 3-point++: a new technique for 3D manipulation of virtual objects.

Nguyen, T.-T.H., Duval, T., Pontonnier, C., 2014. A new direct manipulation technique for immersive 3D virtual environments. In: International Conference on Artificial Reality and Telexistence.

Osawa, N., 2008. Two-handed and one-handed techniques for precise and efficient manipulation in immersive virtual environments. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Remagnino, P., Porikli, F., Peters, J., Klosowski, J., Arns, L., Chun, Y.K., Rhyne, T.-M., Monroe, L. (Eds.), Advances in Visual Computing. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 987–997.

Park, K.-B., Choi, S.H., Lee, J.Y., Ghasemi, Y., Mohammed, M., Jeong, H., 2021. Hands-free human–robot interaction using multimodal gestures and deep learning in wearable mixed reality. IEEE Access 9, 55448–55464. http://dx.doi.org/10.1109/ACCESS.2021.3071364.

Pierce, J.S., Pausch, R., 2002. Comparing voodoo dolls and HOMER: exploring the importance of feedback in virtual environments. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 105–112.

Poupyrev, I., Billinghurst, M., Weghorst, S., Ichikawa, T., 1998. The Go-Go interaction technique: Non-linear mapping for direct manipulation in VR.

Rantamaa, H.-R., Kangas, J., Kumar, S.K., Mehtonen, H., Järnstedt, J., Raisamo, R., 2023. Comparison of a VR stylus with a controller, hand tracking, and a mouse for object manipulation and medical marking tasks in virtual reality. Appl. Sci. (ISSN: 2076-3417) 13 (4), http://dx.doi.org/10.3390/app13042251, URL https://www.mdpi.com/2076-3417/13/4/2251.

Song, P., Goh, W.B., Hutama, W., Fu, C.-W., Liu, X., 2012. A handle bar metaphor for virtual object manipulation with mid-air interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12, Association for Computing Machinery, New York, NY, USA, ISBN: 9781450310154, pp. 1297–1306. http://dx.doi.org/10.1145/2207676.2208585, URL http://dx.doi.org/10.1145/2207676.2208585.

Vercoulen, J.H., Swanink, C.M., Fennis, J.F., Galama, J.M., van der Meer, J.W., Bleijenberg, G., 1994. Dimensional assessment of chronic fatigue syndrome. J. Psychosom. Res 38 (5), 383–392.

Voss, A., Schroeder, R., Heitmann, A., Peters, A., Perz, S., 2015. Short-term heart rate variability—influence of gender and age in healthy subjects. PLoS One 10 (3), e0118308.

Wilkes, C., Bowman, D.A., 2008. Advantages of velocity-based scaling for distant 3D manipulation. In: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology. VRST '08, Association for Computing Machinery, New York, NY, USA, ISBN: 9781595939517, pp. 23–29. http://dx.doi.org/10.1145/1450579.1450585, URL http://dx.doi.org/10.1145/1450579.1450585.

Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J., 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 143–146.

Xiaolong Liu is a Ph.D. student in the School of Computer Science and Engineering of Beihang University, China. His current research focuses on virtual reality, augmented reality, and HCI.

Lili Wang received her Ph.D. degree from Beihang University, Beijing, China. She is a professor with the School of Computer Science and Engineering of Beihang University and a researcher with the State Key Laboratory of Virtual Reality Technology and Systems. Her interests include virtual reality, real-time rendering and HCI.

Wei Ke received his Ph.D. degree from the School of Computer Science and Engineering, Beihang University. He is currently a professor at the Faculty of Applied Sciences, Macao Polytechnic University. His research interests include programming languages, image processing, computer vision, and tool support for component-based engineering and systems.

Sio Kei Im received his Ph.D. degree in Electronic Engineering from Queen Mary University of London (QMUL), United Kingdom. He is a professor at the Faculty of Applied Sciences, Macau Polytechnic University. His research interests include video coding, image processing, machine learning for NLP and multimedia.