A Mixed Reality Training System for Hand-Object Interaction in Simulated Microgravity Environments

Kanglei Zhou[1]* Chen Chen[1]* Yue Ma[1] Zhiying Leng[1] Hubert P. H. Shum[2] Frederick W. B. Li[2] Xiaohui Liang[1,3]†

[1] State Key Laboratory of Virtual Reality Technology and Systems, Beihang University [2] Durham University [3] Zhongguancun Laboratory
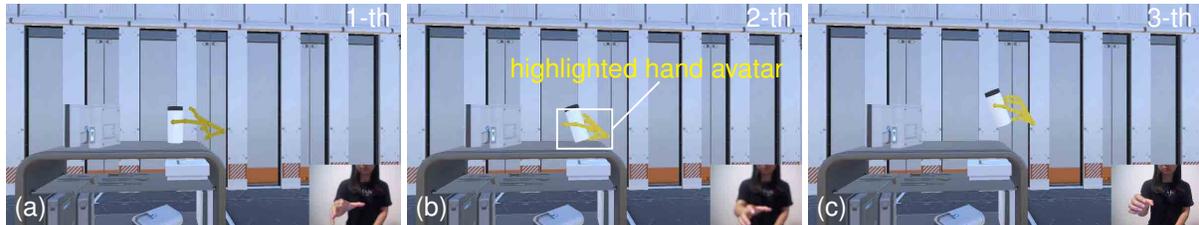
Figure 1: Object grasping in a simulated microgravity environment: (a) depicts a session where the object remains ungrasped; (b) and (c) show sessions where the object is successfully grasped. The highlighted hand avatar offers visual feedback upon contact.

ABSTRACT

As human exploration of space continues to progress, the use of Mixed Reality (MR) for simulating microgravity environments and facilitating training in hand-object interaction holds immense practical significance. However, hand-object interaction in microgravity presents distinct challenges compared to terrestrial environments due to the absence of gravity. This results in heightened agility and inherent unpredictability of movements that traditional methods struggle to simulate accurately. To this end, we propose a novel MR-based hand-object interaction system in simulated microgravity environments, leveraging physics-based simulations to enhance the interaction between the user's real hand and virtual objects. Specifically, we introduce a physics-based hand-object interaction model that combines impulse-based simulation with penetration contact dynamics. This accurately captures the intricacies of hand-object interaction in microgravity. By considering forces and impulses during contact, our model ensures realistic collision responses and enables effective object manipulation in the absence of gravity. The proposed system presents a cost-effective solution for users to simulate object manipulation in microgravity. It also holds promise for training space travelers, equipping them with greater immersion to better adapt to space missions. The system reliability and fidelity test verifies the superior effectiveness of our system compared to the state-of-the-art CLAP system.

Index Terms: Human-centered computing—Human computer interaction—Interaction paradigms—Mixed/augmented reality

1 INTRODUCTION

As manned spaceflight technology continues to advance, an increasing number of individuals are embarking on space exploration journeys. Whether on Earth or in the vast expanse of space, human hands play a vital role in our daily lives, enabling us to interact with our environment and manipulate objects [1]. However, it is important to recognize that there are notable disparities in the perception of self-motion during hand-object interaction between the gravitational environment of Earth and the microgravity conditions experienced in space. This dissimilarity arises due to a phenomenon

---

*denotes that these authors contributed equally to this work.
†e-mail: liang_xiaohui@buaa.edu.cn (corresponding author)

known as sensory conflict [2], where the interpretation of cues, particularly from the vestibular system, is altered when transitioning from a microgravity environment back to Earth. Therefore, it is necessary for space travelers to conduct extensive and specialized training prior to their space missions, which facilitates their rapid adaptation to the unique conditions of microgravity.

Traditional microgravity simulation methods [3, 4, 5], such as parabolic flights and neutral buoyancy tanks, have inherent limitations in accurately replicating the full experience of microgravity. These methods provide only partial weightlessness and struggle to recreate the complex environmental factors encountered in space. Moreover, the accessibility of these methods is restricted, making them less practical and costly for training purposes. To overcome these drawbacks, virtual/augmented/mixed reality (VR/AR/MR) technology has emerged as a promising solution for hand-object interaction in microgravity. Particularly, MR offers distinct advantages that address the limitations of traditional methods. It provides a more realistic and immersive training experience that closely resembles microgravity conditions. Additionally, MR-based training [6, 7] is more accessible, cost-effective, and versatile compared to traditional microgravity simulations.

In this work, we propose a novel hand-object interaction system for simulated microgravity environments. This allows the public to experience space-based manipulation using visual feedback, bypassing the need for force feedback devices. Importantly, our system can refine space travelers' perception of their own motion, aiding them in navigating the sensory challenges induced by microgravity. Utilizing MR technology, the system simulates genuine microgravity hand-object interaction. This immersive method ensures space travelers rapidly adapt to microgravity for space missions.

Achieving a realistic sense of interaction between the real hand and virtual objects is a multi-faceted challenge that involves accurate hand gesture tracking and enabling the virtual hand to effectively interact with virtual objects. The former can be addressed using various hardware devices and algorithms [1, 8]. In this work, we employ the Leap Motion device [9] to capture the movements of users' hands and generate corresponding virtual hand representations. The latter, enabling the real hand to penetrate virtual objects in a realistic and natural manner seamlessly, is a more complex challenge. It stems from the inherent disparities between physical and virtual objects and the necessity to ensure coherent and intuitive interactions between them. To overcome this obstacle, researchers actively explore approaches in advanced physics-based simulations [10, 11] and visual systems [12, 13]. Considering the advantages of being able to accurately model the physical prop-

erties and behaviors of virtual objects, we specifically focused on leveraging physics-based simulations to enhance hand-object interaction in virtual environments.

Hand-object interaction in microgravity presents unique challenges compared to interactions in terrestrial environments due to the absence of gravitational forces. In microgravity, objects exhibit remarkable agility and present manipulation challenges during interactions, particularly in scenarios involving subtle tapping and forceful slapping. This poses a limitation for existing physics-based methods [10, 11] that rely on penetration force calculations after hands are inserted into objects, making them less effective in simulating such interactions. To address this, we propose a new physics-based hand-object interaction model that merges an impulse-based simulation with penetration contact dynamics. The impulse-based approach models instantaneous momentum changes during contact interactions, precisely capturing the dynamics of microgravity hand-object interaction. By accounting for the forces and impulses during contact, our model achieves realistic collision responses and object manipulation without gravity. This method allows for a nuanced simulation of hand-object interaction in microgravity, yielding a more immersive and accurate virtual experience. By addressing unique challenges in microgravity, our approach surpasses traditional physics-based models, offering a specialized solution for realistic interactions in microgravity.

We have conducted a reliability and fidelity test to evaluate the effectiveness of our system. The results highlight that participants perceived our system as more natural and realistic than the state-of-the-art system, CLAP. They reported an enhanced sense of immersion and described the interactions as richer and more engaging. Such feedback underscores our method's ability to amplify the overall user experience, offering a more accurate representation of hand-object interaction in microgravity. Moreover, we undertook quantitative assessments by timing task completions and tallying failed operations during training. These metrics offered concrete insights into our system's performance. Collectively, these evaluations corroborate the efficacy of our approach.

The main contributions are summarized as follows:

- The design and development of a novel MR-based hand-object interaction system in microgravity environments, providing a cost-effective solution to experience object interaction in microgravity through visual feedback.

- The proposal of a physics-based hand-object interaction model that combines impulse-based simulation with penetration contact dynamics, accurately capturing the dynamics of hand-object interaction in microgravity and enabling realistic collision responses and object manipulation in microgravity.

- The evaluation of the proposed MR-based hand-object interaction system through a reliability and fidelity test, demonstrating its effectiveness in providing a more realistic and natural experience compared to the state-of-the-art.

2 RELATED WORK

This section begins with an overview of microgravity simulation techniques, followed by an introduction to hand-object interaction.

2.1 Microgravity Simulation

Simulating microgravity, which is experienced in space environments, is crucial for training space travelers and studying the effects of weightlessness on human activities [5, 7]. According to the microgravity simulation support facility of NASA [14], various traditional microgravity simulation methods have been utilized to replicate partial weightlessness on Earth.

Among them, parabolic flights and neutral buoyancy tanks are two commonly used methods. Parabolic flights [15, 16] involve specially designed aircraft flying in a parabolic trajectory. During the upward portion of the trajectory, the aircraft experiences a period of freefall, creating a brief sensation of weightlessness for passengers on board. While parabolic flights offer valuable firsthand experiences of microgravity, they are limited by the short duration of weightlessness and the logistical challenges associated with conducting experiments in flight. Neutral buoyancy tanks [17, 18], on the other hand, provide an environment that mimics the feeling of weightlessness by using water to counteract buoyancy forces. Space travelers and researchers can perform tasks and experiments in the water, experiencing a reduced sense of gravity. However, neutral buoyancy tanks have their limitations, including the necessity of bulky equipment and the inability to replicate certain aspects of the space environment, such as the absence of air resistance. While these methods offer some level of simulation, they fall short of reproducing the full experience of microgravity and replicating the complex environmental factors encountered in space. Additionally, they are often limited in accessibility and cost, making them less practical for extensive training purposes.

2.2 Hand-Object Interaction

Hand-object interaction is a fundamental aspect of MR experiences, as it enables users to engage with and manipulate virtual objects in a more intuitive and immersive manner [19]. We will provide an overview of the advancements made in enhancing hand-object interaction in MR from two perspectives: hand perception techniques and interaction relationship reasoning.

Hand Perception Technique   According to the way user hand perception is achieved, these methods can be categorized into tracking-based methods and controller-based methods.

On the one hand, tracking-based methods [10, 20, 21, 22, 23] utilize technologies such as motion capture or depth sensing to accurately track and interpret the movements and gestures of the user's hands in real-time response. The Leap Motion device is often utilized for hand-object interaction due to its high accuracy and precision in tracking hand movements and gestures. Delrieu et al. [23] utilized the Leap Motion device to establish a connection between a virtual hand and a tracked hand, enabling grasping assistance through the implementation of virtual springs. Our approach also leverages the Leap Motion device for tracking human hands. Furthermore, Zhou et al. [1] utilized the Microsoft HoloLens 2 to track human hands, enabling robust hand manipulation in MR. These methods allow for more natural and intuitive interaction with virtual objects, as the user's hand movements are directly translated into corresponding actions within the virtual environment.

On the other hand, controller-based methods [24, 25] rely on handheld input devices or controllers that provide a means for users to interact with virtual objects. Oprea et al. [25] presented a grasping system that leverages headset tracking and motion controllers, allowing a human operator to embody virtual human or robot agents and freely navigate and interact with objects in virtual environments. The system is capable of handling various object geometries without the need for predefined grasp animations, automatically fitting the fingers to object shape and geometry. Additionally, Han et al. [24] utilized Oculus Quest 2 head-mounted displays and VR controllers to enable users to manipulate objects with dexterity in virtual environments. By mapping the MR controller to the virtual hand and using a deep neural network, the system synthesized dynamic hand motions and joint orientations for a realistic and immersive hand-object interaction experience. While these methods can offer precise control, they may lack the level of realism and naturalness achieved by tracking-based methods.

To avoid the limitation of controller-based methods, we incorporate tracking-based techniques to achieve more natural and realistic free-hand perception [26], allowing for natural interactions with virtual objects. Our system enhances the user experience by providing highlighting hand avatars upon contact.

Interaction Relationship Reasoning According to interaction relationship reasoning, these methods can be broadly categorized into physics-based and learning-based methods.

On the one hand, physics-based methods [10, 20, 23, 27] in hand-object interaction relationship reasoning focus on modeling the physical properties and behaviors of objects and hands. These methods typically involve applying principles of physics, such as Newtonian mechanics and rigid body dynamics, to simulate realistic interactions between the hand and objects. They aim to accurately represent the forces, collisions, and constraints involved in hand-object interaction, providing a physically grounded simulation of the interaction dynamics. Höll et al. [10] employed the Coulomb friction model to identify contact points between the user's real hand and virtual objects in order to simulate various actions, including pushing, pulling, grasping, and dexterous manipulations. Furthermore, several previous works [28, 29] have employed a distance field to represent the spatial relation between a hand and an object, enabling accurate calculations of the proximity between the hand and object, and facilitating realistic hand-object interaction and manipulation. These methods require a comprehensive understanding of the physical properties of objects and hands and often involve complex calculations and simulations.

On the other hand, learning-based methods [12, 24, 30, 31, 32] for hand-object interaction relationship reasoning leverage machine learning and artificial intelligence techniques to infer and reason about the interaction between hands and objects. These methods involve training models on large datasets of hand-object interaction to learn patterns, relationships, and context-aware information. Fan et al. [12] proposed a large-scale dataset to bridge the gap between human and machine understanding of hand-object interaction. It supports consistent motion reconstruction and interaction field estimation, aiming to reconstruct and estimate spatial-temporally consistent hand-object interaction. Deep learning, reinforcement learning, and other learning algorithms can capture the complex and subtle aspects of hand-object interaction [33, 34]. Streli et al. [30] introduced a data-driven method to detect finger pinching for off-screen selection events. It enhances the interaction by incorporating a haptic actuator in their wrist device, providing users with tactile feedback to confirm their selections and enabling them to interact with virtual content through actions like selecting, grabbing, and dropping. Moreover, Han et al. [24] incorporated reinforcement learning to build a visually and physically plausible hand manipulation system. While learning-based methods allow for more flexible and context-aware interaction behaviors, the need for extensive training data makes it challenging to generalize well to new objects or scenarios that were not included in the training set. Additionally, they may struggle with capturing complex and subtle aspects of hand-object interaction, as it can be difficult to represent and model all the intricate dynamics and constraints involved.

Our work addresses the constraints of learning-based methods in hand-object interaction by leveraging the advantages of physics-based techniques. Through developing a novel physics-based method, we can achieve more realistic and accurate simulations of hand-object interaction in virtual settings, thus elevating the fidelity and authenticity of interactions within MR.

3 METHODOLOGY

In this section, we first provide an overview of the entire framework. Subsequently, we introduce the key phases of the system.

3.1 Framework Overview

The pipeline of our system is shown in Fig. 2. It consists of three key components: (a) a hand tracking module for capturing and tracking hand movements, (b) a collision detection module to detect interactions between the hand and objects, and (c) an interaction relationship reasoning module to simulate natural interaction.
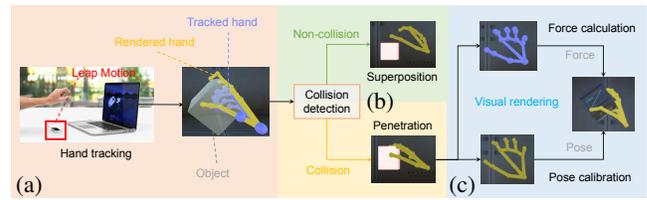

Figure 2: The pipeline of our system: (a) hand tracking, (b) collision detection, and (c) interaction relationship reasoning.

Through a simple yet effective window filtering method, the hand tracking module can precisely capture hand movements in microgravity, enabling stable object manipulation. In microgravity, the dynamics of hand-object interaction deviate notably from those in gravity-based scenarios. The collision detection module is tailor-made to accommodate these distinct dynamics posed by microgravity, ensuring precise detection of contact and collisions between the hand and virtual objects. This leads to realistic and consistent interactions, even without gravitational forces. The interaction relationship reasoning module factors in the specific characteristics of hand-object interaction in microgravity. Incorporating impulse-based simulation, it examines the spatial relationships, contact points, and the involved forces. This permits intelligent reasoning and informed decision-making. Consequently, the system can adeptly adapt and offer a more intuitive and responsive interaction experience in microgravity settings.

Incorporating these modules within a microgravity context paves the way for groundbreaking hand-object interaction. Users can undertake various tasks, such as assembling components, handling equipment, and running experiments, which are peculiar to microgravity environments. This system presents a genuine and efficient platform for the microgravity experience.

3.2 Hand Tracking

In microgravity environments, accurate tracking of hand poses and movements is essential for hand-object interactions. The Leap Motion device is selected due to its advanced algorithms, which offer precise hand tracking. This allows users to interact with virtual objects using natural hand movements, eliminating the need for extra sensors or equipment. The device records the hand's positions, orientations, and movements, providing a dependable basis for simulating interactions between hands and virtual objects. The hand skeleton,

Figure 3: The visualization of the hand skeleton tracked by Leap Motion.

depicted in Fig. 3 and comprising five fingers, a palm, and a wrist, is monitored by the Leap Motion device.

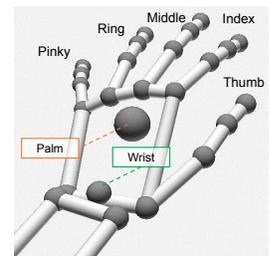In this work, we find that the lack of gravitational forces in microgravity exacerbates the impact of noise on hand-object interaction. Small disturbances and noise can significantly affect hand movements and object interactions, posing challenges in grasping and manipulating objects accurately. As shown in Fig. 4(a), the trajectory of the hand exhibits initial fluctuations due to noise but remains able to grasp the object. However, over time, the movement becomes increasingly unstable, requiring constant adjustments to maintain the grip. This negatively impacts the user experience.

To address the issue of unreliable raw data and potential hand shaking during prolonged operation in microgravity environments,
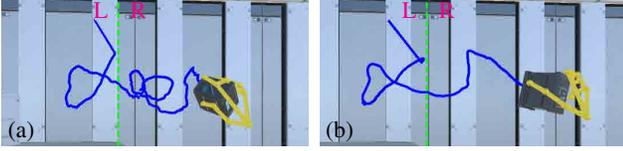
Figure 4: The visualization of motion trajectories of two manipulation experiments under different conditions: (a) the original without filtering and (b) our method using the filtered strategy.

Figure 5: The illustration of a single-point collision.

we have employed a simple yet effective window filtering strategy, which can be methodically represented as:

$$\boldsymbol{x}_i = \frac{1}{F} \sum_{f=0}^{F} \boldsymbol{x}_{i-f}, \tag{1}$$

where F is the filter window size. In microgravity hand-object interaction, balancing noise suppression and system responsiveness is essential. Preliminary tests indicated that a small F inadequately filtered system noise, producing jerky trajectories. Conversely, a larger F introduced perceptible lag, risking real-time MR interactions. Empirical evaluations revealed $F = 3$ as optimal, balancing noise reduction and system reactivity.

This simple yet effective filtering method can boost both trajectory stability and precision during hand-object interactions in microgravity. As illustrated in Fig. 4, the difference is clear: while Fig. 4(a) displays an unfiltered trajectory wherein the user struggles to manipulate an object to complete a specific task, Fig. 4(b) presents a smoother, filtered version that allows the same user to effortlessly finish the same task. With this approach, we guarantee a stable and consistent user experience in microgravity, devoid of the irregularities seen in unfiltered movements.

3.3 Collision Detection

In hand-object interaction, accurate collision detection is crucial for determining the contact points between the tracked hands and the virtual objects. This is essential to achieve realistic and natural interactions, as it prevents the tracked hands from penetrating or intersecting with the virtual objects. To accomplish this, we utilize both tracked hands and rendered hands.

Tracked hands act as virtual counterparts of the user's actual hand within the virtual environment. They faithfully replicate the movements and positions of the user's real hand, ensuring seamless integration between the physical and virtual realms. On the other hand, rendered hands have the responsibility of providing users with a natural and visually appealing representation of their hand movements during interactions with virtual objects. By delivering accurate visual feedback, rendered hands enhance the overall user experience and contribute to the realism of the interaction.

The collision detection process determines if a collision has occurred by comparing the position and orientation of the tracked hands with the geometry of the virtual objects in the scene. For the multipoint collision, it is essential to compute the single-point collision independently for each contact point. Consequently, our subsequent discussion focuses on the single-point collision. As depicted in Fig. 5, a single-point collision between the tracked hand and the virtual object is illustrated, highlighting the contact point $\boldsymbol{c}$. From this, we can determine the contact normal vector $\boldsymbol{n} \in \mathbb{R}^3$. For objects with continuous smooth surfaces, the contact normal aligns with the surface normal of both objects at the point $\boldsymbol{c}$. In this scenario, the surface normals of the two objects are parallel but point in opposite directions.

Knowing the position and orientation of the tracked hand and the virtual object, we can compute the contact points w.r.t. either one
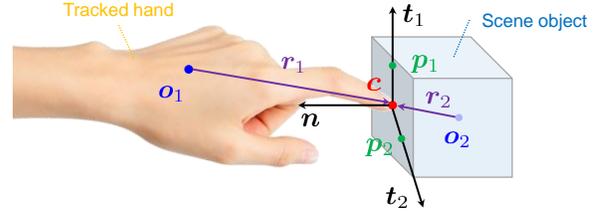
of these objects, i.e.,

$$\boldsymbol{c}_1 = \boldsymbol{c}_2, \text{ s.t. } \begin{cases} \boldsymbol{c}_1 = \boldsymbol{r}_1 + \boldsymbol{o}_1 \\ \boldsymbol{c}_2 = \boldsymbol{r}_2 + \boldsymbol{o}_2 \end{cases}, \tag{2}$$

where the subscripts 1 and 2 refer to the tracked hand and the virtual object, $\boldsymbol{o}$ is the position of the center of mass, and $\boldsymbol{r}$ represents a vector from $\boldsymbol{o}$ to the contact position. Using the body frame BF_i, as outlined in the literature [35], we can rewrite $\boldsymbol{r}_i$ as:

$$\boldsymbol{r}_i = \mathbf{R}_i \left[\boldsymbol{r}_i\right]_{\mathrm{BF}_i}, \tag{3}$$

where $\mathbf{R}_i$ is the rotation matrix of the body frame of i-the object. Now, integrating Eq. (3) into Eq. (2), we obtain:

$$\boldsymbol{c}_i = \mathbf{R}_i \left[\boldsymbol{r}_i\right]_{\mathrm{BF}_i} + \boldsymbol{o}_i. \tag{4}$$

If we differentiate Eq. (4) w.r.t. time, we get the contact velocity of the contact point on body i, i.e.,

$$\dot{\boldsymbol{c}}_i = \dot{\mathbf{R}}_i \left[\boldsymbol{r}_i\right]_{\mathrm{BF}_i} + \dot{\boldsymbol{o}}_i. \tag{5}$$

The columns of $\mathbf{R}_i$ are unit vectors along the axes of BF_i, which are body-fixed vectors. Consequently, we can write the time derivative of a body-fixed vector as the cross product of the angular velocity ω and the body-fixed vector, i.e., Eq. (5) is reformulated as:

$$\begin{aligned} \dot{\boldsymbol{c}}_i &= \left[\omega_i \times \mathbf{R}_i^x \,|\, \omega_i \times \mathbf{R}_i^y \,|\, \omega_i \times \mathbf{R}_i^z\right] \left[\boldsymbol{r}_i\right]_{\mathrm{BF}_i} + \dot{\boldsymbol{o}}_i \\ &= \omega_i \times \left(\mathbf{R}_i \left[\boldsymbol{r}_i\right]_{\mathrm{BF}_i}\right) + \dot{\boldsymbol{o}}_i. \end{aligned} \tag{6}$$

Next, we simplify Eq. (6) to obtain:

$$\begin{aligned} \dot{\boldsymbol{c}}_i &= \omega_i \times \left(\mathbf{R}_i \left[\boldsymbol{r}_i\right]_{\mathrm{BF}_i} + \boldsymbol{o}_i - \boldsymbol{o}_i\right) + \dot{\boldsymbol{o}}_i \\ &= \omega_i \times \left(\boldsymbol{c}_i - \boldsymbol{o}_i\right) + \dot{\boldsymbol{o}}_i \\ &= \omega_i \times \boldsymbol{r}_i + \dot{\boldsymbol{o}}_i. \end{aligned} \tag{7}$$

Now, we can calculate the relative contact velocity v in the normal direction between the tracked hand and the virtual object:

$$v = \boldsymbol{n}^{\top} \left(\dot{\boldsymbol{c}}_1 - \dot{\boldsymbol{c}}_2\right). \tag{8}$$

The relative contact velocity tells us something about how the two objects are moving relative to each other. Then, we know that:

1. If $v > 0$, a separating contact occurs, indicating that in the future, regardless of the size of the time step we consider, the objects will no longer be in contact at the contact point $\boldsymbol{c}$.

2. If $v = 0$, a resting contact occurs, suggesting that the contact point $\boldsymbol{c}$ will persist in the future regardless of how small a time step we take.

3. If $v < 0$, a colliding contact occurs, implying that without any counteraction, such as applying a collision impulse at the contact point, the two objects will penetrate in the future regardless of the size of the time step we consider.
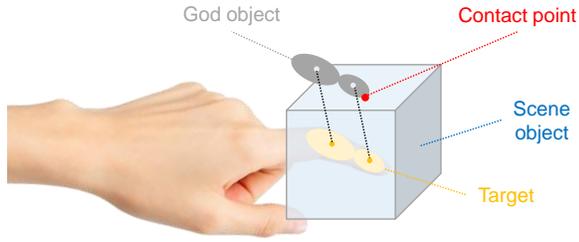
Figure 6: The illustration of the God-object approach.

When no collision is detected, the positions and orientations of the tracked hand and the rendered hand remain consistent, indicating proper alignment between the two. However, if a collision is detected, it indicates that the tracked hands have penetrated or intersected with the virtual objects. This typically occurs when a point on the tracked hand is closer to the surface of the virtual object than a predefined threshold. In such cases, we find the closest 3D point on the object's surface to the contact point.

One challenge we face is updating the contact points during object interaction, particularly when there is slippage along the surface of an object. To address this, we employ the God-Object approach [36, 37], a well-known method for updating contact points on the surface of objects. This involves treating the virtual object as a single cohesive entity, referred to as the God-Object, and updating the contact points based on the interactions between the hand and this virtual entity. As shown in Fig. 6, the tracked hand is experiencing penetration with the virtual object, whereas the rendered hand remains outside of the object. To facilitate this interaction, each tracked finger is associated with a target object per phalanx, while the corresponding rendered finger is represented by a single God object per phalanx. The contact point between the rendered hand and the virtual object enables the computation of direct manipulation, allowing actions such as pushing the object downwards or exerting force in a specific direction. When there is no contact between the user's hand and the virtual object, no force is applied, and movements are determined by the user's intention. However, upon contact, two key factors need to be considered for visual rendering: the depth of penetration of the hand into the object and the specific dynamics required to simulate the interaction accurately. By using the God-Object approach, we can accurately track and update the contact points, ensuring realistic and responsive hand-object interaction within the virtual environment.

3.4 Interaction Relationship Reasoning

In microgravity environments, the traditional penetration contact dynamics approaches [10, 11, 38] that primarily focus on collision detection and response when objects intersect or overlap, may not be suitable for accurately simulating nuanced hand-object interaction. In such environments, where the effects of gravity are minimal or absent, the concept of objects penetrating or intersecting with each other loses its significance. Additionally, penetration-based dynamics may not capture the subtleties of hand-object manipulation in microgravity, such as objects sliding or rotating along the hand surface without separation. These limitations can result in unrealistic simulation results and a reduced sense of immersion. To address them and achieve realistic hand-object interaction in microgravity, we incorporate the impulse along with penetration contact dynamics. The impulse-based method focuses on simulating the exchange of impulses between objects during collisions, rather than relying solely on the concept of penetration and separation. By considering the transfer of momentum and forces between the hand and the object, we can accurately simulate nuanced hand-object interaction in microgravity.

In this work, we follow a two-phase process to simulate hand-object interaction. The first phase is the impulse phase, as described in Sect. 3.4.1, where we utilize an impulse-based method when the real hand contacts the virtual object. This phase focuses on calculating and applying appropriate impulses to model the initial impact and response of the interaction. Once the impulse phase is completed, we transition to the second phase, known as the penetration phase, as outlined in Sect. 3.4.2. In this phase, we incorporate penetration contact dynamics to handle the ongoing contact forces and interactions between the hand and the object. This includes considering factors such as friction, stability, and the prevention of object penetration. This comprehensive approach enables us to achieve a realistic and immersive simulation of hand-object interaction within the unique environment of microgravity.

3.4.1 The Impulse Phase

In our study, we specifically focus on the colliding case and aim to apply a collision impulse to a single point of contact before penetration occurs. This is crucial for enhancing the interaction response in microgravity environments. To achieve this, we utilize Newton's collision law [39] as a foundation for implementing the collision impulse. By incorporating this law, we can accurately calculate and apply the necessary impulse to prevent penetration and simulate realistic collision behavior during hand-object interaction.

First, we assume that the collision impulse is parallel with the $\boldsymbol{n}$-axis. Then, we can obtain:

$$\mathbf{J} = j\boldsymbol{n}, \tag{9}$$

where j is the magnitude of the collision impulse. We will also assume that only the relative contact velocity in the normal direction changes. Newton's impact law defines a simple linear relation between the initial relative contact normal velocity, u_{init}, and the final relative contact normal velocity, u^{final},

$$u^{\text{final}} = -\epsilon \cdot u^{\text{init}}, \tag{10}$$

where ϵ is the normal restitution coefficient with values limited to the interval $(0, 1)$. The restitution coefficient can be thought of as a measure of bounce. So the change in relative contact velocity in the normal direction can be written as:

$$
\begin{aligned}
\Delta \boldsymbol{u} \cdot \boldsymbol{n} &= u^{\text{final}} - u^{\text{init}} \\
&= -\epsilon \cdot u^{\text{init}} - u^{\text{init}} \\
&= -(1 + \epsilon)\boldsymbol{u}^{\text{init}} \cdot \boldsymbol{n},
\end{aligned} \tag{11}
$$

where $\boldsymbol{u}^{\text{init}}$ can be estimated by using Eq. (8). According to the impulse-based momentum relation [35], we have:

$$\Delta \boldsymbol{u} = \mathbf{K}j\boldsymbol{n}, \tag{12}$$

where $\mathbf{K}$ is the collision matrix that is determined by the mass of the objects. We are only interested in the normal direction, so we take the dot product with the contact normal:

$$\Delta \boldsymbol{u} \cdot \boldsymbol{n} = (\mathbf{K}j\boldsymbol{n}) \cdot \boldsymbol{n} = \boldsymbol{n}^{\top}\mathbf{K}\boldsymbol{n}j. \tag{13}$$

Combining Eq. (11) and Eq. (13), we have

$$j = \frac{-(1 + \epsilon)\boldsymbol{u}^{\text{init}} \cdot \boldsymbol{n}}{\boldsymbol{n}^{\top}\mathbf{K}\boldsymbol{n}}. \tag{14}$$

Thus, we can insert Eq. (14) into Eq. (9) to obtain the impulse $\mathbf{J}$.

Second, we apply the impulse $\mathbf{J}$ to the virtual object, which changes the linear velocity v with the amounts:

$$\Delta \boldsymbol{v}_{\text{obj}} = \mathbf{J}/m_{\text{obj}}, \tag{15}$$

where m_{obj} denotes the mass of the virtual object. Thus, we can obtain the motion status of the virtual object.

3.4.2 The Penetration Phase

We utilize both the normal and tangential components of contact force in hand-object interaction in the penetration phase. The contact force for each phalanx in contact with an object is computed and applied to the corresponding contact area. To account for surface friction, we utilize the Coulomb friction model, which calculates tangential friction forces based on the tangential component of contact forces. This model distinguishes between static friction and dynamic friction. Static friction occurs at stable contact points with a higher friction coefficient, while dynamic friction arises when objects are in relative motion, such as sliding along the surface of the hand. The determination of friction type is based on a friction cone criterion, allowing for accurate simulation of different frictional behaviors during hand-object interaction.

As shown in Fig. 6, the i-th contact point on the object's surface is denoted as c_i. We use p_k $(k = 1, 2, \cdots, K)$ to denote the 3D centroid of the k-th hand phalanx. There are K hand phalanxes that are tracked by the Leap Motion device. Then, we can get the contact force f_i^{cont} by:

$$f_i^{\text{cont}} = \gamma \left(c_i - p_k \right), \tag{16}$$

where we set $\gamma = 100$, as same as the previous work [10]. The normal component $f_i^{\text{n-cont}}$ of f_i^{contact} can be calculated as:

$$f_i^{\text{n-cont}} = \left(f_i^{\text{cont}} \cdot n_i \right) n_i, \tag{17}$$

where n_i is the surface normal vector of the object mesh at the contact point c_i. Obviously, the tangential component $f_i^{\text{t-cont}}$ of f_i^{cont} can be represented as:

$$f_i^{\text{t-cont}} = f_i^{\text{cont}} - f_i^{\text{n-cont}}. \tag{18}$$

Next, we discuss how to determine the tangential force $f_i^{\text{T-cont}}$, which depends on whether or not the contact force f_i^{cont} is inside the friction cone. The friction cone is defined by the Coulomb friction model as a cone with the vertex corresponding to the contact point and the axis along the surface normal n_i. When the contact force f_i^{cont} is inside the friction cone, we have:

$$f_i^{\text{T-cont}} = f_i^{\text{t-cont}}, \text{ s.t. } \begin{cases} f_i^{\text{cont}} \cdot n_i > 0 \\ \left\| f_i^{\text{t-cont}} \right\| \leq \mu_i \left(f_i^{\text{cont}} \cdot n_i \right), \end{cases} \tag{19}$$

where μ_i is the static friction coefficient at the surface location of c_i. Otherwise, $f_i^{\text{T-cont}}$ can calculated as:

$$f_i^{\text{T-cont}} = \phi_i f_i^{\text{t-cont}}, \tag{20}$$

where ϕ_i is the dynamic friction coefficient at the surface location of c_i. The coefficients ϕ_i, μ_i are set as same as the previous work [10]. Finally, we apply the forces $f_i^{\text{T-cont}}$, $f_i^{\text{T-cont}}$ and $f_i^{\text{T-cont}}$.

The contact force plays a crucial role in allowing the object to slide along the hand surface or, conversely, enabling a firm grasp by counteracting gravity. The magnitude and direction of the contact force determine these effects while considering the friction properties of the surface. By appropriately adjusting the contact force, we can control the sliding or grasping behavior of the object, ensuring a realistic and interactive hand-object interaction experience.

4 EXPERIMENTS

We present the implementation details of our system in a simulated microgravity environment. We showcase the quantitative results obtained from experiments and measurements, evaluating the effectiveness of our approach compared with state-of-the-art approaches. Finally, we present the system reliability and fidelity test.
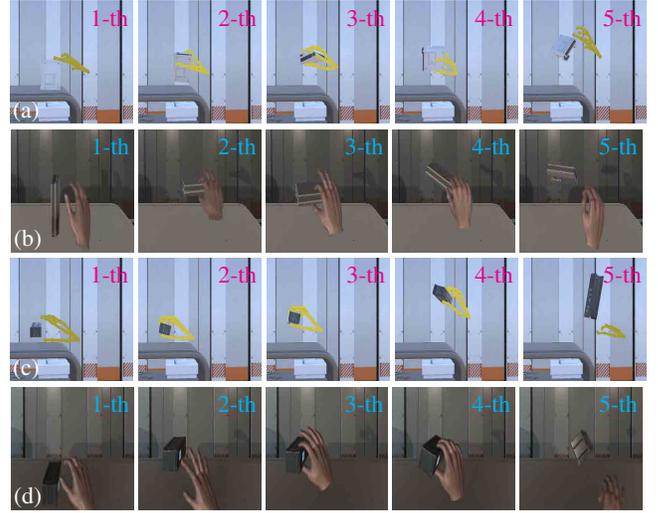

Figure 7: Comparative visualizations of our method and CLAP [20]. (a) Ours: flat object grab, (b) CLAP: flat object grab, (c) Ours: slender object grab, and (d) CLAP: slender object grab.

4.1 Implementation Details

We implemented our system on a laptop, which is equipped with an 11th Gen Intel(R) Core(TM) i7-11800H processor, operating at a frequency of 2.30 GHz. It had a total memory capacity of 16.0 GB, with 15.8 GB available for use. The system is built by Unity with a Leap Motion device. These hardware and software components provide the necessary computational power, graphics rendering capabilities, and MR interaction support to simulate and visualize the hand-object interaction in our system effectively.

4.2 Comparisons with the State-of-the-Art

We evaluate the performance of our system by comparing it to an open-source state-of-the-art CLAP system [20] in grasping two different cuboid-shaped objects, as illustrated in Fig. 7. The figure presents five frames from each method, depicting various stages of the task execution from left to right. This analysis enables us to evaluate the efficacy and robustness of our system in comparison to CLAP. For a comprehensive view of the entire procedure, please refer to the provided supplementary video.

When assessing the capability of our system in grasping a slender object (Fig. 7(c)) against that of the flat object (Fig. 7(a)), it's evident that the slender object presents more challenges. The extended form and dimensions of this elongated cuboid necessitate heightened precision in hand movements and coordination for effective grip and manipulation. Furthermore, its slim profile augments the potential for collisions with other virtual entities or surfaces, emphasizing the need for meticulous navigation and maneuvers. Contrasting with a prior study [20] illustrated in Figs. 7(b) and 7(d), our approach showcases enhanced efficacy, adeptly handling both objects. This underscores our system's proficiency in navigating the intricacies linked to manipulating elongated cuboid items in simulated microgravity settings.

Furthermore, our system accommodates a range of hand-object interaction techniques, such as ball grasping, cube grasping, ball catching, pinching (securing a slender object using limited fingers), lifting (gripping an object and elevating it), and gently nudging objects, as depicted in Fig. 8. Our conducted tests underscore the user-friendly nature of our approach, with participants adeptly completing all tasks, affirming the method's efficacy. Such a diverse array of interaction techniques not only equips users with hands-on experience but also bolsters proficiency in manipulating tasks within

Figure 8: The visualization of diverse hand-object interaction demos. Our system supports a wide range of hand-object interaction actions, accommodating various object shapes and manipulation tasks.
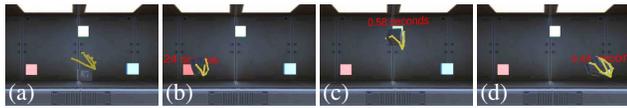

Figure 9: The visualization of different task procedures: (a) the initial scene, (b) the first place, (c) the second place, and (d) the third place.

Figure 10: The visualization of the test procedure timeline.

simulated microgravity contexts. This prepares them more effectively for tangible space expeditions.

4.3 Reliability and Fidelity Test

To address the constraints inherent to authentic microgravity environments and the limited availability of astronauts, our work primarily concentrates on evaluating the system's reliability and fidelity under standard gravitational conditions. This evaluation lays the groundwork for training future astronauts for space missions. Our test unfolds in two phases. Firstly, we argue that a dependable training system should consistently guide beginners in refining their skills. By comparing the performance of participants trained with our system to those who were not, we can validate our system's reliability. Secondly, we contrast user feedback from our system with a state-of-the-art method to gauge our system's simulation fidelity.

Apparatus and Participants   All participants utilize the same apparatus during the whole test, comprising a laptop equipped with Unity and Leap Motion. Further details can be found in Sect. 4.1.

We have recruited 30 participants, including 20 males and 10 females aged between 18-30 (avg=23.63, std=3.44), with a diverse age distribution. This selection ensures representation from a demographic familiar with technology and adaptable to new interfaces. Among them, 3 participants are left-handed and the remaining are right-handed. Regarding previous experience, 8 participants have no experience with MR. This diverse group of participants is selected to provide a comprehensive evaluation of our system. To evaluate the training reliability of our system, we divided participants equally into two unbiased groups: the veteran (V) group and the beginner (B) group, ensuring gender balance within each group.

Tasks   During training, participants are required to do simple manipulation operations to be familiar with the experimental con-

ditions. During testing, participants are required to perform three pick-and-place tasks on differently shaped objects. These objects include a cube (C) shown in Fig. 8(f), a flat cuboid (F) shown in Fig. 7(a), and a long cuboid (L) shown in Fig. 7(c). The order of objects and their corresponding test order is determined through a pseudo-random process to minimize transfer and learning biases.

To enhance external validity, each pick-and-place task consists of three stages with diverse difficulty levels: manipulation, maintenance, and release. (1) In the manipulation stage, participants pick up the target object from a specified position, as can be seen in Fig. 9(a). (2) During the maintenance stage, they hold the target object from first place to third place for a minimum duration of three seconds, as Figs. 9(b) to 9(d). When the target object is reached the corresponding place, the mark is highlighted and the timer is triggered. (3) Finally, participants release or drop the object.

Evaluation Metrics   To measure task completion performance and user feedback, we use both quantitative and qualitative metrics.

Quantitatively, we measure the time taken to perform the task and the number of attempts. Specifically, the dependent variables are (1) task completion time and (2) the number of attempts (instances where the object was dropped during manipulation).

Qualitatively, participants evaluate the system's naturalness and efficiency using an 11-point Likert scale. (1) The naturalness rating gauges the extent to which the interaction mirrors real-world hand-object manipulations, asking: "How natural did the sensation of the average task completion experience feel? (0: Extremely unnatural, 10: Extremely natural)". (2) The efficiency rating measures the fluency and effectiveness of task performance, posing the question: "How would you rate the efficiency of the average task completion experience? (0: Extremely inefficient, 10: Extremely efficient)". These Likert-scale questions are adapted from [23].

Procedure   The test procedure initiates with an orientation to the objectives. Subsequently, participants are given comprehensive instructions about the system and the manipulation tasks. During
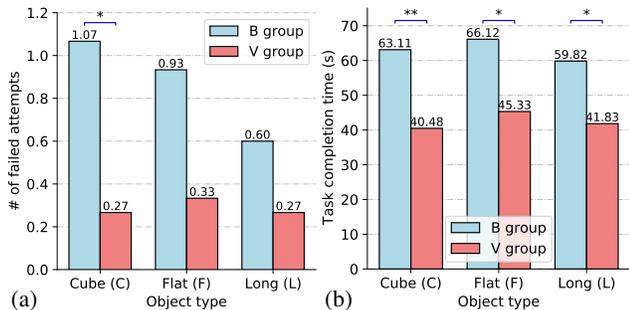
Figure 11: Bar plots of different objects and groups with respect to (a) the number of attempts and (b) the task completion time (s).

Figure 12: The rating distribution plot of our system and CLAP [20].

Table 1: Comparative results between our system and CLAP [20].

Ours	Time (s)	# attempts	CLAP	Time (s)	# attempts
	51.7 ± 19.4	0.9 ± 0.3		101.6 ± 58.0	2.2 ± 2.9

the training phase, the V group engages in a training session with our system, while the B group directly advances to the first stage. In this initial stage, every participant leverages our system to tackle three distinct tasks for each respective object. Key metrics, including task completion time and the number of attempts, are documented to appraise the system's efficacy. For the second stage, participants interact with the state-of-the-art system for the purpose of hand-object manipulation, serving as a benchmark. Next, participants are required to fill out a form regarding the naturalness and efficiency of the pick-and-place experience, rating them on an 11-point Likert scale. The procedure is encompassed within a 40-minute bracket, detailed further in Fig. 10.

Effectiveness of System Reliability Fig. 11 shows the bar plots, illustrating the impact of varying object types and participant groups on the number of attempts, as well as task completion times (s). An ANOVA has been carried out for every independent variable. The significance levels are represented using stars: '*' for $p < 0.05$, '**' for $p < 0.01$, and '***' for $p < 0.001$. Significantly, none of the variables violated the assumption of sphericity. Given the multiple comparisons made, p-values from these comparisons underwent Bonferroni correction. This procedure strengthens the validity of our results, especially when assessing performance disparities across distinct variables. In Fig. 11(b), the post hoc analysis indicates a significant effect of the participant group on task completion time (with $p < 0.05$). Notably, the B group's completion time surpasses that of the V group. Even though the participant group does not exhibit a marked influence on flat and long objects as seen in Fig. 11(a), the B group's attempts are approximately double those of the V group. These observations highlight the reliability of our system in assisting users with various object manipulations, thereby showcasing the large potential for the effectiveness of the training performance in microgravity. Moreover, Fig. 11 reveals that users often face heightened challenges when endeavoring to grasp flat or elongated objects. This is attributable to their minimized grip contact area and difficulties in force distribution.

Effectiveness of the System Fidelity In this test, we continue to use the state-of-the-art CLAP system [20] as our baseline, given its robust performance. We comprehensively compare our system with CLAP and the results are shown in Tab. 1 and Fig. 12. The data from Tab. 1 suggests that our system is more operationally efficient, characterized by reduced task completion time and attempts compared to CLAP. Moreover, the ratings presented in Fig. 12 depict a distinct disparity between the two systems in terms of naturalness and efficiency. This difference underscores our system's heightened ability to offer users a more authentic sensation of hand-object interaction compared to CLAP. These results underscore that our system outperforms CLAP in delivering an enhanced user experience. This further verifies our system's fidelity for simulating hand-object interaction and shows great potential in space mission training.
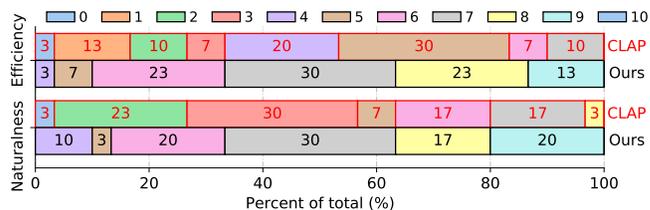
5 CONCLUSIONS AND DISCUSSIONS

Our innovative MR-based system, incorporating impulse-based simulation with penetration contact dynamics, effectively captures the complex dynamics essential for a realistic simulation of hand-object interactions in simulated microgravity environments. It not only presents a cost-effective way for the general public to experience space object manipulation but also establishes a crucial training tool for space travelers. Experimental results highlight the advantages of our system over the state-of-the-art, showcasing the potential in enhancing space mission preparedness. Despite these contributions, several areas still require further exploration:

Reliance Solely on Visual Feedback The current rendition of our system chiefly provides an immersive sensation via visual feedback. Although this form of feedback has proven effective in various scenarios, the intricacies of microgravity interactions emphasize the importance of tactile feedback. A lack of haptic feedback might impede the attainment of a truly immersive experience. To address this limitation, future work should consider integrating cutting-edge haptic rendering technologies, like the high-precision tactile gloves referenced in the previous work [40], ensuring a more genuine sensation of weightlessness.

Demographic Limitations in Testing The effectiveness of our system, especially concerning space traveler training for specialized space missions, remains speculative given the scant astronaut engagement in our work. Though the system is attuned to cater to a wider audience for experiencing object manipulation in microgravity, there's a potential oversight of the specific challenges astronauts grapple with in authentic microgravity conditions. To enhance the system's fidelity and relevance, we aim to collaborate more intensively with astronauts in future work, leveraging their firsthand insights to fine-tune the system.

Real-Virtual Hand Incongruence In our system, the user's hands are represented using an MR skeletal model, which can potentially lead to visual inconsistencies when juxtaposed with the users' actual hands. Such inconsistencies might diminish the overall immersion by causing perceptual discord. It's crucial to address these discrepancies to ensure a truly immersive experience. Integrating high-fidelity rendering with tactile feedback can be instrumental in resolving these issues. Additionally, the infusion of cognitive phenomena, exemplified by the "rubber hand illusion", and its associated synchronous stimuli could amplify the user's perceived agency over the virtual avatar.

REFERENCES

[1] K. Zhou, Z. Cheng, H. P. Shum, F. W. Li, and X. Liang, "Stgae: Spatial-temporal graph auto-encoder for hand motion denoising," in *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 41–49, IEEE, 2021.

[2] T. L. Lonner and T. K. Clark, "The efficacy of vr as a countermeasure for astronaut motion sickness in post-flight water landings," in *2023 IEEE Aerospace Conference*, pp. 1–8, IEEE, 2023.

[3] T. Rybus and K. Seweryn, "Planar air-bearing microgravity simulators: Review of applications, existing solutions and design parameters," *Acta Astronautica*, vol. 120, pp. 239–259, 2016.

[4] G. Albrecht-Buehler, "The simulation of microgravity conditions on the ground.," *ASGSB bulletin: publication of the American Society for Gravitational and Space Biology*, vol. 5, no. 2, pp. 3–10, 1992.

[5] J. Regnard, M. Heer, C. Drummer, and P. Norsk, "Validity of microgravity simulation models on earth," *American journal of kidney diseases*, vol. 38, no. 3, pp. 668–674, 2001.

[6] G. Vailland, L. Devigne, F. Pasteau, F. Nouviale, B. Fraudet, E. Leblong, M. Babel, and V. Gouranton, "Vr based power wheelchair simulator: Usability evaluation through a clinically validated task with regular users," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 420–427, IEEE, 2021.

[7] R. Herranz, R. Anken, J. Boonstra, M. Braun, P. C. Christianen, M. de Geest, J. Hauslage, R. Hilbig, R. J. Hill, M. Lebert, *et al.*, "Ground-based facilities for simulation of microgravity: organism-specific recommendations for their use, and recommended terminology," *Astrobiology*, vol. 13, no. 1, pp. 1–17, 2013.

[8] X. Zhou, Q. Wan, W. Zhang, X. Xue, and Y. Wei, "Model-based deep hand pose estimation," *arXiv preprint arXiv:1606.06854*, 2016.

[9] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013.

[10] M. Höll, M. Oberweger, C. Arth, and V. Lepetit, "Efficient physics-based implementation for realistic hand-object interaction in virtual reality," in *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 175–182, IEEE, 2018.

[11] C. B. Zilles and J. K. Salisbury, "A constraint-based god-object method for haptic display," in *Proceedings 1995 ieee/rsj international conference on intelligent robots and systems. Human robot interaction and cooperative robots*, vol. 3, pp. 146–151, IEEE, 1995.

[12] Z. Fan, O. Taheri, D. Tzionas, M. Kocabas, M. Kaufmann, M. J. Black, and O. Hilliges, "Arctic: A dataset for dexterous bimanual hand-object manipulation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12943–12954, 2023.

[13] J. Zheng, Q. Zheng, L. Fang, Y. Liu, and L. Yi, "Cams: Canonicalized manipulation spaces for category-level functional hand-object manipulation synthesis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 585–594, 2023.

[14] Y. Zhang, J. T. Richards, J. L. Hellein, C. M. Johnson, J. Woodall, T. Sorenson, S. Neelam, A. M. J. Ruby, and H. G. Levine, "Nasa's ground-based microgravity simulation facility," *Plant Gravitropism: Methods and Protocols*, pp. 281–299, 2022.

[15] V. Pletser, "European aircraft parabolic flights for microgravity research, applications and exploration: A review," *Reach*, vol. 1, pp. 11–19, 2016.

[16] A. Aissiou, S. Jha, K. Dhunnoo, Z. Ma, D. Li, R. Ravin, M. Kunze, K. Wong, and A. Adesida, "Transcriptomic response of bioengineered human cartilage to parabolic flight microgravity is sex-dependent," *npj Microgravity*, vol. 9, no. 1, p. 5, 2023.

[17] C. Sun, S. Chen, J. Yuan, and Z. Zhu, "A six-dof buoyancy tank microgravity test bed with active drag compensation," *Microgravity Science and Technology*, vol. 29, no. 5, pp. 391–402, 2017.

[18] P. Yang, X. Wang, Y. Liu, C. Yan, and X. Wang, "Numerical simulation of vapor configuration and bubbles coalescence in cryogenic propellant tank under microgravity," *Microgravity Science and Technology*, vol. 35, no. 2, p. 20, 2023.

[19] Q. Feng, H. P. H. Shum, and S. Morishima, "Resolving hand-object occlusion for mixed reality with joint deep learning and model optimization," *Computer Animation and Virtual Worlds*, vol. 31, no. 4–5, p. e1956, 2020.

[20] M. Verschoor, D. Lobo, and M. A. Otaduy, "Soft hand simulation for smooth and robust natural interaction," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 183–190, IEEE, 2018.

[21] J.-S. Kim, M. Jeon, and J.-M. Park, "Multi-hand direct manipulation of complex constrained virtual objects," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3235–3240, IEEE, 2019.

[22] W. Quan, H. Yang, C. Han, and Y. Li, "Realistic interaction system for human hand in virtual environments," in *2020 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, pp. 772–778, IEEE, 2020.

[23] T. Delrieu, V. Weistroffer, and J. P. Gazeau, "Precise and realistic grasping and manipulation in virtual reality without force feedback," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 266–274, IEEE, 2020.

[24] D. Han, R. Lee, K. Kim, and H. Kang, "Vr-handnet: A visually and physically plausible hand manipulation system in virtual reality," *IEEE Transactions on Visualization and Computer Graphics*, 2023.

[25] S. Oprea, P. Martinez-Gonzalez, A. Garcia-Garcia, J. A. Castro-Vargas, S. Orts-Escolano, and J. Garcia-Rodriguez, "A visually realistic grasping system for object manipulation and interaction in virtual reality environments," *Computers & Graphics*, vol. 83, pp. 77–86, 2019.

[26] A. D. Blaga, M. Frutos-Pascual, C. Creed, and I. Williams, "Freehand grasping: An analysis of grasping for docking tasks in virtual reality," in *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 749–758, IEEE, 2021.

[27] H. Hu, X. Yi, H. Zhang, J.-H. Yong, and F. Xu, "Physical interaction: Reconstructing hand-object interactions with physics," in *SIGGRAPH Asia 2022 Conference Papers*, pp. 1–9, 2022.

[28] H. Zhang, Y. Ye, T. Shiratori, and T. Komura, "Manipnet: neural manipulation synthesis with a hand-object spatial representation," *ACM Transactions on Graphics (ToG)*, vol. 40, no. 4, pp. 1–14, 2021.

[29] H. Yu, C. Cheang, Y. Fu, and X. Xue, "Hando: a hybrid 3d hand–object reconstruction model for unknown objects," *Multimedia Systems*, pp. 1–15, 2022.

[30] P. Streli, R. Armani, Y. F. Cheng, and C. Holz, "Hoov: Hand out-of-view tracking for proprioceptive interaction using inertial sensing," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2023.

[31] L. Yang, X. Zhan, K. Li, W. Xu, J. Li, and C. Lu, "Cpf: Learning a contact potential field to model the hand-object interaction," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 11097–11106, 2021.

[32] S. Christen, M. Kocabas, E. Aksan, J. Hwangbo, J. Song, and O. Hilliges, "D-grasp: Physically plausible dynamic grasp synthesis for hand-object interactions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20577–20586, 2022.

[33] Z. Chang, G. A. Koulieris, and H. P. Shum, "On the design fundamentals of diffusion models: A survey," *arXiv preprint arXiv:2306.04542*, 2023.

[34] K. Zhou, R. Cai, Y. Ma, Q. Tan, X. Wang, J. Li, H. P. Shum, F. W. Li, S. Jin, and X. Liang, "A video-based augmented reality system for human-in-the-loop muscle strength assessment of juvenile dermatomyositis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 29, no. 5, pp. 2456–2466, 2023.

[35] K. Erleben, J. Sporring, K. Henriksen, and H. Dohlmann, *Physics-based animation*, vol. 79. Charles River Media Hingham, 2005.

[36] M. Ortega, S. Redon, and S. Coquillart, "A six degree-of-freedom god-object method for haptic display of rigid bodies with surface properties," *IEEE transactions on visualization and computer graphics*, vol. 13, no. 3, pp. 458–469, 2007.

[37] J. Jacobs, M. Stengel, and B. Froehlich, "A generalized god-object method for plausible finger-based interactions in virtual environments," in *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 43–51, IEEE, 2012.

[38] Y. Shi, L. Zhao, X. Lu, T. Hoang, and M. Wang, "Grasping 3d objects with virtual hand in vr environment," in *The 18th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, pp. 1–8, 2022.

[39] E. Guendelman, R. Bridson, and R. Fedkiw, "Nonconvex rigid bodies with stacking," *ACM transactions on graphics (TOG)*, vol. 22, no. 3, pp. 871–878, 2003.

[40] M. Aiple and A. Schiele, "Pushing the limits of the cybergrasp™ for haptic rendering," in *2013 IEEE international conference on robotics and automation*, pp. 3541–3546, IEEE, 2013.