GeoCtrl: Finger-based Geometric Mapping and Speed-Adaptive Gain for In-Hand Object Rotation in AR/VR

Seo Young Oh*
KAIST UVR Lab.
KAIST HCI Tech Lab.

Taejun Son
KAIST UVR Lab.

Juyoung Lee
KAIST UVR Lab.

Sang Ho Yoon†
KAIST HCI Tech Lab.

Woontack Woo†
KAIST UVR Lab.
KAIST KI-ITC ARRC

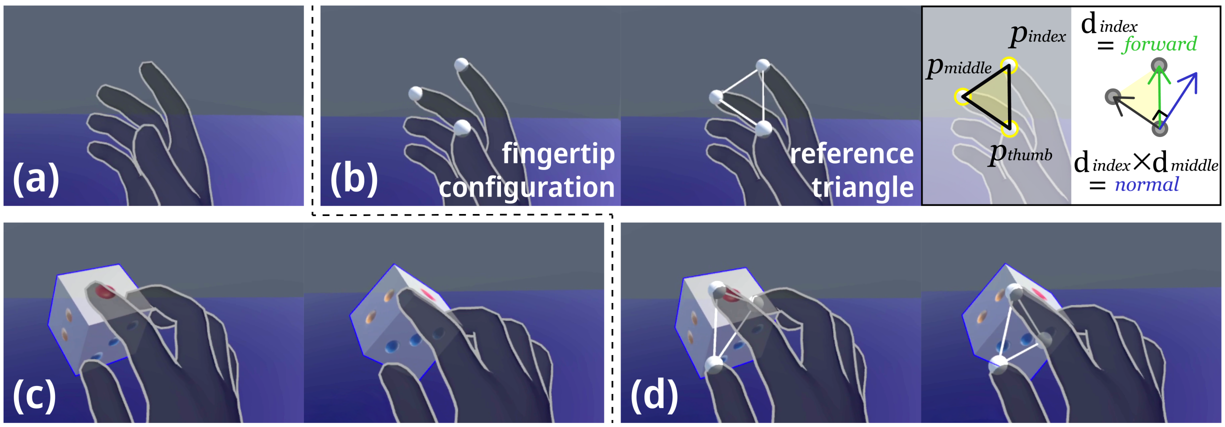

Figure 1: GeoCtrl enables dexterous in-hand rotation of a virtual object through finger-level geometric mapping: (a) while the normal virtual hand is visible to the user, (b) an internal reference triangle is formed based on three fingertips, including thumb, index, and middle, which convert geometric changes in finger configuration into the virtual object's rotation. (c) When the user traces the desired motion of the object with fingers, (d) the orientation and morphology changes of the reference triangle are directly applied to the object's rotation. GeoCtrl supports gestural approximations without requiring precise motor control.

Abstract

We propose a geometry-based virtual object rotation technique that leverages triangular configuration of fingertips and a speed-responsive Control-Display (CD) gain. Rather than replicating real-world physics which often lead to unnatural mid-air behavior in the absence of physical feedback, we leverage finger-level geometry to enable low-effort object rotation. We map the orientation and morphology of a triangle formed around fingertips directly to the object's rotation, assuming that users tend to trace the object's intended motion when physical constraints are missing. To enhance the proposed interaction and accommodate natural motor limits, we implement an adaptive CD gain that scales geometric input according to the fingertip movement speed. In a user evaluation, our method yielded shorter task completion time than physics-based rotation, though longer than a commodity hand-interaction technique, while significantly reducing physical demand relative to both reference conditions without increasing overall workload. We demonstrate that geometry-driven mapping and dynamic gain adaptation enable low-effort, finger-level rotational control of virtual objects in immersive environments.

Index Terms: Dexterous Manipulation, Object Rotation, Hand Interaction, 3D Interaction Techniques, Adaptive Control-Display Gain, Virtual and Augmented Realities

*E-mails: {oh.seo, tj.son, ejuyoung, sangho, wwoo}@kaist.ac.kr.
†Co-corresponding authors.

1 Introduction

Advancements in visualization technologies for Augmented Reality (AR) and Virtual Reality (VR) have paved the way for professional XR applications within increasingly complex 3D environments. For instance, tasks that demand high precision and fidelity, such as surgical simulations, immersive education, and assembly of industrial machinery, now necessitate interaction with digital assets at a granular level. While these professional scenarios require precise and dexterous manipulation, current interaction techniques often fall short when handling intricate virtual objects. As controller-free, bare-hand interaction becomes central to the scalability and accessibility of XR, intuitive techniques that support dexterous rotation of virtual objects remain an essential challenge.

To achieve realistic in-hand manipulation, previous studies have largely focused on high-fidelity physics simulation that replicates real-world constraints such as friction, gravity, and collisions [19, 22, 53]. Recently, Lougiakis et al. [33] (publicly available hand physics simulation systems) constructed interlinked articulations on the NVIDIA PhysX 4 engine to robustly model realistic interaction with virtual objects. However, these models often result in an interaction gap in mid-air XR environments, where the essential haptic cues and physical support of a real object are absent, even with an advanced physics engine. Without the necessary tactile feedback, users struggle to execute the delicate finger coordination required by physics engines [9], leading to unstable and unintuitive rotation.

Instead, we propose GeoCtrl, a geometric rotation technique that enables mid-air finger manipulation without demanding exact motor precision. We derived our method from the observation that without physical supports, users tend to gesturally represent the object's intended motion with their fingers [1, 43]. For instance, a user's fingertips might collectively trace a clockwise arc when intending to rotate a virtual object clockwise. Our system forms an imaginary reference triangle using three fingertips, which are the minimum required for a stable grasp, and directly maps these geometric changes to the transformation of the virtual object. The proposed mechanism alleviates the user's burden of replicating exact physical behaviors by accommodating coarse gestural approximations of the desired rotation.

To enhance the efficiency and controllability of the geometric mapping mechanism, we implemented a speed-responsive CD gain that adapts to fingertip movement speed. While a direct, unmodified mapping provides a realistic illusion, it may not be the most efficient way to rotate a virtual object. Drawing on prior work in traditional 2D and recent 3D interaction that utilizes adaptive CD gain [39, 49, 59], we expected the CD gain adaptation would enhance both interaction controllability and efficiency. Through a controlled study, we empirically identified the CD gain curve that best facilitates both the ballistic and corrective phases of rotation.

We conducted a within-subject user evaluation with 24 participants, comparing GeoCtrl against a physics-based hand interaction system and a commodity hand interaction system on a dexterous object docking task. GeoCtrl significantly reduced task completion time compared to the physics-based system, confirming that geometric mapping provides more effective rotational control in the absence of haptic feedback. In terms of perceived workload, GeoCtrl imposed significantly lower physical demand than both reference conditions, highlighting the advantage of finger-level dexterous interaction. These results suggest that GeoCtrl offers more efficient, lower-effort finger-level rotational control, complementing rather than replacing commodity hand interaction.

Our contribution lies in the proposal and validation of a geometry-driven manipulation technique, as well as the implementation and exploration of a dynamic CD gain suitable for the proposed method. We expect this technique to help connect fine finger coordination with dexterous manipulation in XR. In the following sections, we elaborate on the geometry mapping and CD gain adaptation mechanisms. Also, we describe the experimental exploration of the gain curve and report comparative results against the existing standard method and recent physics-based rotation methods.

2 Related Works

2.1 Rotation Techniques for Virtual Objects

Rotation of virtual objects has long been a challenging problem in 3D interaction, with common approaches relying on gizmo-based interfaces or handheld controllers [20, 49]. Various conventional input modalities have been explored for 3D object manipulation, including mouse, touch, and tangible input [4] and rolling input devices [44]. In particular, direct-touch techniques for 3D manipulation on 2D surfaces provide well-established examples of mapping finger contact coordinates from the input space into a distinct virtual output space [48, 35, 17]. Custom transformation axes for degree-of-freedom separation [36], pin-based constrained manipulation [16], and adaptive scaled mapping [49] have also been proposed to improve precision of rotation with controllers.

To enable more dexterous in-hand rotation, dedicated haptic and tangible input devices have been proposed. Most approaches include finger-level haptic sensation with a graspable controller [31, 32] or ball-shaped controllers [27, 28, 51] to leverage real-world hand dexterity. Tangible proxies such as spherical objects and everyday items have also been employed as passive controllers for object manipulation in AR [11, 12]. However, previous approaches require hardware beyond bare hands, leaving finger-level rotation without physical support underexplored.

The closest bare-hand work investigates the degree-of-freedom separation between finger and wrist. This reflects a similar principle to constraint-based approaches [37], but does not specifically address fine-grained in-hand rotation. More broadly, most bare-hand finger interaction techniques target system control tasks such as interface navigation and indirect pointing [26, 29], rather than object manipulation. GeoCtrl addresses this gap by introducing a geometric mapping that translates finger movements directly into in-hand object rotation without requiring physical hardware.

2.2 Realistic Interaction and Physics Simulation

Physics simulation is considered the most natural approach for bare-hand interaction with virtual objects, aiming to replicate real-world dynamics without additional hardware. These simulations model soft tissue deformation [21, 23, 55], contact constraints [53], and friction to replicate real-world hand-object interaction [19, 22]. However, achieving realistic physics behavior depends on users producing precise contact forces and finger configurations that occur naturally only in the presence of physical resistance. Without haptic feedback, users cannot replicate the precise motor behaviors these engines assume [9], resulting in artifacts such as interpenetration and unstable object trajectories. Consequently, few physics solutions have proven usable in practice [10, 33].

To bridge this gap, researchers have explored substitutes such as visuo-haptic cues that simulate contact [41] and algorithmic grasp generation that synthesizes plausible contact configurations [24, 54]. Yet none fully compensates for the absence of physical contact during object manipulation. Rather than replicating constraints users cannot feel, GeoCtrl accommodates coarse gestural approximations through geometric mapping, eliminating the burden of exact behavior reproduction.

2.3 Control-Display Gain for 3D Interaction

The concept of CD gain has long been used to overcome the limits of human motor control, facilitating faster and more precise target acquisition and object manipulation. CD gain has been studied in both 2D pointing tasks [8, 13] and 3D interaction [14, 45]. More recently, CD gain has been employed in fine-grained 3D selection and manipulation tasks, including precise direct selection [59], multi-target interaction [25], and force-modulated raycasting [42]. Beyond pointing and translation, CD gain has also been applied to 3D rotation where non-isomorphic rotational mappings were shown to reduce task completion time without loss of accuracy [47, 46]. PRISM interaction demonstrated scaled rotation of virtual objects [15], and AMP-IT applied adaptive scaled mapping to high-precision 3D manipulation tasks [49]. However, existing rotational gain techniques relied on physical devices or hand-level object manipulations. GeoCtrl introduces a relationship between finger movement and in-hand object rotation, applying the concept of CD gain on dexterous finger-level interaction.

In terms of gain function design, some systems employ a fixed gain for predictable, stable interaction [8, 13, 42], while others use adaptive gain that adjusts to the interaction context, such as target proximity [5], target distance [45], or hand configuration [25, 59]. Among adaptive strategies, speed-responsive gain is the most prevalent approach, grounded in the natural duality of human movement [14, 30, 56]. Users move faster during the ballistic phase for coarse adjustments and slower during the corrective phase for fine corrections, following the two-phase model of motor control [40]. While speed-responsive gain has been extensively studied for pointing and translation, its application to finger-level rotation remains underexplored. GeoCtrl extends these principles with behavior-responsive gain adapting to fingertip speed, specifically targeting finger-level virtual object rotation. By tuning the gain curve to facilitate ballistic and corrective phases, we enhance dexterous control without sacrificing efficiency.

3 System Design

3.1 Geometry-based Rotation

GeoCtrl provides an intuitive in-hand rotation technique for virtual objects without requiring exact replication of real-world behaviors. In the absence of physical support and tactile feedback, a user's

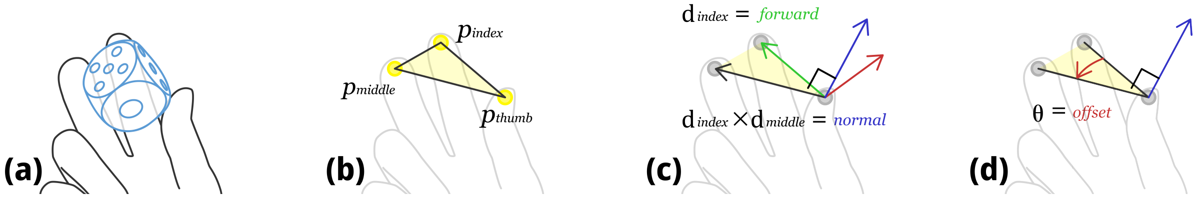

Figure 2: Geometry-based rotation mechanism: (a) a grab is initiated with three fingers, (b) the fingertips form a reference triangle, (c) its orientation $Q_{\text{orientation}}$ is set by the forward vector $\mathbf{d}_{\text{index}}$ and the normal $\mathbf{d}_{\text{index}} \times \mathbf{d}_{\text{middle}}$, and (d) the offset Q_{offset} by the angle θ between $\mathbf{d}_{\text{index}}$ and $\mathbf{d}_{\text{middle}}$. The combined change $Q = Q_{\text{orientation}} * Q_{\text{offset}}$ is applied to the object.

finger coordination can easily conflict with the virtual object's constraints, leading to deviations from reality such as interpenetration. Consequently, users struggle to perform the precise motor actions required by physics simulations, leading to unintended transformations and diminished interaction quality. To bridge this gap, we propose a geometric mechanism that maps fingertip configurations directly to the virtual object's orientation.

Our approach assumes that users gesturally represent an object's intended motion when physical proxies are missing. For instance, a user would trace a clockwise arc with their fingertips to indicate an intent to rotate a virtual object clockwise along the corresponding axis. While various hand postures can hold small objects, we focus on the three-fingered precision pinch, the minimal configuration affording stable grasp control [34, 58], using the thumb, index, and middle fingers as reference points.

When the user initiates a grab gesture with the three fingers (Fig.2(a)), an imaginary reference triangle is formed by the coordinates of the three fingertip (Fig.2(b)). To ensure interaction stability, a 1-Euro filter [7] is applied on each fingertip position to mitigate tracking jitter ($f_c = 3.0$, $\beta = 0.66$). The initial orientation of the triangle, $Q_{\text{orientation}}$ is then determined by a forward vector $\mathbf{d}_{\text{index}} = \text{norm}(p_{\text{index}} - p_{\text{thumb}})$ and a normal vector derived from the cross-product of $\mathbf{d}_{\text{index}}$ and $\mathbf{d}_{\text{middle}} = \text{norm}(p_{\text{middle}} - p_{\text{thumb}})$ (Fig.2(c)). Acknowledging the convention that the thumb plays a dominant role in hand activities, we designated the thumb as the anchor point for these geometric calculations.

However, $Q_{\text{orientation}}$ alone does not capture the in-plane motion of the middle finger. When p_{middle} moves within the triangle's plane, both the forward vector $\mathbf{d}_{\text{index}}$ and the plane normal $\mathbf{d}_{\text{index}} \times \mathbf{d}_{\text{middle}}$ remain unchanged, leaving $Q_{\text{orientation}}$ unaffected. To fully integrate the movements of all three fingertips, we introduce an additional "offset" calculation to account for morphological changes in the triangle. The angular displacement of $\mathbf{d}_{\text{middle}}$ relative to $\mathbf{d}_{\text{index}}$ is used to derive a rotation along the normal axis, denoted as Q_{offset} (Fig.2(d)). The final geometry of the reference triangle is represented by the combined quaternion $Q_{\text{curr}} = Q_{\text{orientation}} * Q_{\text{offset}}$, ensuring that both the spatial orientation and the internal configuration of the fingers affect the result. When the middle finger instead moves purely along $\mathbf{d}_{\text{middle}}$, radially relative to the thumb, the direction of $\mathbf{d}_{\text{middle}}$ is preserved, so $Q_{\text{orientation}}$, Q_{offset}, and Q_{curr} all remain unchanged. By design, such radial extension induces no rotation.

The relative rotation between consecutive frames, defined as $\Delta Q = Q_{\text{curr}} * Q_{\text{prev}}^{-1}$, is directly applied to the virtual object's transformation such that $Q_{\text{obj,curr}} = \Delta Q * Q_{\text{obj,prev}}$. Unlike force-based models, this mapping remains robust even when users perform coarse or relaxed movements that would otherwise cause stability issues in physics engines. All calculations are performed in the local wrist coordinate frame to decouple finger manipulation from wrist rotation. The object's own orientation serves as continuous feedback of rotation progress, and no separate visual rotation indicator is provided.

3.2 Adaptive Control-Display Gain

In both traditional 2D and contemporary 3D interfaces, the application of CD gain is a well-established method for mitigating the inherent limitations of human motor control [8, 13, 42]. Adaptive CD gain, in particular, benefits users by facilitating both a rapid approach to a target and high-precision adjustments [25, 45, 59]. While GeoCtrl maintains a realistic illusion, the mapping does not require a 1:1 ratio, and CD gain integration can further enhance efficiency and controllability.

Speed-responsive CD gain has been frequently proposed across various interaction paradigms to optimize performance based on the user's operation speed [14, 15, 30, 49, 57]. The core principle of such mechanisms relies on the duality of human movement. Users typically exhibit high speed during the ballistic stage for large-scale adjustments and lower speed during the corrective stage for fine-tuning. Consequently, high fingertip speeds trigger higher gains for large-scale rotations, whereas slower movements allow for fine adjustments.

To implement this speed-responsive CD gain, we adopted a sigmoid function modeled after prior literature [39, 56]:

$$g(v) = \frac{g_{\max} - g_{\min}}{1 + e^{k(v_{\inf} - v)}} + g_{\min} \tag{1}$$

where v represents the maximum speed among the three fingertips in each frame, and k controls the steepness of the gain transition. To apply the gain, the rotational increment ΔQ is converted into an angle-axis representation, where the angle is scaled by $g(v)$ to produce the modulated transformation ΔQ_g. This resulting transformation is then applied to the object's orientation, $Q_{\text{obj,curr}} = \Delta Q_g * Q_{\text{obj,prev}}$ [46, 47]. Consistent with our geometric mapping, all computations are performed within the local wrist coordinate frame to independently handle finger-level manipulation. The detailed parameter tuning for the sigmoid function is described in Section 4.

3.3 Implementation Details

Aside from the core mechanisms, GeoCtrl incorporates two supplementary features. A grab-and-release logic manages state transitions without physics-based contact handling, and a clutching behavior lets users temporarily decouple input and output spaces to reset their hand pose. These features are ancillary to the primary contributions of geometric mapping and CD gain, and represent one workable implementation open to refinement as better solutions emerge.

In our implementation, a grab is triggered when all three fingertips make contact with the virtual object, thereby establishing the initial reference triangle. The release occurs when contact is lost for two or more fingertips. During manipulation, the virtual object is snapped to P_{obj}, the angle-weighted centroid of the reference triangle, $P_{\text{obj}} = \frac{\sum_n p_n/\theta_n}{\sum_n 1/\theta_n}$, where θ_n are the interior angles. This is only a positioning choice, separate from the rotation mapping, as rotation is applied in the object's own frame rather than

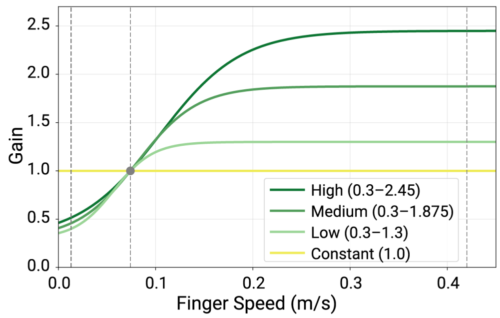

Figure 3: The four CD gain curves explored in Study 1. The Const condition applies no scaling. The Low, Mid, and High conditions each follow a sigmoid function that attenuates gain during slow finger movements and amplifies it during fast movements, at increasing rates. All sigmoid curves share the same inflection point v_{inf} and minimum gain $g_{\min}$, but differ in the maximum gain.

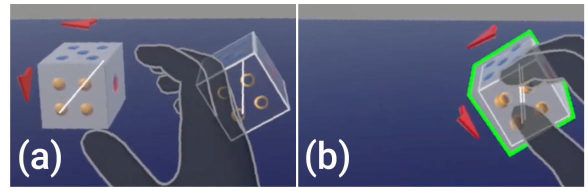

Figure 4: The docking task used in Study 1. (a) A solid die (object) is to be rotated to match the orientation of an outlined die (target). The rotation axis is indicated by a white line and two red arrows on the object. (b) A green outline indicates successful docking within the task completion threshold.

about P_{obj}. A characteristic of the precision grip is that the index and middle fingers remain close together while the thumb opposes them. Standard triangle centers are set by the triangle's overall geometry rather than by singling out the opposing thumb, so they drift toward the two close fingers or fall outside the grasp. Because the thumb vertex subtends the sharpest interior angle in this configuration, inverse-angle weighting gives it the greatest influence and restores the anchor to the center of the grip.

The clutching feature is essential for enabling continuous rotation when users reach their kinematic limits, such as when fingers become fully curled or pinched. Since such physical constraints naturally induce involuntary pause during finger movement, we implemented a dwell-based clutching logic that triggers when a momentary pause of the fingertips over 0.15 s is detected. This lets users reset finger positions without altering the object's orientation. Clutching exits automatically when the object is released.

4 Study 1: Exploration of the CD Gain Curve

GeoCtrl applies a speed-responsive CD gain on top of the geometric mapping to enhance rotation efficiency. Typically, the range of optimal CD gain values can be calculated from a device's resolution or operational range [8, 40]. Our geometric mapping uses a reference triangle that creates an indirect link between finger motion and object rotation, making the analytic derivation of optimal CD gain difficult to apply. Therefore, we conducted an empirical study to identify the gain curve that best balances efficiency, controllability, and user experience. Specifically, we compared four Gain Level conditions varying in sensitivity to finger speed across a range of target Rotation Angles and axes to determine the most effective configuration.

4.1 Construction of CD Gain Curves

We experimented with four different CD Gain Levels: Constant (Const), Low, Mid, High. In the Const condition, no scaling factor was applied, and the object rotation was directly mapped to the rotation of the reference triangle. The remaining three conditions, Low, Mid, and High, were designed to attenuate rotation during slow finger movements and amplify it during fast movements, at different rates. To achieve this, we manually tuned sigmoid functions of Eq. 1 for each condition.

First, we determined a reference fingertip speed to serve as both the unity-gain point (CD ratio = 1) and the inflection point v_{inf} of the sigmoid function. As there is no widely accepted standard for finger speed, we estimated the maximum and minimum speeds from prior work on finger-based Fitts' Law tasks. Sridhar et al. [52] investigated pointing tasks within the full effective range of each finger's MCP (Metacarpophalangeal) joints, while Balakrishnan et al. [2] focused on the left-right motion of the index finger, which involves more subtle movements. We used the results of the former to determine $v_{\max}$ and the results of the latter for $v_{\min}$, yielding $v_{\max} = 0.42$ m/s and $v_{\min} = 0.013$ m/s. We then took the geometric mean of $v_{\max}$ and $v_{\min}$, assuming a logarithmic distribution of finger movement speed, to obtain $v_{\text{inf}} = 0.074$ m/s.

Using these thresholds, we defined three gain curves (Fig.3). In all three conditions, $g_{\min}$ was set to 0.3, identified as the minimum effective CD gain in pilot testing. Below this value, the gain was too low to rotate the object in practice. $g_{\max}$ was set to 2.45, beyond which rotational overshoot became clearly noticeable. For the Low condition, the gain span was set to 1.0 above $g_{\min}$. The High condition spanned the full range from $g_{\min}$ to $g_{\max}$. The Mid condition used the average gain span of the two.

4.2 Experiment Design

Rotation Angle was included as a second within-subject independent variable with three levels: 30°, 120°, and 210°. The three angles, spaced 90° apart, were chosen to span from fine adjustment to larger-scale rotations requiring multiple clutching operations. They were presented in ascending order of anticipated difficulty. We avoided multiples of 90° because pilot participants judged orientation by the die's silhouette rather than its faces, which becomes ambiguous at axis-aligned poses. Even though 210° exceeds 180° and the opposite rotation would be shorter, participants were instructed to follow the given direction. To cover a diverse range of spatial orientations, we varied the rotation axis across the positive and negative directions of the x, y, and z axes, resulting in 6 axes total. The axes were presented in a randomized order within each block, and each participant completed 3 sets per block. Combined with 4 Gain Level conditions and 3 Rotation Angle levels, this yielded a total of 216 trials per participant.

We employed a docking task, which is a common experimental task in studies of rotational interactions [3, 15, 31], with two 4 cm semi-transparent dice. The task was to grab the object to control (a solid die), rotate the object in the designated direction, and match the orientation and position of the object to the target (an outlined die) (Fig.4). The object to control and the target to match were presented 30 cm in front of the participant with a 10 cm gap between, so that the dice were clearly separated but did not require excessive translation. The Rotation Angle was communicated to participants both verbally and as on-screen text, and the rotation axis was indicated by a line with red arrows on the object. A green outline indicated the successful docking (Fig.4(b)), and the task was completed when the docked status was maintained for 1 s. The grab status was indicated with a thin blue outline. The offset thresholds for task completion were 2 cm and 10 °. The position threshold was set with a large tolerance, as the experiment focused on rotational movements. We used a threshold-based dock-and-hold criterion rather than a release-based one, which would conflate an

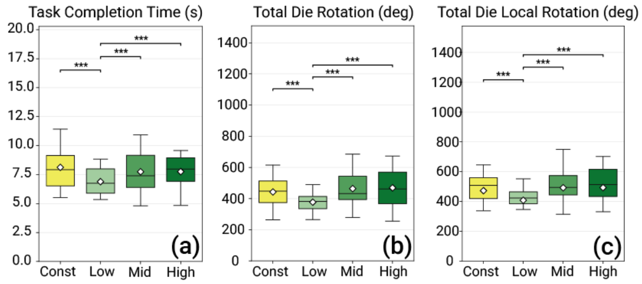

Figure 5: Study 1 results across four Gain Level conditions: (a) task completion time, (b) accumulated die rotation, and (c) accumulated die local rotation in the wrist coordinate frame.

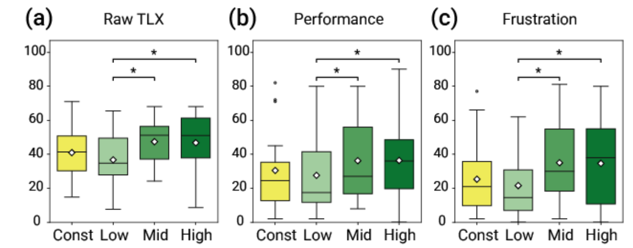

Figure 6: NASA-TLX workload scores across the four Gain Level conditions in Study 1: (a) overall score and (b) performance, and (c) frustration subscales.

intentional release with the release-based clutching. The 10° orientation threshold follows the size-normalized formula of Frees et al. [15]. Participants could use both hands, switching the active hand by releasing the object.

16 participants (22–$35, M = 27.7, SD = 3.96$, 12 males and 4 females) volunteered for the experiment through the institute's online community board. All participants were right-handed. The participants had varying levels of experience in realistic hand interaction in XR, ranging from no experience to expert users. The institutional review board approved the experiment and informed consent was collected from all participants. Participants were compensated with \$10 for 1 hour of participation. After being introduced to the GeoCtrl technique, participants were given as much time as they wished to practice the interaction and become familiar with the task. In each Gain Level condition, participants completed three blocks of Rotation Angles, each consisting of three sets of six rotation axes. The order of Gain Level conditions was counterbalanced across participants.

Along with task completion time, we measured total die rotation, defined as the cumulative orientation change across frames, to assess rotational efficiency across Gain Level conditions. We also recorded total die local rotation, defined as the sum of rotation angles measured in the wrist-based coordinate frame, to measure the rotation induced by finger movements. It should be noted that total die local rotation can be larger than total die rotation, as wrist rotation can counteract finger-based rotation. Perceived workload was assessed using the NASA Task Load Index (NASA-TLX) [18], with scores ranging from 0 to 100.

4.3 Results

All inferential analyses were conducted using Linear Mixed Models (LMM) with random intercepts per participant, accounting for trial nesting and inter-individual variability. Omnibus tests were performed via likelihood ratio tests (LRT), comparing a full model against an intercept-only null model fitted with maximum likelihood (ML). The magnitude of fixed and random effects is reported via marginal/conditional R^2 (R^2_m, R^2_c) following Nakagawa & Schielzeth [38]. Pairwise post-hoc comparisons were conducted using Wald z-tests on Estimated Marginal Mean (EMM) contrasts derived from REML fixed-effect estimates, with Benjamini-Hochberg false discovery rate (FDR) correction. Outliers were identified using Interquartile Range (IQR)-based detection on task completion time, yielding no outlier trials or participants[1].

Gain Level had a significant effect on task completion time ($\chi^2(3) = 30.626, p < .001, R^2_m = 0.008, R^2_c = 0.065$). Post-hoc comparisons showed that task completion time was shorter in the Low condition than in Const, Mid, and High (all $p < .001$) (Fig.5(a)). Including Rotation Angle and its interaction also significantly predicted task completion time ($\chi^2(11) = 676.400, p < .001, R^2_m = 0.167, R^2_c = 0.225$), and the Low condition's advantage held across all three angles.

Gain Level also had a significant effect on total die rotation ($\chi^2(3) = 40.070, p < .001, R^2_m = 0.011, R^2_c = 0.070$), with total die rotation in the Low condition being smaller than in Const, Mid, and High (all $p < .001$) (Fig.5(b)). The interaction model for total die rotation likewise improved on the null ($\chi^2(11) = 1282.756, p < .001, R^2_m = 0.288, R^2_c = 0.358$), with the Low condition producing smaller rotation than all other levels at every angle.

Gain Level also significantly affected total die local rotation ($\chi^2(3) = 33.127, p < .001, R^2_m = 0.010, R^2_c = 0.049$), with total die local rotation in the Low condition being smaller than in Const, Mid, and High (all $p < .001$) (Fig.5(c)). The same pattern held for total die local rotation ($\chi^2(11) = 1585.993, p < .001, R^2_m = 0.353, R^2_c = 0.394$), where the Low condition again produced the smallest rotation across all angles.

For the NASA-TLX, Gain Level had a significant effect on Raw TLX ($\chi^2(3) = 10.432, p = .015, R^2_m = 0.070, R^2_c = 0.590$), Performance ($\chi^2(3) = 8.925, p = .030, R^2_m = 0.023, R^2_c = 0.842$), and Frustration ($\chi^2(3) = 10.801, p = .013, R^2_m = 0.059, R^2_c = 0.668$) (Fig.6). In all three subscales, the Low condition was rated more favorably than High and Mid (all $p < .05$), while no significant difference was found between Low and Const. Physical Demand showed a significant omnibus effect ($\chi^2(3) = 8.396, p = .038, R^2_m = 0.057, R^2_c = 0.577$), but no pairwise comparisons survived BH correction. Mental Demand, Temporal Demand, and Effort showed no significant omnibus effects.

4.4 Discussion

In summary, the Low gain condition consistently outperformed the alternatives across all three objective measures. Gain Level had a significant main effect on all three measures (all $p < .001$), and the advantage of the Low condition was maintained across all Rotation Angles in the interaction models. The Low condition reduced task completion time and total rotational movement compared with all other conditions, reaching significant differences in the majority of comparisons across angles. The Mid and High conditions did not differ significantly from each other and were comparable to the Const condition in task completion time, while producing greater total rotation. The Low condition also produced lower Raw TLX, Frustration, and Performance ratings than Mid and High (all $p < .05$), with no significant difference from Const.

The performance advantage of the Low condition over the Const condition suggests that attenuating slow finger movements improved rotational efficiency, reducing overshoots and instabilities. The Mid and High conditions instead reduced usability relative to Const. Because users can briefly exceed v_{inf} even during fine adjustments, the steep gain rise above v_{inf} in these conditions turned such exceedances into disproportionately large rotations. The Low condition avoided this, as its gain stays near or below unity across

[1] Analyses were performed in Python 3.13.12 using statsmodels 0.14.6.

the practical movement range. With clear improvements in task performance in the Low condition without increased subjective workload relative to the Const condition, we adopted the Low gain curve as the CD gain configuration for GeoCtrl in the subsequent system evaluation.

5 Study 2: System Evaluation

Having established the reliable CD gain configuration, we conducted a comparative evaluation to assess the benefits of GeoCtrl. We hypothesized that, in the absence of physical support and haptic feedback, realistic physics simulation would fail to provide sufficiently effective rotational control. To empirically verify this, we compared GeoCtrl against a physics-based hand interaction system in a dexterous manipulation scenario. We also compared the proposed system against a commodity hand interaction system, which is designed primarily for robustness and ease of use. We examined whether dexterous rotational control offers measurable advantages and validated that GeoCtrl remains comparably intuitive.

5.1 Implementation and Apparatus

From Section 4, we observed that unwanted rotation mostly occurs when the fingers are too close together. To address this, we added another clutching logic that deactivates rotation when the triangle area falls below an empirically determined threshold. We chose triangle area over interior angles because it captures absolute fingertip proximity, which scale-invariant interior angles cannot. Inspection of the triangle area distribution from Section 4 ($M = 4.39, SD = 4.45, \min = 0.001$) revealed a long lower tail. We set the threshold at the 5th percentile (< 0.70), which excluded frames with insufficiently separated fingertips while retaining 95% of all recorded frames.

We selected two reference conditions for comparison. For the Physics condition, we employed HPTK+ [33], one of the state-of-the-art physics-based hand simulation systems for realistic interaction with virtual objects. The HPTK+ system was deployed on Meta Quest 3 with Unity 6000.2.10f1. To match the experimental environment, the object ignored gravity and inertia to only be affected by the hand movement. The object's movement was cancelled immediately after contact was lost with the hands, effectively preventing the object from indefinitely drifting away from the hands. No other physical modifications were made on HPTK+ logic, to keep the hand interaction mechanism unchanged.

For the HandLevel condition, we used the Touch Hand Grab feature in the Building Blocks of Meta XR Interaction SDK 85.0.0. While this feature does not support dexterous finger movements during grab, its robustness makes it a representative commodity approach. The HandLevel condition was deployed on Meta Quest 3 with Unity 2022.3.62f1, the same environment as GeoCtrl. On grab, the system automatically snapped the object to the hand in the most feasible and stable position. Both visual and physical hands were fixed on grab, disabling the physical interaction between fingers and the grabbed object.

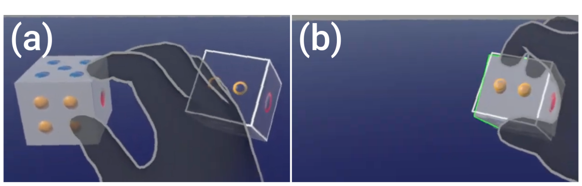

Figure 7: The docking task used in Study 2. (a) The object die is to be matched to a target die rotated 135° around a random axis, without explicit rotation axis guidance. (b) A green outline indicates successful docking within the completion threshold.

5.2 Experiment Design

As in Section 4, a traditional docking task with two 4 cm dice was employed. The object die was presented without any visual aids for rotation, and the target die was rotated 135 ° with respect to a random axis (Fig. 7(a)). The experimental task involved a visual search process to reflect use case scenarios where users determine the desired rotation by inspecting the object during the manipulation. The docking offset allowed for task completion were 2 cm and 10 ° and the docked status had to be maintained for 1 s to finish the trial. During the task, participants could use either hand, switching by releasing the object.

As before, a green outline indicated the successful docking between the object and the target (Fig. 7(b)). For the HandLevel and GeoCtrl conditions, the grab status was indicated with thin blue outlines to help users recognize the interaction state. The clutching state of GeoCtrl condition, in which the grabbed object is fixed to the wrist coordinate and not affected by finger movements, was visible with thicker outlines. For Physics condition, which does not explicitly define interaction states such as grab but provides continual contact interaction, the outline feedback was provided only to indicate the task completion.

We recruited 24 participants (22–46, $M = 27, SD = 4.44$, 11 males and 13 females) through the institute's online community board. All participants were right-handed. The participants' experience of realistic hand interaction in XR varied widely. The experiment was approved by the institutional review board and all participants provided their informed consent. Participants received \$10 for 1 hour of participation. The experiment was conducted using a within-subject design with three conditions, whose order was counterbalanced across participants. In each condition, participants practiced the technique until confident, then completed 3 blocks of 10 trials. After each condition, participants completed the NASA-TLX, a 7-point questionnaire, and optional open-ended comments on their experience. The questionnaire comprised the agency subscale of the Virtual Embodiment Questionnaire (VEQ) [50] and five custom items assessing Predictability (whether the object moved as expected), Realism (realism of the manipulation sensation), Physical Comfort, Understandability (ease of understanding the technique), and Familiarization (familiarity gained over the session). After all conditions, participants indicated the most and least precise technique and the most and least preferred technique, each accompanied by an open-ended explanation.

We recorded accumulated die rotation and translation, and, since the conditions afforded finger-based manipulation to different extents, wrist movements as well. We also measured time on target, defined as the interval from the object first reaching the completion threshold to the trial end, as participants tended to struggle to stabilize the object after losing secure control.

5.3 Quantitative Results

To maintain a consistent analytical framework, the same LMM procedure described in Section 4 was applied to all objective measures and NASA-TLX subscales, with condition sum-coded. No outlier trials or participants were identified with task completion time using IQR-based detection.

Condition had a significant effect on task completion time ($\chi^2(2) = 333.686, p < .001, R^2_m = 0.126, R^2_c = 0.244$) and time on target ($\chi^2(2) = 256.287, p < .001, R^2_m = 0.109, R^2_c = 0.135$). For both measures, all pairwise comparisons were significant (all $p < .001$), with HandLevel performing best, followed by GeoCtrl, and Physics worst (Fig. 8(a–b)).

Total die rotation differed significantly across conditions ($\chi^2(2) = 318.019, p < .001, R^2_m = 0.124, R^2_c = 0.222$). Die rotation was greatest in GeoCtrl, followed by Physics, and smallest in HandLevel (all $p < .001$). Total die translation also differed significantly ($\chi^2(2) = 581.753, p < .001, R^2_m = 0.215, R^2_c = 0.306$),

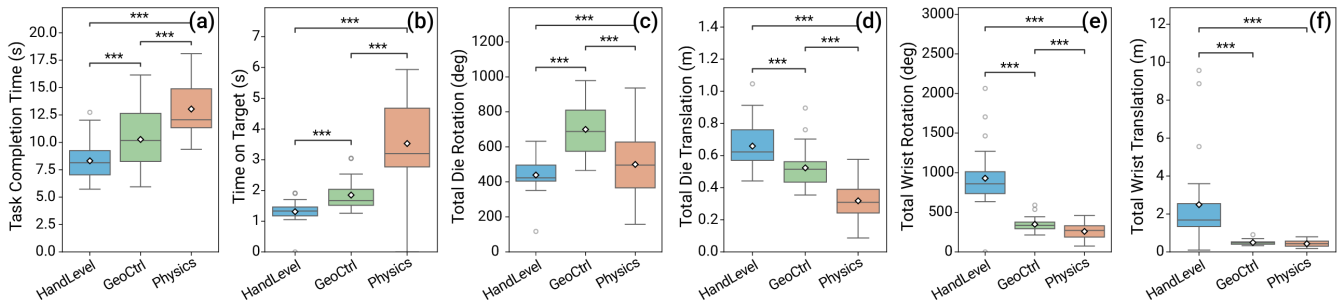

Figure 8: Quantitative results of Study 2 across the three interaction techniques: (a) task completion time, (b) time from the first collision on target to the trial end, (c) total die rotation and (d) translation, (e) total wrist rotation and (f) translation.

with the opposite ordering: HandLevel produced the greatest translation, followed by GeoCtrl, and Physics the least (all $p < .001$) (Fig. 8(c–d)).

Condition significantly affected total wrist rotation ($\chi^2(2) = 867.468, p < .001, R^2_m = 0.308, R^2_c = 0.379$) and total wrist translation ($\chi^2(2) = 262.812, p < .001, R^2_m = 0.107, R^2_c = 0.171$). Wrist rotation was greatest in HandLevel, followed by GeoCtrl, and smallest in Physics (all $p < .001$). For wrist translation, HandLevel was significantly greater than both GeoCtrl ($p < .001$) and Physics ($p < .001$), with no significant difference between the latter two (Fig. 8(e–f)).

We additionally analyzed the heatmap of participants' wrist movements (Fig. 9). Condition significantly affected both the workspace area ($\chi^2(2) = 14.327, p < .001$) and the KDE entropy ($\chi^2(2) = 11.999, p < .01$), which measures how uniformly wrist movements were distributed. For both measures, GeoCtrl was smaller than HandLevel ($p < .01$) and Physics ($p < .05$), with no difference between the two baselines, indicating that GeoCtrl confined manipulation to a more concentrated workspace.

For the NASA-TLX, condition had a significant effect on all subscales (Table 1, Fig. 10). Physics received significantly higher ratings than HandLevel and GeoCtrl on Raw TLX, Mental Demand, Temporal Demand, Effort, and Frustration (all $p \leq .01$), and lower perceived Performance than both (all $p < .05$). Physical Demand showed a different pattern: GeoCtrl was rated lower than both HandLevel ($p < .001$) and Physics ($p < .001$), with no significant difference between HandLevel and Physics. No significant difference was found between HandLevel and GeoCtrl on any other subscale.

For the post-condition questionnaire, we report mean per condition with Friedman omnibus statistics, along with Wilcoxon signed-rank post-hoc comparisons with Bonferroni correction (Table 2). $*$ marks conditions rated significantly higher than Physics. GeoCtrl and HandLevel did not differ on any measure. The VEQ agency subscale showed no significant effect of condition ($\chi^2(2) = 2.264, p = .322$), and its four constituent items were individually non-significant as well. Condition significantly affected all five custom items. Predictability was rated higher in both HandLevel and GeoCtrl than in Physics (both $p < .05$), with no significant difference between HandLevel and GeoCtrl. Realism and Physical Comfort were higher in GeoCtrl than in Physics (both $p < .05$), the only significant contrasts for these two items. Understandability and Familiarization were higher in HandLevel than in Physics (both $p < .01$). GeoCtrl and HandLevel did not differ significantly on any item.

Table 1: NASA-TLX omnibus LRT results ($df = 2$).

Subscale	χ^2	p	R^2_m	R^2_c
Raw TLX	25.402	$< .001$	0.169	0.621
Mental Demand	10.432	.005	0.078	0.500
Physical Demand	23.254	$< .001$	0.186	0.532
Temporal Demand	17.420	$< .001$	0.125	0.554
Performance	17.361	$< .001$	0.150	0.460
Effort	22.334	$< .001$	0.159	0.579
Frustration	20.022	$< .001$	0.170	0.487

GeoCtrl was rated the most preferred technique (67%, versus 29% for HandLevel and 4% for Physics) and the least preferred by 17% (HandLevel 8%, Physics 75%). For precision, HandLevel and GeoCtrl tied as most precise (46% each), while Physics was rated least precise by 71%.

5.4 Qualitative Feedback

Open-ended feedback, collected with the perceived precision and preference, was analyzed following Reflexive Thematic Analysis [6]. A Large Language Model (LLM) applied a researcher-

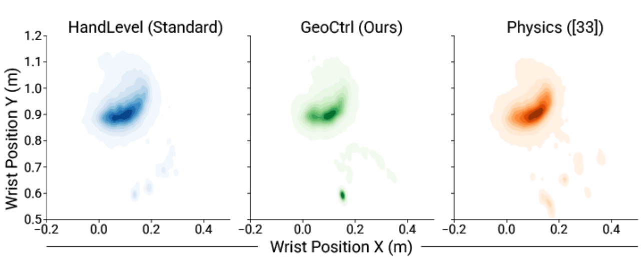

Figure 9: Spatial distribution of wrist position during Study 2, aggregated across participants per condition. The heatmap reflects the range and extent of wrist-level movement required under each technique. All participants were right-handed.

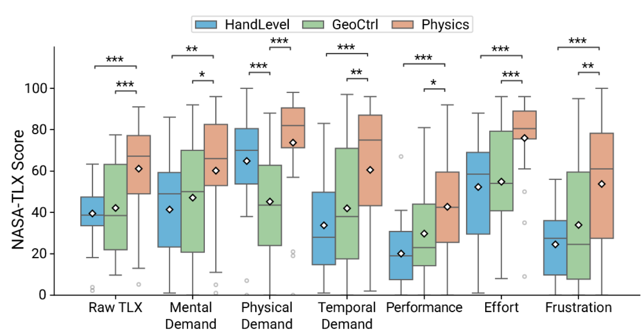

Figure 10: NASA-TLX workload scores across the three conditions in Study 2.

Table 2: 7-point Likert questionnaire results ($df = 2$).

Measure	HandLevel	GeoCtrl	Physics	χ^2	p
VEQ Agency	4.81	5.25	4.56	2.264	.322
Predictability	5.00∗	4.54∗	3.17	15.407	< .001
Realism	3.96	4.67∗	3.25	8.237	.016
Physical Comfort	3.62	4.67∗	2.54	13.153	.001
Understandability	6.00∗	5.04	4.58	9.270	.010
Familiarization	5.71∗	5.38	4.50	10.364	.006

defined codebook of 23 codes across seven groups, after which a researcher reviewed and corrected all assignments.

The codes were mapped onto higher-order themes, iteratively revised, and cross-checked independently by three researchers. Four themes emerged, Intuitiveness and Learnability, Paradox of Physics-based Interaction, Physical Load and Movement, and Finger-Centric Manipulation Strategy, each described below.

6 Discussion

6.1 Mental Model Alignment in Enacting-based Interaction

Our initial assumption was that in the absence of physical support, users will gesturally enact an object's intended rotation with their fingers [1, 43]. This assumption was broadly supported by participant behavior. Although we did not directly measure enacting behavior, the experimental outcomes suggest that GeoCtrl was understood and adopted with minimal instruction. Participants required only a brief explanation of the grab logic, after which they could successfully perform the manipulation tasks. Notably, among 24 participants, we did not detect any participant-level outlier by task completion time under the GeoCtrl condition using IQR (Interquartile Range) outlier detection ($Q_1 = 8.260, Q_3 = 12.645$).

Introducing finger-level control might raise a concern that such dexterous manipulation would impose additional cognitive or physical burden. Despite the added complexity of a high degree of freedom, GeoCtrl did not increase workload compared to the straightforward HandLevel condition. None of the NASA-TLX subscales nor the overall score revealed a significant difference between the two conditions (Fig.10). This suggests that the geometric mapping mechanism eases the complexity of finger coordination, yielding a workload comparable to wrist-level manipulation despite the added finger-level DOF. Accordingly, GeoCtrl was preferred by 67 % of participants over the alternatives.

Post-study qualitative feedback reinforced this finding. Among participants who preferred GeoCtrl, 10 reported that they could "effortlessly" complete tasks using the proposed interaction. Three participants explicitly noted that GeoCtrl employed "familiar" gestures, with two of them stating that GeoCtrl required hand movements most similar to real-world object manipulation (P11: *"the way of using my hands resembled rotating an object in reality,"* P5: *"most similar to rotating a small object in the real world, so it felt familiar."*). The post-condition questionnaire corroborated this, with GeoCtrl rated comparably to HandLevel on Predictability, Understandability, and Familiarization, and above Physics on Predictability (Table 2). These observations suggest that the geometric mapping mechanism aligns with users' mental models of in-hand rotation, and that GeoCtrl remained approachable across participants with varying XR experience.

6.2 The Paradox of Physics-based Interaction

Even though physics-based simulation approaches the exact replication of reality, the loss of human dexterity in mid-air conditions cannot be easily compensated by simulation fidelity alone. Our results reveal a fundamental paradox that realistic physics simulation does not necessarily yield a satisfactory experience because of changes in user behavior. Under the Physics condition, we observed frequent jitter, involuntary contact loss, and object instability throughout trials. This inconsistency translated directly into significantly higher task load (Fig.10).

The difficulty in maintaining a stable grasp was further reflected in the time-on-target metric. Participants spent an excessive amount of trial time recovering from jitter and re-establishing contact rather than performing intentional rotation. While Physics recorded smaller total die transformations than GeoCtrl (Fig.8(c–d)), we argue this reflects interaction instability rather than efficiency, as confirmed by the time metrics (Fig.8(a–b)). Physics also produced a wider workspace area and higher KDE entropy than GeoCtrl (Fig.9), indicating that participants were not executing finger-level rotation but primarily relied on wrist movement in response to grab failures. In contrast, GeoCtrl users maintained a more fixed wrist position, presumably performing stable rotation primarily through finger movements.

Aligned with higher workload, some participants expressed visible frustration with Physics. Qualitative responses described the object's behavior as *"unexpected"* (P20) and *"different from what I know in reality"* (P2), creating *"the greatest gap from real-world"* (P13). This discrepancy is particularly damaging for physics simulation, as participants struggled to perform the finger-level movements the engine demanded. Interestingly, rather than attempting to grab the object, several participants resorted to pushing or nudging the die to complete the task. This highlights that the fidelity offered by physics simulation is undermined when the haptic feedback and physical support of a real object are absent.

These difficulties reflect the interaction itself rather than measurement error. Because all three conditions used the same hand-tracking system, tracking quality cannot account for the differences between conditions. The behavioral change instead follows from the absence of tactile cues and contact forces that guide real-world manipulation, which remain missing in VR regardless of haptic hardware [9]. Notably, participants rated GeoCtrl as more realistic than Physics (Table 2), even though only Physics simulates real physics, underscoring that perceived realism depends on usable interaction rather than simulation fidelity alone.

6.3 Physical Demand and Loss of Dexterous Control in Hand-level Interaction

The simplicity and robustness of the commodity approach came at the cost of higher physical demand and reduced dexterous control. Since HandLevel is insensitive to finger movements after grabbing, rotation was necessarily performed at the wrist or arm level, requiring larger-amplitude body movements. This was reflected in higher wrist transformation and die translation under HandLevel (Fig.8(d)) and in the wide spatial spread of wrist movements (Fig.9).

Participants rotated the object less with HandLevel (Fig.8(c)) because rotating it further with restricted movement demanded more physical effort. Subjective responses confirmed this burden, with 8 participants reporting that HandLevel required large body movements and induced substantial physical fatigue (P16: *"caused significant fatigue,"* P9: *"my wrist hurt more."*). The NASA-TLX physical demand subscale was significantly higher under HandLevel compared to GeoCtrl (Fig.10). In contrast, GeoCtrl enabled rotation primarily through small finger movements, which participants described as physically comfortable.

Beyond physical comfort, we considered whether finger-level control preserved the user's sense of agency, a potential cost of mediating manipulation through the CD gain rather than direct motion. The VEQ agency subscale revealed no significant difference across conditions, and GeoCtrl was rated numerically highest (Table 2). This indicates that GeoCtrl did not sacrifice agency relative to either reference, even though its geometric mapping breaks 1:1 correspondence. Consistent with this, one participant noted a stronger sense of synchronization when the fingers played a larger role (P19: *"when fingers played a stronger role, it seemed to provide a greater sense of synchronization."*). Under HandLevel, where the visual hand locks onto the object upon grab and decouples from finger motion, no participant spontaneously mentioned agency.

Taken together, GeoCtrl provides composite benefits of the two compared approaches. Physics pursues high-fidelity realism but fails to deliver it in practice, in the absence of haptic feedback and physical support. HandLevel is robust and accessible but sacrifices dexterous control by locking finger movements after the grab. GeoCtrl balances these tradeoffs, enabling finger-level dexterous interaction without the instability of physics simulation or the physical burden of wrist-level manipulation, at the cost of longer task completion time than HandLevel (P6: *"it felt like a combination of the advantages of the other ones."*). The geometric mapping accommodates the coarse gestural movements of users by eliminating the need for exact motor precision while still engaging the fine motor capabilities.

6.4 Finger-centric Control Sensitivity and CD Gain

One consistent theme across both Physics and GeoCtrl was that participants found finger-level movements to be overly sensitive. While participants appreciated the advantages of GeoCtrl's finger-centric control for fast object inspection (P24: *"I could explore the die more quickly,"* P6: *"I first quickly explored with fingers."*), several reported experiencing unintended rotations triggered by finger motion. This sensitivity is an inherent trade-off, as the same responsiveness that enables dexterous control also amplifies small, unintended movements, pointing to the necessity of CD gain.

Our initial assumption was that users would require amplified input to perform large-scale ballistic rotations. However, the more critical function of the CD gain in GeoCtrl is to diminish fine finger movements rather than amplify them since wrist-level movement already provides sufficient ballistic input. In practice, the gain's dampening behavior for small-displacement inputs was more consequential to usability than its amplification at large displacements. A more carefully tuned gain curve, with stronger damping in the low-displacement regime, would reduce unintended rotations and make GeoCtrl more suitable for precise manipulation tasks.

The importance of a "freeze" control mechanism was also affirmed by participants (P24: *"I liked being able to freely alternate between moving and fixing,"* P23: *"when unable to fix the die or my hand, the manipulation felt uncomfortable."*). As fine finger movements can induce both efficient rotation and inadvertent transformation, the capacity for deliberate pause helps users maintain intentional control. Thus, refined CD gain application and an improved clutch mechanism will form a complementary pair to support more precise manipulation with GeoCtrl.

Because GeoCtrl combines the geometric mapping and CD gain with ancillary mechanisms, object positioning, grab-release, clutching, and triangle-area thresholding, their individual contributions cannot be fully separated in our study. Our observations indicate that these ancillary mechanisms primarily refine stability and precision, and a formal ablation isolating each component remains an important direction for future work.

6.5 Limitation and Future Work

GeoCtrl also revealed shortcomings that shaped user preference. Although it was the most preferred technique overall (67%), it was the least preferred for 17% of participants and was judged most precise about as often as HandLevel, not more. Open-ended responses attribute this to the finger-level sensitivity noted above, where small unintended finger motions produced unwanted rotations. For some participants, this unreliability outweighed the technique's efficiency. Improving usability therefore depends on damping these motions, which more careful gain tuning could achieve. The clearer boundary is task speed, since GeoCtrl's longer completion time makes it less attractive when raw speed dominates. GeoCtrl thus offers the greatest advantage in tasks that prioritize low-effort, fine-grained rotation over speed.

The current study used a simple docking task with a single cube-shaped object and fixed success thresholds, limiting generalizability. Future work should broaden the task parameters, object geometries, and complexity, since object shape may engage the fingertip-triangle mapping in ways a cube does not reveal. The CD gain space also warrants fuller exploration, as the optimal curve may differ across precision requirements and further tuning could mitigate the sensitivity issues participants reported.

Although our implementation included grab and clutching logic, a more thorough investigation of these features could substantially improve the usability of GeoCtrl. Integration with a physics engine for grabbing may further enhance stability while retaining the benefits of geometric mapping. Developing adequate visual and pseudo-haptic augmentation could also enrich the interaction experience. Finally, exploring concrete application scenarios such as assembly tasks or immersive education would further validate the generalizability of GeoCtrl.

7 Conclusion

Dexterous bare-hand rotation of virtual objects remains challenging in XR, where the haptic cues and physical support of a real object are absent. We presented GeoCtrl, a technique that maps geometric changes of finger configurations to the rotation of a virtual object. Our approach accommodates coarse gestural approximations rather than demanding the motor precision assumed by physics-based approaches. A speed-responsive CD gain, empirically optimized through a controlled study, further enhances efficiency and controllability of the interaction.

Relative to the high-fidelity physics simulation, GeoCtrl improved task efficiency and provided more effective rotational control in the absence of haptic feedback. It imposed the lowest physical demand of the three techniques and added no workload over the commodity condition, though it did not match that condition's raw completion speed. The study also highlighted that the primary role of the CD gain is damping small finger movements rather than amplifying large-scale input, and that the clutch mechanism was essential for maintaining intentional control.

Our work has demonstrated the potential of geometry-driven manipulation as a practical foundation for dexterous hand interaction in XR, particularly where low-effort, fine-grained rotation matters more than raw task speed. We believe that GeoCtrl inspires future interaction techniques to prioritize gestural intent over motor precision and physical replication, enabling dexterous hand interaction in XR environments with growing complexity.

Disclosure on Gen AI Usage

In accordance with IEEE guidelines, the authors used Claude Code for linguistic editing and to debug data-analysis scripts. All intellectual contributions remain the sole responsibility of the authors.

Acknowledgments

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. RS-2024-00397663, Real-time XR Interface Technology Development for Environmental Adaptation; No. IITP-2026-RS-2024-00436398, ITRC(Information Technology Research Center); No. IITP-2022(2026)-RS-2022-00156435, Virtual Convergence Support Program to Nurture the Best Talents).

REFERENCES

[1] R. Arora, R. H. Kazi, D. M. Kaufman, W. Li, and K. Singh. Magicalhands: Mid-air hand gestures for animating in vr. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, pp. 463–477. ACM, 2019. doi: 10.1145/3332165.3347942 1, 8

[2] R. Balakrishnan and I. S. MacKenzie. Performance differences in the fingers, wrist, and forearm in computer input control. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, pp. 303–310, 1997. 4

[3] J. Bergström, T.-S. Dalsgaard, J. Alexander, and K. Hornbæk. How to evaluate object selection and manipulation in vr? guidelines from 20 years of studies. In *proceedings of the 2021 CHI conference on human factors in computing systems*, pp. 1–20, 2021. 4

[4] L. Besançon, P. Issartel, M. Ammi, and T. Isenberg. Mouse, tactile, and tangible input for 3d manipulation. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 4727–4740, 2017. 2

[5] R. Blanch, Y. Guiard, and M. Beaudouin-Lafon. Semantic pointing: improving target acquisition with control-display ratio adaptation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 519–526, 2004. 2

[6] V. Braun and V. Clarke. Reflecting on reflexive thematic analysis. *Qualitative research in sport, exercise and health*, 11(4):589–597, 2019. 7

[7] G. Casiez, N. Roussel, and D. Vogel. 1€ filter: a simple speed-based low-pass filter for noisy input in interactive systems. In *Proceedings of the SIGCHI Conference on human factors in computing systems*, pp. 2527–2530, 2012. 3

[8] G. Casiez, D. Vogel, R. Balakrishnan, and A. Cockburn. The impact of control-display gain on user performance in pointing tasks. *Human–computer interaction*, 23(3):215–250, 2008. 2, 3, 4

[9] M. Chessa, G. Maiello, L. K. Klein, V. C. Paulun, and F. Solari. Grasping objects in immersive virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1749–1754. IEEE, 2019. 1, 2, 8

[10] 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. 2

[11] D. Englmeier, J. Dorner, A. Butz, and T. Hollerer. A tangible spherical proxy for object manipulation in augmented reality. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 221–229. IEEE, 2020. doi: 10.1109/vr46266.2020.00041 2

[12] X. Fan, X. Huang, and R. Xiao. Tangiar: Markerless tangible input for immersive augmented reality with everyday objects. In *Proceedings of the 2025 31st ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11. ACM, Dec 2025. doi: 10.1145/3756884.3766028 2

[13] C. Forlines, D. Vogel, and R. Balakrishnan. Hybridpointing: fluid switching between absolute and relative pointing with a direct input device. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pp. 211–220, 2006. 2, 3

[14] S. Frees and G. Kessler. Precise and rapid interaction through scaled manipulation in immersive virtual environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, pp. 99–106. IEEE, 2005. doi: 10.1109/vr.2005.1492759 2, 3

[15] S. Frees, G. D. Kessler, and E. Kay. Prism interaction for enhancing control in immersive virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 14(1):2–es, 2007. 2, 3, 4, 5

[16] P. C. Gloumeau, W. Stuerzlinger, and J. Han. Pinnpivot: Object manipulation using pins in immersive virtual environments. *IEEE transactions on visualization and computer graphics*, 27(4):2488–2494, 2020. 2

[17] M. Hancock, S. Carpendale, and A. Cockburn. Shallow-depth 3d interaction: design and evaluation of one-, two- and three-touch techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1147–1156. ACM, New York, NY, USA, 2007. doi: 10.1145/1240624.1240798 2

[18] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988. 5

[19] O. Hilliges, D. Kim, S. Izadi, M. Weiss, and A. Wilson. Holodesk: direct 3d interactions with a situated see-through display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2421–2430, 2012. 1, 2

[20] K. Hinckley, J. Tullio, R. Pausch, D. Proffitt, and N. Kassell. Usability analysis of 3d rotation techniques. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pp. 1–10, 1997. 2

[21] K. Hirota and K. Tagawa. Interaction with virtual object using deformable hand. In *2016 IEEE Virtual Reality (VR)*, pp. 49–56. IEEE, 2016. 2

[22] 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. 1, 2

[23] J. Jacobs and B. Froehlich. A soft hand model for physically-based manipulation of virtual objects. In *2011 IEEE virtual reality conference*, pp. 11–18. IEEE, 2011. 2

[24] H. Jiang, D. Weng, Z. Song, X. Dongye, and Z. Zhang. Dexhand: dexterous hand manipulation motion synthesis for virtual reality. *Virtual Reality*, 27(3):2341–2356, 2023. 2

[25] J. Kim, S. Park, Q. Zhou, M. Gonzalez-Franco, J. Lee, and K. Pfeuffer. Pinchcatcher: Enabling multi-selection for gaze+ pinch. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2025. 2, 3

[26] J. Kim, Y. Zhang, and S. H. Yoon. T2iray: Design of thumb-to-index based indirect pointing for continuous and robust ar/vr input. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2025. 2

[27] S. Kim and G. Lee. Evaluating an in-hand ball-shaped controller for object manipulation in virtual reality. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 10–19. IEEE, 2024. 2

[28] S. Kim, Y. A. Shim, and G. Lee. Exploration of form factor and bimanual 3d manipulation performance of rollable in-hand vr controller. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–11. ACM, 2022. doi: 10.1145/3562939.3565625 2

[29] K. Kin, C. Wan, K. Koh, A. Marin, N. C. Camgöz, Y. Zhang, Y. Cai, F. Kovalev, M. Ben-Zacharia, S. Hoople, et al. Stmg: A machine learning microgesture recognition system for supporting thumb-based vr/ar input. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pp. 1–15, 2024. 2

[30] W. A. König, J. Gerken, S. Dierdorf, and H. Reiterer. Adaptive pointing–design and evaluation of a precision enhancing technique for absolute pointing devices. In *IFIP Conference on Human-Computer Interaction*, pp. 658–671. Springer, 2009. 2, 3

[31] J. Lee, M. Sinclair, M. Gonzalez-Franco, E. Ofek, and C. Holz. Torc: A virtual reality controller for in-hand high-dexterity finger interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13. ACM, 2019. doi: 10.1145/3290605.3300301 2, 4

[32] C. Lim, T. Kim, and G. Lee. Twinspin: A virtual ball in a vr controller enabling in-hand 3dof rotation. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*, pp. 1–13, 2025. 2

[33] C. Lougiakis, J. J. González, G. Ganias, A. Katifori, M. Roussou, et al. Comparing physics-based hand interaction in virtual reality: custom soft body simulation vs. off-the-shelf integrated solution. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 743–753. IEEE, 2024. 1, 2, 6

[34] C. L. MacKenzie and T. Iberall. *The grasping hand*, vol. 104. Elsevier, 1994. 3

[35] A. Martinet, G. Casiez, and L. Grisoni. Integrality and separability of multitouch interaction techniques in 3d manipulation tasks. *IEEE Transactions on Visualization and Computer Graphics*, 18(3):369–380, 2012. doi: 10.1109/TVCG.2011.129 2

[36] D. Mendes, M. Sousa, R. Lorena, A. Ferreira, and J. Jorge. Using

custom transformation axes for mid-air manipulation of 3d virtual objects. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, pp. 1–8, 2017. 2

[37] T. Mikkelsen, Q. Zhou, M. N. Lystbæk, Y. Liu, H. Gellersen, and K. Pfeuffer. Dof-separation for 3d manipulation in xr: Understanding finger-wrist separation to simultaneously translate and rotate objects. In *2025 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1246–1256. IEEE, 2025. 2

[38] S. Nakagawa and H. Schielzeth. A general and simple method for obtaining r2 from generalized linear mixed-effects models. *Methods in ecology and evolution*, 4(2):133–142, 2013. 5

[39] M. Nancel, O. Chapuis, E. Pietriga, X.-D. Yang, P. P. Irani, and M. Beaudouin-Lafon. High-precision pointing on large wall displays using small handheld devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 831–840, 2013. 2, 3

[40] M. Nancel, E. Pietriga, O. Chapuis, and M. Beaudouin-Lafon. Mid-air pointing on ultra-walls. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 22(5):1–62, 2015. 2, 4

[41] E. Normand, C. Pacchierotti, E. Marchand, and M. Marchal. Visuo-haptic rendering of the hand during 3d manipulation in augmented reality. *IEEE Transactions on Haptics*, 17(2):277–291, 2024. 2

[42] S. Y. Oh, J. Seo, J. Lee, B. Yoon, S. H. Yoon, and W. Woo. Forcectrl: Hand-raycasting with user-defined pinch force for control-display gain application. *IEEE Transactions on Visualization and Computer Graphics*, 32(3):2529–2540, 2026. doi: 10.1109/tvcg.2025.3647547 2, 3

[43] S. Pei, A. Chen, J. Lee, and Y. Zhang. Hand interfaces: Using hands to imitate objects in ar/vr for expressive interactions. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–16. ACM, 2022. doi: 10.1145/3491102.3501898 1, 8

[44] G. Perelman, M. Serrano, M. Raynal, C. Picard, M. Derras, and E. Dubois. The roly-poly mouse: Designing a rolling input device unifying 2d and 3d interaction. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, pp. 327–336, 2015. 2

[45] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: non-linear mapping for direct manipulation in vr. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pp. 79–80, 1996. 2, 3

[46] I. Poupyrev, S. Weghorst, and S. Fels. Non-isomorphic 3d rotational techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 540–547, 2000. 2, 3

[47] I. Poupyrev, S. Weghorst, T. Otsuka, and T. Ichikawa. Amplifying spatial rotations in 3d interfaces. In *CHI'99 Extended Abstracts on Human Factors in Computing Systems*, pp. 256–257, 1999. 2, 3

[48] J. L. Reisman, P. L. Davidson, and J. Y. Han. A screen-space formulation for 2d and 3d direct manipulation. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology*, pp. 69–78. ACM, New York, NY, USA, 2009. doi: 10.1145/1622176.1622190 2

[49] F. Rodrigues, A. Giovannelli, L. Pavanatto, H. Miao, J. C. de Oliveira, and D. A. Bowman. Amp-it and wisdom: Improving 3d manipulation for high-precision tasks in virtual reality. In *2023 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 303–311. IEEE, 2023. 2, 3

[50] D. Roth and M. E. Latoschik. Construction of the virtual embodiment questionnaire (veq). *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3546–3556, 2020. doi: 10.1109/TVCG.2020.3023603 6

[51] B. Schmeier, J. P. Kopetz, B. Kordts, and N. Jochems. Manipulating virtual objects in augmented reality using a new ball-shaped input device. In *12th Augmented Human International Conference*, pp. 1–8. ACM, 2021. doi: 10.1145/3460881.3460935 2

[52] S. Sridhar, A. M. Feit, C. Theobalt, and A. Oulasvirta. Investigating the dexterity of multi-finger input for mid-air text entry. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 3643–3652, 2015. 4

[53] A. Talvas, M. Marchal, C. Duriez, and M. A. Otaduy. Aggregate constraints for virtual manipulation with soft fingers. *IEEE transactions on visualization and computer graphics*, 21(4):452–461, 2015. 1, 2

[54] H. Tian, C. Wang, D. Manocha, and X. Zhang. Realtime hand-object interaction using learned grasp space for virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2623–2635, 2018. 2

[55] 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. 2

[56] H. Wang, L. Sidenmark, F. Weidner, J. Newn, and H. Gellersen. Headshift: Head pointing with dynamic control-display gain. *ACM Transactions on Computer-Human Interaction*, 32(1):1–28, 2025. 2, 3

[57] C. Wilkes and D. A. Bowman. Advantages of velocity-based scaling for distant 3d manipulation. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, pp. 23–29, 2008. 3

[58] S. Zhai, P. Milgram, and W. Buxton. The influence of muscle groups on performance of multiple degree-of-freedom input. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 308–315, 1996. 3

[59] F. Zhu, L. Sidenmark, M. Sousa, and T. Grossman. Pinchlens: Applying spatial magnification and adaptive control-display gain for precise selection in virtual reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1221–1230. IEEE, 2023. doi: 10.1109/ismar59233.2023.00139 2, 3