PHI-NOM: Physics-Based Hand Interactions for Natural Object Manipulation in Virtual Reality

Phi Nguyen
NE-XC
NASA
Kennedy Space Center
phi.h.nguyen@nasa.gov

Wessley Dennis
NE-XC
NASA
Kennedy Space Center
wessley.a.dennis@nasa.gov

Antonio Pego
NE-XC
NASA
Kennedy Space Center
antonio.pego-1@nasa.gov

***Abstract*—Training for space missions or surgical procedures is inherently expensive, logistically complex, and resource-intensive, making immersive virtual reality (VR) an appealing alternative for rehearsing mission-critical tasks. However, current VR systems often rely on handheld controllers or gesture recognition, failing to simulate the physics necessary for equipment handling and assembly tasks. This paper introduces PHI-NOM (Physics-Based Hand Interactions for Natural Object Manipulation). This controller-free VR framework provides biomechanically accurate, real-time hand interactions by modeling fingertip contact and grasp physics. Built using motion capture and Unreal Engine, PHI-NOM employs a new algorithm for physics-based object manipulation, enabling natural actions such as grasping, lifting, rotating, and passing objects without the need for controllers, predefined gestures, or intermediary devices. Unlike existing methods that rely on computationally intensive inverse kinematics or physics solvers, PHI-NOM achieves low computational overhead through a lightweight model that provides visually accurate real-world physics. This design enables real-time interaction, supports diverse object geometries, allows for two-handed manipulation, and facilitates both individual and collaborative training scenarios, including cooperative tool use and system assembly. By enhancing tactile realism, PHI-NOM promotes motor learning and improves skill transfer from virtual to real-world environments.**

***Index Terms*—Virtual Reality, Simulation Training, Hand-Object Manipulation**

I. INTRODUCTION

Over the past decade, virtual reality (VR) has transformed from a niche technology into a versatile tool for immersive training across many industries. From aerospace and healthcare to manufacturing and education, VR simulations offer cost-effective, scalable, and risk-free environments where users can practice complex tasks without the consequences of real-world errors. VR has proven its value in enhancing skill acquisition and knowledge retention by allowing trainees to rehearse procedures, explore dynamic systems, and make critical decisions in controlled settings. However, the effectiveness of VR training is closely tied to the realism of user interactions within the simulated environment.

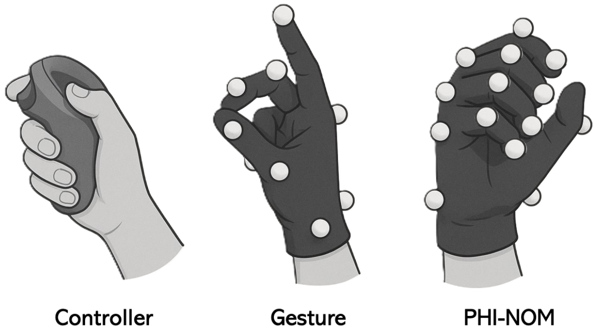

Fig. 1. Methods for grabbing a lever in virtual reality.

While visual and auditory fidelity in VR has advanced considerably, natural hand-object interaction remains a significant challenge. The current state of VR relies on handheld controllers or gesture recognition for world interaction(Fig. 1), which fail to replicate the dynamics of real-world hand movements. These methods restrict users to simplified input methods, breaking the intuitive connection with how the user would interact with the world in real life. The limitations of these interaction methods are especially problematic in training scenarios that depend on fine motor skills and muscle memory. For example, a maintenance technician practicing complex equipment repairs in VR may struggle to build proper hand-eye coordination if forced to use a controller to grab and manipulate tools. Similarly, a medical student attempting a delicate surgical procedure may fail to develop the precise, coordinated hand movements required for success because their virtual hand interactions lack the tactile and proprioceptive feedback of real-world manipulation. Moreover, without realistic hand-object interactions, students miss opportunities to practice essential teamwork skills, such as passing instruments to peers during procedures, an integral aspect of surgical collaboration. Thus, there is a critical need for VR systems that offer more realistic, physics-based hand-object interactions. Such systems must enable users to grasp, lift, pass, and manipulate objects using their hands naturally, just as they would in the real world, without the limitations of controllers or predefined gestures. This paper introduces PHI-NOM (Physics-Based Hand Interactions for Natural Object Manipulation), a controller-free VR framework designed to overcome these limitations. PHI-NOM leverages advanced motion capture and physics modeling to simulate real-time, biomechanically accurate hand interactions, enabling users to perform complex tasks with high fidelity and skill transferability. In this paper, we present the architecture and technical foundation of PHI-NOM and an analysis of its performance in comparison to controllers.

II. RELATED WORK

Realistic hand-object interaction in virtual reality (VR) has been a critical area of research, with multiple approaches developed over the years. These techniques can be categorized into four main categories: Gesture-Based Interaction, Animation-Driven Techniques, Physical Proxy Systems, and Physics-Based Grasping Frameworks.

A. Gesture-Based Interaction

Gesture-based interaction relies on symbolic gestures, such as pinch or grip, to manipulate virtual objects [2]. These techniques are intuitive because they map user actions to predefined gestures. Typically, systems like Leap Motion [14] or Microsoft Kinect [15] use skeletal tracking to identify specific hand poses and trigger interactions. Critically, because these gestures are symbolic rather than physically grounded, they do not foster muscle memory development. A user who repeatedly performs a pinch gesture in VR does not develop the fine motor skills needed for precise object manipulation in the real world.

B. Animation-Driven Techniques

Animation-driven methods use predefined animations or grasp taxonomies to simulate hand-object interactions [3]. These systems rely on predefined animations for grasping, such as open and closed hand animations, which are triggered based on user input. While these approaches ensure visual consistency, they lack real-time responsiveness because the user's actual hand movements do not directly influence the animation. This results in a perceptual disconnect where the user sees a smooth animation but cannot make real-time adjustments. The fixed nature of these animations also prevents the development of muscle memory, as users are not actively controlling their finger orientations or adapting their grip in response to object shape.

C. Physical Proxy Systems

Physical proxy systems use real objects or devices that users can touch to simulate virtual interactions. These can be as simple as 3D-printed shapes that match the size and shape of virtual objects, or as advanced as gloves with built-in sensors that provide a sense of touch. For example, if a user is supposed to grab a virtual wrench, they might instead hold a real, 3D-printed wrench that feels similar. This makes interactions feel more realistic because users can touch something. However, these systems have limitations. Each proxy is designed for a specific task, so they lack flexibility. A 3D-printed wrench cannot suddenly become a screwdriver. Additionally, these physical proxies might not always match what the user sees in VR, causing confusion. For instance, the user might be holding a real object but seeing a different virtual object, which can break the sense of immersion and prevent them from developing accurate muscle memory.

D. Physics-Based Grasping Frameworks

Physics-based frameworks simulate how objects behave and respond to user interactions [6,7] by calculating physical properties like weight, friction, and force. These systems typically rely on three core components: collision detection, which determines when the user's virtual hand touches an object; rigid body dynamics, which simulate how the object moves or reacts based on the user's input; and inverse kinematics, a method where the system calculates how the user's fingers should bend or position themselves to match a pre-programmed grasping pose. For example, when a user reaches for a virtual wrench, the system detects the collision between the user's hand and the wrench model. It then calculates the contact points—specific locations on the wrench where the fingers should touch—and uses inverse kinematics to adjust the virtual fingers, making them appear to grasp the wrench. The system may even simulate force feedback, making the wrench resist movement if the user tries to push it. However, the reliance on inverse kinematics introduces a critical limitation. Instead of allowing the user's fingers to naturally conform to the object's shape, the system forces the fingers into predefined poses. This means users are not directly controlling their virtual fingers, but rather their movements are filtered through a complex solver that determines the final pose. As a result, the user has less direct control over fine motor skills, which are essential for building muscle memory. Fine motor skills are best developed through continuous, user-driven adjustments, not automatic pose corrections."

E. Comparison with PHI-NOM

PHI-NOM builds on the foundation of physics-based approaches, sharing similarities such as detecting collisions between the user's hand and virtual objects while avoiding gesture detection or grasp taxonomies. However, while traditional physics-based systems rely on colliders and inverse kinematics, PHI-NOM uses ray-traced intersections between fingertips and object surfaces. This allows for the calculation of contact points and surface normals, directly mapping these detections to user inputs. This design enables PHI-NOM to maintain low-latency, real-time performance without the overhead of complex solvers. By directly mapping the user's finger positions and maintaining consistent spatial relationships, PHI-NOM promotes the development of muscle memory, allowing users to build real-world fine motor skills that transfer beyond the virtual environment.

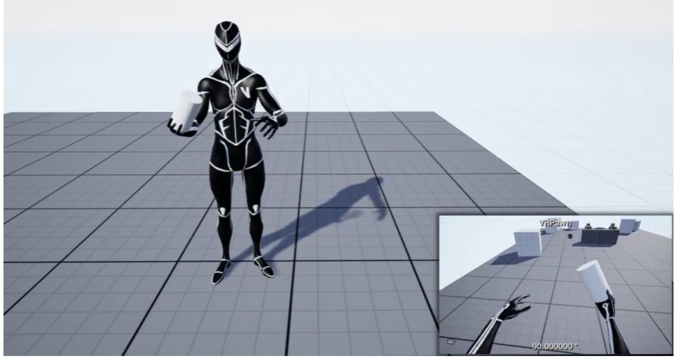

Fig. 2. Third and first person perspective of grabbing with PHI-COM.

III. USE

This section provides a detailed overview of how PHI-NOM is set up and used within a VR environment, including the hardware, software architecture, and data flow that enable precise, controller-free hand interactions. PHI-NOM is a modular software framework designed for integration within Unreal Engine environments (Fig. 2), enabling controller-free object manipulation in virtual reality (VR) applications. This paper includes the use of the Vicon motion capture system in conjunction with VR head-mounted displays (HMDs); however, the framework is hardware-agnostic. Any motion capture system capable of accurately tracking the position and orientation of the user's hands and fingers can be used to drive the interaction. The setup used for this paper involves users wearing lightweight gloves equipped with motion capture markers (Fig. 1) compatible with the Vicon system. These markers allow the system to track hand and finger movements with high precision, which are then streamed to a server running the Unreal Engine session. In Unreal Engine, the software references the phalange bones from the streamed data to retrieve their positions and rotations, which are then used to calculate graspability. The Unreal Engine environment, hosted on a high-performance workstation, receives the incoming motion capture data and uses it to drive a virtual hand avatar. This avatar mirrors the user's real-time movements, which PHI-NOM uses to track finger location and rotation in the virtual world. Motion capturing acts as an intermediary by mapping the user's body from the real world to virtual reality. Notably, PHI-NOM is not a plug-and-play solution, but rather a highly flexible framework that can be adapted to a range of VR training applications. While the underlying architecture remains consistent, minor customization may be necessary to suit specific scenarios.

IV. GRASPING SYSTEM

The PHI-NOM grasping system is a physics-based framework designed to achieve natural, realistic hand-object interactions in virtual reality (VR). Unlike conventional methods that rely on gesture recognition or pre-defined animations, PHI-NOM uses a continuous, physics-driven approach that accurately captures fingertip contact, grasp validation, and object manipulation. PHI-NOM employs a ray-casting method for real-time geometric contact detection. Each fingertip is represented as a ray origin, and the system continuously calculates the intersection points between these rays and the surfaces of virtual objects. When a fingertip ray intersects an object's surface, the contact point and surface normal are computed and stored. This direct geometric calculation ensures that every interaction is physically grounded, bypassing the limitations of gesture libraries or animation-based interactions. Grasp validity is determined through a dynamic, angle-based criterion. For a grasp to be recognized, at least the thumb and one opposing finger must make simultaneous contact with the object, and the angle between their surface normals must exceed a specified threshold, typically 120 °, which was determined experimentally. Row 4 of Fig. 3 demonstrates how this method allows users to dynamically hold objects with different fingers as seen in real life. This angular requirement mirrors the natural opposition between fingers during real-world grasping, preventing unintentional grasps—a common issue seen in gesture-based systems—and ensuring secure object handling. Once a grasp is established, PHI-NOM supports continuous physics-based manipulation. Rows 1 through 3 of Fig. 3 show how users can naturally lift, rotate, pass, or reorient objects with their hands. This approach supports both single-handed and two-handed operations, including tasks such as passing an object between hands to reorient or reposition the object. To enhance stability, PHI-NOM implements a temporal validation window. This means a grasp is only considered active if the contact is maintained continuously over a defined period, preventing accidental interactions from momentary touch. Furthermore, PHI-NOM is designed to handle complex object geometries as seen in Row 1 of Fig. 3. Whether interacting with simple shapes or high-polygon models, the system is designed to be robust and dynamically detects and responds to diverse object contours. This scalability allows users to manipulate objects of varying complexity, from simple tools to intricate mechanical components.

For a technician practicing equipment assembly, PHI-NOM would allow them to naturally reach out, grip, and manipulate tools in various ways to discover the most comfortable and secure grip. If they lose hold of an object due to a poor grip, they can adjust their technique and build muscle memory that directly translates to real-world performance.

A. Algorithm

In PHI-NOM, hand-object interaction is modeled through real-time geometric and physical analysis of fingertip contact. Each fingertip is represented as a ray-casting origin. When a ray intersects with an object's mesh, the corresponding contact point and surface normal are computed to determine grasp feasibility.

Let $\vec{p}_i \in \mathbb{R}^3$ denote the position of the i^{th} fingertip, and $\vec{q}_i \in \mathbb{R}^3$ the closest intersection point on the object's mesh. A contact is considered valid if the Euclidean distance between the fingertip and the surface point satisfies:

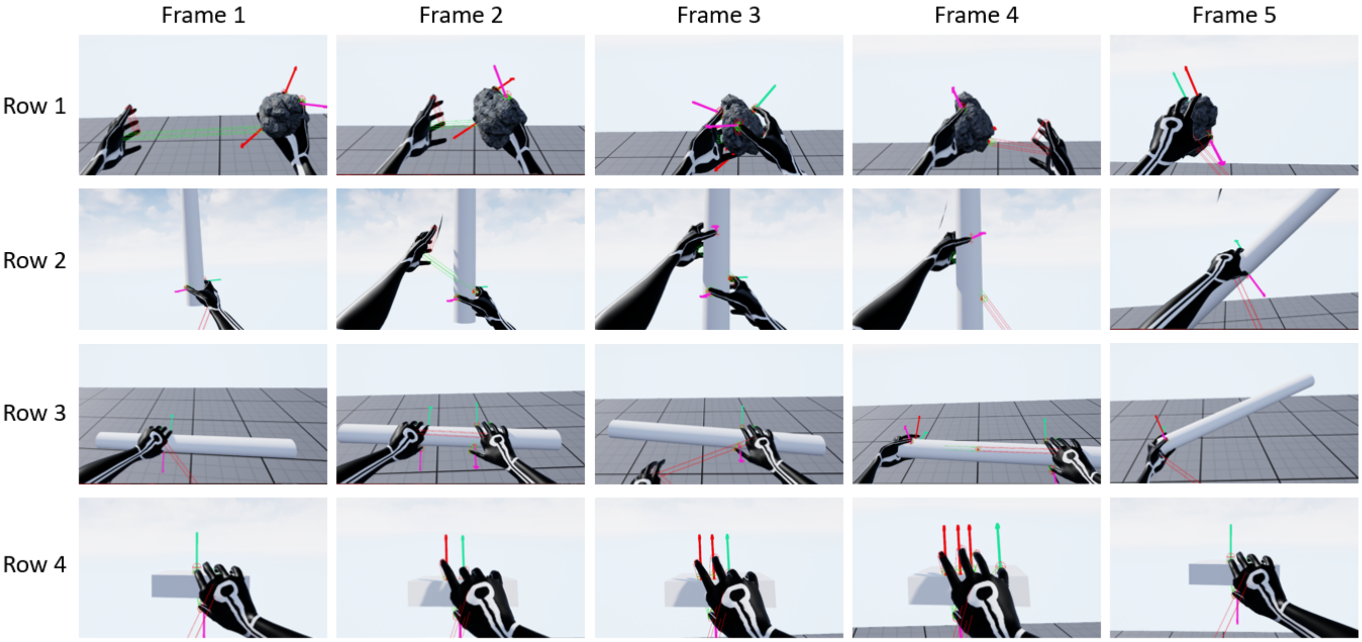

Fig. 3. Examples demonstrating PHI-NOM's capabilities in real-time grasping and manipulation. The arrows represent surface normal vectors at the fingertips, which determine grasp validity. Pink arrows indicate the surface normals of the thumb, while red arrows represent normals from fingers that do not meet the grasp criteria due to incorrect distance or angle. In contrast, green arrows show normals from fingers that satisfy both the distance and angle criteria, making them eligible for grasping. Row 1 illustrates the process of grabbing and passing a high-polygon object (a rock). Row 2 demonstrates a simulated climbing action on a long cylindrical object, where users alternate hand grips to move upward. Row 3 highlights the versatility of gripping the same cylinder, allowing users to hold it from the side with one hand, two hands, or even by grasping its circular end. Finally, Row 4 showcases fine control, where users can adjust their finger positions while maintaining a secure grip, provided that the thumb and at least one other finger consistently meet the grasp criteria.

$$\|\vec{p}_i - \vec{q}_i\| < \epsilon_d, \tag{1}$$

where ϵ_d is a tunable grasp threshold accounting for tracking jitter and model alignment.

To assess whether opposing fingers are aligned in a physically plausible manner, the angle between their contact surface normals $\vec{n}_1$ and $\vec{n}_2$ is evaluated—Fig. 3 illustrates these vectors:

$$\theta = \cos^{-1}\left(\frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n}_1\|\,\|\vec{n}_2\|}\right) = \cos^{-1}\left(\vec{n}_1 \cdot \vec{n}_2\right). \tag{2}$$

A valid grasp requires that this angle exceeds a minimum threshold θ_{min}, typically around $120°$:

$$\theta > \theta_{\text{min}}. \tag{3}$$

Finally, to ensure the grasp is intentional and stable, not the result of momentary or accidental contact, a temporal validation window Δt is enforced. A grasp is only considered active if both spatial and angular criteria are satisfied continuously over the time interval:

$$\text{Grasp}_{\text{valid}}(t) = \begin{cases} 1, & \begin{array}{l}\text{if } \forall \tau \in [t - \Delta t, t]: \\ \|\vec{p}_i - \vec{q}_i\| < \epsilon_d \text{ and } \theta > \theta_{\text{min}}\end{array} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

This formulation avoids snapping artifacts, unnatural hand locking, and premature object attachment—issues common in systems using automatic pose fitting. This method also preserves full articulation of each finger, allowing for smooth, anatomically accurate hand motion throughout the grasping sequence.

Notably, the geometric approach allows the system to support arbitrarily complex object geometries, including asymmetric or high-polygon meshes, and allows grasping from multiple orientations. It scales gracefully to both single- and two-handed manipulation, including collaborative tasks in multi-user VR scenarios. Throughout all operations, the user retains complete control over the virtual hand, free from gesture classification, animation blending, or rigid inverse kinematics, enabling deeply immersive and physically plausible interactions.

V. PERFORMANCE ANALYSIS

A. *Participants*

To evaluate PHI-NOM's effectiveness and realism, we conducted a user study involving 12 participants with varying levels of familiarity with VR. Each participant was asked to perform a series of four assembly tasks first using (1) a controller-based system where objects were selected via trigger press, (2) our physics-based PHI-NOM framework that allowed direct, natural interaction without predefined gestures or input devices.

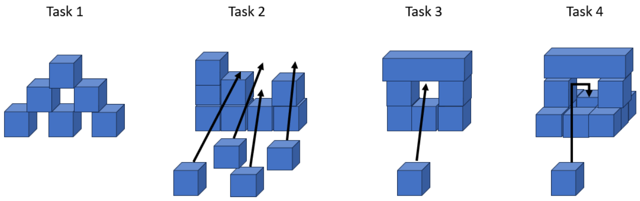

Fig. 4. Assembling tasks for user study.

B. Procedure

Each participant completed a series of four structured assembly tasks designed to evaluate their ability to interact with virtual objects using two different VR interaction methods: a traditional controller-based system and the PHI-NOM framework. As illustrated in Fig. 4, these tasks varied in complexity and spatial precision, providing a comprehensive assessment of user performance:

- Task 1: Participants were instructed to assemble a basic pyramid structure by stacking cubic blocks, testing their ability to accurately position objects using the chosen interaction method.
- Task 2: Participants were tasked with stacking cubes on top of existing structures, requiring precise vertical placement and spatial awareness. The arrows in the figure indicate the intended placement positions for the cubes.
- Task 3: Users were required to insert a cube through an open frame without touching the sides, testing their fine motor control and depth perception.
- Task 4: Participants had to insert a cube through a narrow opening and navigate it to a specific target location within a complex structure, testing both their precision and their ability to manipulate objects in confined spaces.

Each participant performed all four tasks using the controller-based system first, followed by the PHI-NOM system. This consistent order ensured that any performance improvements were due to the interaction method and not user familiarity with the tasks. For each task, participants were instructed to complete the assembly as quickly and accurately as possible, and their performance was measured in terms of completion time, number of errors, and qualitative feedback on the ease of use and sense of immersion.

C. Data Collection and Scoring

To further evaluate the effectiveness of PHI-NOM compared to traditional controller-based interactions, participants completed an anonymous feedback form immediately after completing the tasks. This form consisted of five questions designed to assess user perceptions of realism, intuitiveness, focus, challenge, and training preference. Each question was scored using a seven-point Likert scale, where participants indicated their level of agreement or disagreement with each statement:

- Question 1: ”Grabbing objects with PHI-NOM felt more life-like than grabbing using controllers.”
- Question 2: ”Grabbing objects with PHI-NOM felt more intuitive than grabbing using controllers.”
- Question 3: ”Grabbing objects with PHI-NOM required more focus than grabbing with controllers.”
- Question 4: ”Grabbing using PHI-NOM was more challenging than grabbing with controllers.”
- Question 5: ”If training for a real-world task, I'd prefer to practice with PHI-NOM over practicing with controllers.”

Participants responded to each question using a seven-point scale, where their responses were quantified as follows:

- Strongly Disagree: -3
- Disagree: -2
- Slightly Disagree: -1
- Both Felt Similar: 0
- Slightly Agree: 1
- Agree: 2
- Strongly Agree: 3

We calculated the average score for each question using this point system, and the results are presented in Table I.

TABLE I
AVERAGE SCORES FOR PARTICIPANT FEEDBACK ON PHI-NOM

Question	**Average Score**
Grabbing objects with PHI-NOM felt more life-like than using controllers	2.8
Grabbing objects with PHI-NOM felt more intuitive than using controllers	2.5
Grabbing with PHI-NOM required more focus than using controllers	2.6
Grabbing with PHI-NOM was more challenging than using controllers	2.5
If training for a real-world task, I'd prefer PHI-NOM over controllers	2.2

D. Analysis of User Feedback

The results from the participant feedback survey provide important insights into user perceptions of PHI-NOM compared to traditional controller-based interactions:

- Realism (Question 1, Average Score: 2.8): Participants overwhelmingly agreed that grabbing objects with PHI-NOM felt more life-like than using controllers. This high score reflects PHI-NOM's ability to simulate natural hand-object interactions, providing users with a sense of direct physical contact rather than relying on symbolic gestures or button presses.
- Intuitiveness (Question 2, Average Score: 2.5): Most participants found PHI-NOM to be more intuitive than controllers. This result suggests that the direct, physics-based hand interactions align more closely with users' expectations of how objects should behave, reducing the learning curve.
- Focus (Question 3, Average Score: 2.6): Participants reported that grabbing with PHI-NOM required more focus. This is likely due to two reasons: (1) PHI-NOM demands precise hand movements and careful alignment of the fingers to achieve a valid grasp, unlike controllers, where a single button press is sufficient; and (2) the lack

of force feedback means users must exert extra effort to ensure they don't close their hand too much. Notably, this increased focus might lead to more effective learning, as users must actively engage their fine motor skills and maintain spatial awareness during interactions.
- Challenge (Question 4, Average Score: 2.5): The slight increase in perceived challenge further reinforces that PHI-NOM's realism requires users to actively maintain a secure grip, making interactions more complex. However, this challenge is a necessary aspect of building fine motor skills and muscle memory.
- Training Preference (Question 5, Average Score: 2.2): Despite increased focus and challenge, participants still expressed a preference for training with PHI-NOM over controllers, suggesting that users recognize the value of a more realistic and skill-building interaction method.

The high scores across all questions indicate a strong user preference for PHI-NOM over traditional controller-based methods, despite the increased complexity of interactions. However, the lower score of 2.2 for training preference suggests that while PHI-NOM is taking a step in the right direction, there is still room for improvement in user experience. This score indicates that while users recognize the benefits of PHI-NOM, further refinements—such as enhancing ease of use or reducing cognitive load—could make it even more appealing. Nevertheless, these results align with the core hypothesis that PHI-NOM provides a more realistic and skill-enhancing VR experience.

VI. APPLICATIONS

Although PHI-NOM has yet to be deployed in astronaut training with active crew members, its potential for application spans a wide array of domains where fine motor control and proprioceptive accuracy are critical. Any training scenario that requires precise hand movements, such as surgical procedures, mechanical maintenance, or assembly-line operations, could greatly benefit from PHI-NOM's physically grounded, controller-free interaction paradigm, which emphasizes natural hand-object interaction based on real-time feedback from the user's body position [8].

In aerospace contexts, PHI-NOM can enhance simulations for equipment handling, extravehicular activity (EVA) rehearsal, or microgravity object manipulation, where fine motor skills must be refined in conditions that simulate the constraints of space environments. As highlighted by O'Connor et al. [9], mastering such complex skills in microgravity requires simulation systems that accurately replicate the dynamic and tactile sensations of manipulating tools in space, making PHI-NOM an ideal candidate for these applications. Additionally, PHI-NOM can facilitate training scenarios that do not rely on visual feedback alone, instead using proprioception and touch-based cues to replicate real-world constraints more effectively [10]. In surgical education, PHI-NOM offers the possibility of simulating hand-intensive procedures without the need for costly haptic feedback systems or robotic mannequins. As a result, it presents a more flexible, lower-cost alternative to traditional training methods, which can be restrictive in terms of accessibility and scalability [11]. This model allows students to develop muscle memory for delicate tasks like suturing, tissue handling, or organ manipulation by using their hands directly in a virtual environment, as opposed to relying on mechanical proxies that may limit realism [12]. In industrial and maintenance training, PHI-NOM can be integrated into simulation environments where users need to learn to manipulate tools or components in confined or high-risk settings, such as industrial maintenance or hazardous material handling. Research has shown that physically grounded simulations improve skill transfer by allowing trainees to develop more intuitive task performance, which is critical for safety in high-stakes environments [13]. PHI-NOM's controller-free interaction offers an added benefit: it enables trainees to perform these tasks in environments that reflect real-world spatial and physical constraints.

A key strength of PHI-NOM lies in its support for multi-user collaboration. This feature extends the framework's applications to scenarios where teamwork and shared physical space are paramount, such as spacecraft repair, disaster response coordination, or cooperative robotic assembly. The ability for multiple users to manipulate shared virtual objects in real-time is essential for improving team-based skills such as spatial reasoning, communication, and joint problem-solving, all of which are vital in mission-critical contexts [14].

VII. LIMITATIONS

While PHI-NOM significantly advances the realism and intuitiveness of hand-object interaction in virtual environments, it is not without limitations. Chief among these is the absence of haptic or force feedback. Because users cannot physically feel virtual objects, they must rely entirely on visual cues to determine when contact occurs. This disconnect can hinder proprioceptive precision, making it difficult to know when to stop closing the hand or to recognize when sufficient contact has been made, especially during delicate tasks or with small objects. The integration of haptic or resistive force-feedback gloves could substantially mitigate this issue, enabling more accurate motor learning and a stronger sense of physical realism.

Additionally, PHI-NOM does not currently simulate object mass or inertia. All objects, regardless of their virtual material or geometry, are treated as weightless from the user's perspective. As a result, users cannot develop intuitive expectations about how much force is needed to lift or stabilize a given object. This limits the transferability of strength-based motor skills and may lead to unrealistic behavior when simulating tasks like lifting heavy tools or managing momentum in zero-gravity environments.

These limitations point to important avenues for future work, including the integration of haptics, adaptive physics modeling, and variable-resolution interaction sampling, to further close the gap between virtual and physical training environments.

VIII. CONCLUSION

Realistic hand-object interaction remains a significant frontier in immersive virtual reality training. While many systems have made strides in visual fidelity and gesture recognition, few have captured the nuanced, physics-based dynamics that define real-world manual dexterity. PHI-NOM addresses this challenge by offering a controller-free, biomechanically grounded framework for hand interaction that mirrors the physical constraints and affordances of natural manipulation. Rather than relying on symbolic gestures or pre-defined animations, PHI-NOM empowers users to manipulate objects the same way they would in the real world—using continuous motion, physical geometry, and dynamic feedback. This opens the door to more effective training for tasks where muscle memory and proprioception are critical, such as astronautics, surgery, manufacturing, and robotics.

Using real-time motion capture and physically plausible grasp modeling, PHI-NOM enables users to interact with virtual objects in a manner that is both intuitive and functionally transferable. Across four manipulation tasks, users consistently reported that PHI-NOM provided a more realistic and intuitive experience than traditional controller-based systems, particularly in terms of subjective realism and skill transfer.

The results of the user study also suggest potential avenues for improvement, including the incorporation of haptic feedback and the simulation of mass and inertia. Notably, a key advantage of PHI-NOM is its scalable design, which allows for seamless integration of future enhancements. Since PHI-NOM already organizes precise contact point data for each fingertip, it is well-positioned to support the addition of haptic gloves or other force-feedback devices. These enhancements could leverage existing contact point data to lock rotation or apply force feedback at specific joints, further increasing the realism of interactions. Additionally, PHI-NOM has already been validated for use with motion capture gloves, demonstrating its compatibility with external devices. This flexibility ensures that PHI-NOM can continue to evolve without requiring a complete system overhaul.

The foundation laid by PHI-NOM demonstrates that physics-based fine-grained interaction is both technically feasible and experientially meaningful. As future iterations integrate tactile feedback, adaptive physical modeling, and larger collaborative spaces, PHI-NOM has the potential to set a new standard for realism in VR training, bringing us closer to virtual experiences that truly teach what they aim to simulate.

REFERENCES

[1] A. Hameed, S. Möller, and A. Perkis, “How good are virtual hands? Influences of input modality on motor tasks in virtual reality,” *Journal of Environmental Psychology*, vol. 92, 2023.

[2] A. Schäfer, G. Reis, and D. Stricker, “The Gesture Authoring Space: Authoring Customised Hand Gestures for Grasping Virtual Objects in Immersive Virtual Environments,” in *Proc. Mensch und Computer 2022*, Darmstadt, Germany, 2022, pp. 85–95. doi: 10.1145/3543758.3543766.

[3] F. Kyota and S. Saito, “Fast grasp synthesis for various shaped objects,” *Computer Graphics Forum*, vol. 31, no. 2pt4, pp. 765–774, 2012.

[4] S. Marichal, I. Ezcurdia, R. Morales et al., “Hand-as-a-prop: using the hand as a haptic proxy for manipulation in virtual reality,” *Virtual Reality*, vol. 27, pp. 2911–2925, 2023.

[5] M. Feick, S. Bateman, A. Tang, A. Miede, and N. Marquardt, “Tangi: Tangible Proxies For Embodied Object Exploration And Manipulation In Virtual Reality,” in *Proc. IEEE Int. Symp. Mixed and Augmented Reality (ISMAR)*, Porto de Galinhas, Brazil, 2020, pp. 195–206.

[6] M. Kok, M. H. Everts, and M. Poel, “A Visually Plausible Grasping System for Object Manipulation and Interaction in Virtual Reality Environments,” *Computers and Graphics*, vol. 83, pp. 1–10, Oct. 2019. doi: 10.1016/j.cag.2019.07.003.

[7] K. Nasim and Y. J. Kim, “Physics-based Assistive Grasping for Robust Object Manipulation in Virtual Reality,” *Computer Animation and Virtual Worlds*, vol. 29, no. 3–4, e1820, May 2018. doi: 10.1002/cav.1820.

[8] D. Johnson, J. Williams, and M. Lee, “Controller-free hand interaction in virtual training environments: Benefits and limitations,” Virtual Reality, vol. 25, no. 2, pp. 45–63, 2020.

[9] L. O’Connor, S. Thomas, and D. Richards, “Enhancing astronaut training using immersive virtual environments,” Aerospace Systems, vol. 31, no. 5, pp. 87–101, 2019.

[10] M. Krasny, L. Segall, and M. Whiting, “Proprioceptive feedback in virtual training: A review of controller-free interfaces,” Journal of Human-Machine Interaction, vol. 30, no. 4, pp. 101–115, 2021.

[11] J. Simpson, L. Vickery, and J. Whitfield, “Cost-effective simulation technologies in medical education,” Medical Education, vol. 51, no. 10, pp. 1074–1080, 2017.

[12] J. Rosen, D. Brown, and J. Fung, “The effectiveness of simulation in medical education,” Journal of Surgical Education, vol. 66, no. 4, pp. 256–263, 2009.

[13] M. Gutiérrez and M. Medina, “Industrial training applications using virtual reality and controller-free interaction,” Virtual Reality, vol. 25, no. 1, pp. 25–40, 2021.

[14] T. Tessendorf and M. Wightman, “Collaborative virtual environments for team-based task training,” Computers in Human Behavior, vol. 34, no. 2, pp. 56–67, 2022.

[15] Leap Motion, ”Leap Motion Controller,” [Online]. Available: https://www.leapmotion.com/. [Accessed: May 11, 2025].

[16] Microsoft Kinect, ”Microsoft Kinect Documentation,” [Online]. Available: https://learn.microsoft.com/en-us/previous-versions/windows/kinect/. [Accessed: May 11, 2025].