An Immersive Interface for Colonoscopy Training

Category: Research

Abstract

Extended Reality technologies have the potential to enhance surgical training by providing immersive and controllable simulation environments. However, achieving realistic, natural interaction with complex instruments, such as endoscopes, remains a technical challenge when limited haptic feedback is available. This work presents a virtual reality interface for endoscope manipulation using off-the-shelf hardware, integrating prior research on deformable colon and endoscope simulation. Our system enables natural control and real-time interaction within an immersive virtual environment. A preliminary assessment by medical specialists provided positive feedback on usability and interaction realism, suggesting the feasibility of the proposed simulator. Ongoing work focuses on conducting formal validation.

Index Terms: Virtual Reality, Endoscopy Simulation, Natural Interaction

1 Introduction

Colonoscopy holds significant importance in the detection of various diseases, with a particular emphasis on colorectal cancer (CRC), the second leading cause of cancer-related deaths. Early detection of polyps, adenomas, and CRCs plays a significant role in preventing and lowering mortality rates in Colorectal Cancer (CRC), and colonoscopy is a gold standard technique for detecting and eliminating these polyps [3]. This method is particularly effective, leading to a 67% decrease in the risk of death from any colorectal cancer [4].

However, mastering colonoscopy requires extensive training, and Virtual Reality (VR) simulators have proven to enhance performance in Patient-based Assessments [7]. Besides physical simulators, virtual simulators are crucial for training novice surgeons and for improving patients' comfort and safety. Additionally, VR simulators accelerate students' training in the early stages of practical learning by enabling repeated practice of procedures [12].

The effectiveness of virtual simulators frequently rivals or surpasses that of traditional training methods [15]. The pivotal factor influencing the success of these virtual simulators is the mechanical model used to simulate both the colonoscopy instrument and the colon. The colon can be conceptualized as soft tissue, while the instrument setup is simplified as a flexible tube controlled by a tensioned steel rod within. The primary research focus has been on simulation models to replicate the effects of those key characteristics; however, none have proposed a virtual reality interface for the instruments, which is key to practical use.

In this paper, we present a virtual reality interface and visualization setup replicating an operating room environment. The challenges are manifold, primarily involving how to correctly grasp and manipulate the endoscope, how to compensate for the lack of tactile feedback that surgeons rely on to control the instrument, and how to preserve similarity to real procedural manipulation techniques so that the system maintains instructional value.

The remainder of this paper is organized as follows. In Section 2 we review some related works in the field. Section 3 overviews how we have modeled and simulated the colon and the endoscope, while Section 4 details the virtual reality interface we conceived to allow users to interact with the endoscope. Section 5 presents our preliminary results while informally testing with specialists and Section 6 point out our thoughts and on-going work.

2 Related Work

Traditional colonoscopy simulation relies on physical models [13] made from different materials like latex and other synthetic materials. These models are considered affordable as they can be extensively reused. However, they may come in a wide range of prices and fidelities [5]. *Ex vivo* models made from animal tissue come at a higher cost and fidelity. Training using live animals, generally pigs, is costly and raises ethical concerns, but provides higher fidelity [9].

Virtual simulators, in turn, consist of a hardware component and software that enable the reproduction of different scenarios. The software component simulates soft tissues and rigid bodies. It allows training ranging from basic skills to advanced procedures and can provide objective performance metrics. The hardware part usually requires a real endoscope to be inserted into a mechanical device that reads the movements and may provide force resistance. Although they do not offer the same tactile feedback as physical simulators, they allow for greater case variability. However, they tend to be expensive as they rely on proprietary hardware, which limits their widespread adoption in educational centers. Also, their soft tissues models are extremely simplified, built on game-like components that do not reflect the complexity of real tissue interaction with surgical tools. This choice permits them to run on more affordable computers, but negatively affects their visual realism and that of the forces acting in the colon and instruments [8]. An example of these simulators is the GI Mentor II, by Surgical Science.

Only in recent years have advances in the field of computer science made it possible to use immersive virtual reality for creating simulators [13], which have been used in the early stages of training and have shown a considerable impact on the performance of the trainees and training time [11, 14]. These alternatives are cost-efficient and potentially close to reality, well supporting surgeons practicing and training.

3 Endoscope and Colon Models

Before we present our immersive interface for virtual colonoscopy in Section 4, we briefly summarize here the results of a continuous effort we made in the past years to simulate the colon and endoscope complexity in real time with high realism.

The simulation comprises a pipeline (see Fig. 1) with its key stages, starting with the acquisition of the colon mesh from patient data. The mesh is reconstructed from computerized tomography (CT) images, usually obtained from patients undergoing colonography. Automatic segmentation is possible due to colon insufflation and the dye contrast used in the preparation for the scan. The colon centerline is then extracted and will be used to configure a spline skeleton for the colon. Our endoscope model is also based on a spline skeleton, which allows for a targeted collision detection approach between the colon and the scope tube. Finally, to process the physical deformation of these elements, we propose an XPBD formulation.

Our processing pipeline begins with obtaining a triangle mesh of the colon surface. This can be a reproduction made by a designer anatomist or a reconstruction from medical imaging. While other medical imaging techniques could also provide anatomical data [6], we consider colonography CT data ideal. We followed an approach using open tools [10], capitalizing on The Cancer Imag-

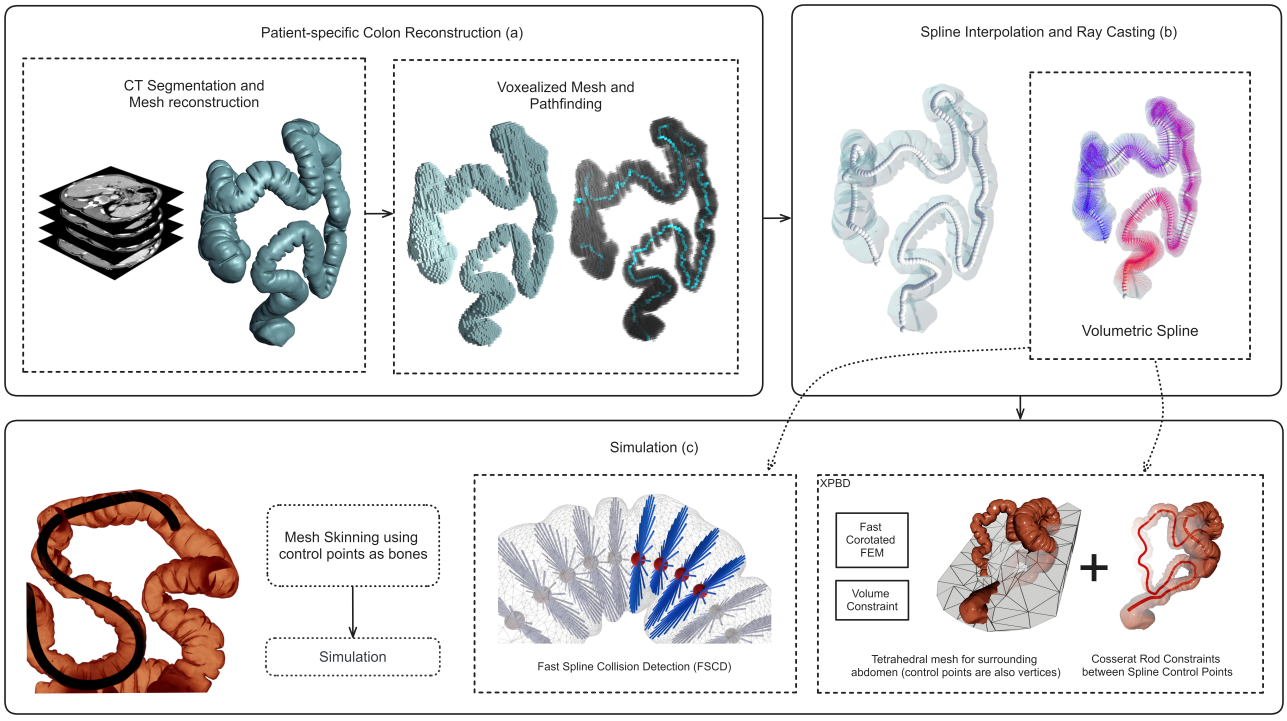

Figure 1: The simulation pipeline is divided into three main parts, starting with the **acquisition and reconstruction** of the colon from a real patient CT (a). Then, the colon centerline is extracted and used to **build a spline skeleton** (b), which will be used to model the colon and the endoscope. The physical deformation is produced using an XPBD formulation and a tailored collision detection method during the **simulation** (c).

ing Archive[1].

Given the surface mesh, we enhance the model with structures necessary to deform the colon. The next step is then to extract the centerline of the colon from the rectum to the cecum. Points on the centerline will be subsequently interpolated to generate a spline that will serve as a backbone for the colon shape. This spline provides a continuous path throughout the colon but has no volume. While the mesh provides volume, it is overly complex to process the detection of multiple collisions while deforming the mesh. We thus propose to enrich the spline with volumetric information, creating a structure we call *volumetric spline*, or simply *volspline*.

Besides the colon, the endoscope is central to colonoscopy. The endoscope is a cylindrical, inextensible, flexible tube carrying the camera's eye and light source. We also modeled the endoscope as a Catmull-Rom *volspline*, with a set of Radiating Control Points (RCPs). An endoscope is simpler to model than the colon. It also does not need to be reconstructed, as its cylindrical shape can be modeled procedurally.

The Fast Spline Collision Detection (FSCD) algorithm is tailored to detect multiple intersections between two deformable tubular structures that permanently touch while moving inside each other, such as an endoscope inside the colon. The method identifies collisions by comparing two volumetric splines. We thus have an inner and an outer *volspline*. At the end, it returns a list of collision tuples used later for collision response.

As the inner *volspline* has a uniform radius all along, we can see each RCP as a sphere. The precision of the inner *volspline* is adjusted by the number of samples, where regions more prone to collision take more samples. For the outer *volspline*, we cannot make that assumption. Therefore, we build trianguloids grouping each radial vector with its adjacent. This reduces the problem to the intersection check between the sphere and the side of the trianguloid representing the outside mesh wall.

An interactive simulation of a colon for a colonoscopy simulation demands a physically plausible response and real-time performance despite the nonlinear complexity of the problem. We employed XPBD and a novel arrangement of constraints to simulate the colon and endoscope deformations [2].

For more details about our colon and endoscope modeling and simulation we recommend reading our previous work [1].

4 Immersive VR Interface

Considering the colon and endoscope models presented in Section 3, we developed a virtual reality interface to provide immersive and realistic interaction with the virtual endoscope. Our interface requires hand tracking as well as simultaneous support for hand and controller input. In the current implementation, we used the Oculus Quest 3 as the head-mounted display, the Unity engine, and the Meta SDK.

4.1 Design Choices

The virtual environment was modeled to resemble a simplified operating room. The main instruments include the intra-colon vision monitor and a mini-map–style visualization providing an overview of the colon and the endoscope (see Fig. 2). Below these monitors, two text-based hint interfaces are displayed. One provides warnings related to the current training procedure, while the other indicates the endoscope's handling state. On the left side of the intra-colon vision monitor, a vertical bar displays the magnitude of the forces applied by the endoscope against the colon walls. This visualization is enabled by the endoscope–colon simulation techniques described previously. Additionally, above the patient, a text box presents instructions relevant to the active training scenario.

[1] https://www.cancerimagingarchive.net

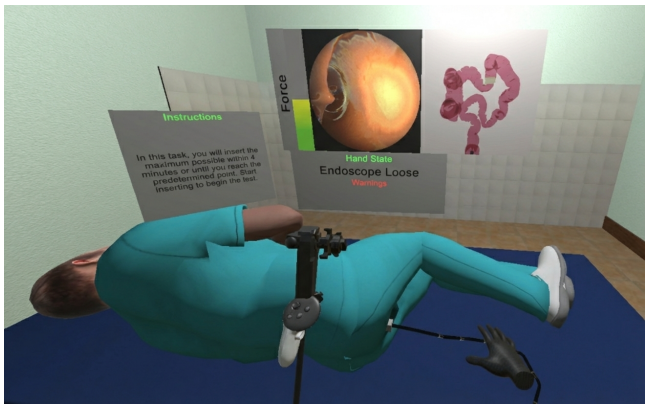

Figure 2: The virtual environment is organized with the intra-colon view displayed centrally, accompanied by the colon minimap, a force-feedback bar, and a text-based hint interface. In the foreground, a virtual operating table is positioned near the user, with the patient model and the inserted endoscope visible. The control head is attached to the user's handheld controller, and an instructional panel is displayed above the patient.

Figure 3: Virtual Endoscope

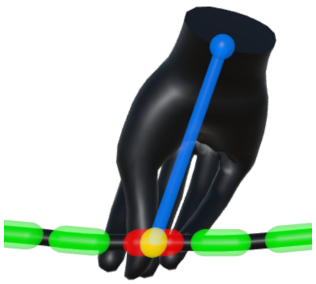

Figure 4: The user's hand interacting with the endoscope. The grabbed segment is highlighted in red. The blue marker denotes the hand's reference origin, and the yellow marker indicates the contact position on the segment. The blue line represents the displacement between these two points

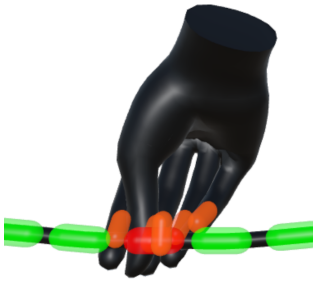

Figure 5: The user's hand sliding in the endoscope

Figure 6: Endoscope released

To reduce computational cost and avoid complex physical constraints at the interface between the user's hand and the endoscope, we chose not to use the same physically accurate endoscope model in the external environment. Poor user performance may lead to unstable simulation conditions, such as large time steps or excessive hand velocity. These conditions can introduce abrupt velocity changes at a given segment of the endoscope, which are propagated along the articulated chain by the physics solver, potentially leading to segment interpenetration and simulation breakdown. In such cases, the simulation would require a restart.

Therefore, we simplified the external endoscope representation and employed a less restrictive physics solver in that environment. Since the external model primarily serves as a visual guide to support hand gestures, full physical accuracy was not required.

4.2 Endoscope Modeling

The virtual endoscope manipulated by the user is illustrated in Fig. 3. It is procedurally generated: the shaft consists of a chain of linked segments, each implemented as a capsule collider connected by a limited-angle joint to the adjacent segment. The Unity physics engine manages and resolves the interactions among these components, without enforcing the more precise and constrained physical behavior required for the internal colon simulation.

4.3 Interaction Model for Endoscope Manipulation

We designed a three-state interaction model to emulate real-world colonoscope handling while remaining compatible with optical hand tracking in VR. The model comprises a **grabbing state**, a **sliding state**, and a **release state**, reflecting clinically observed manipulation strategies.

In the **grabbing state** (Fig. 4), the three fingers are in full contact, forming a closed pinch configuration. When a tube segment is detected at the fingertips, that segment becomes attached to the virtual hand. All manipulation inputs, including insertion, withdrawal, torque, and angulation, are applied directly to the currently grasped segment. This state enables direct mechanical control of the endoscope and corresponds to active instrument actuation in clinical practice.

In the **sliding state** (Fig. 5), the hand maintains a pinch configuration with slight separation between the fingers. This allows the user to reposition the hand along the endoscope shaft without generating control inputs or releasing the instrument. During sliding, the actively controlled segment is dynamically updated according to the hand's position along the tube. This mechanism enables continuous repositioning analogous to real-world handling, where clinicians frequently shift grip location without disengaging the instrument.

In the **release state** (Fig. 6), the hand is open and no contact with the endoscope is established. Consequently, no interaction forces or control inputs are applied.

This state-based formulation separates grasp acquisition, repositioning, and disengagement, reducing ambiguity in input interpretation while preserving procedural realism.

The gestures associated with the three interaction states were defined based on interviews with experienced surgeons responsible for teaching endoscopic manipulation. This process ensured a close correspondence between the virtual interaction model and real-world clinical practice. Although no universally standardized gesture set exists, and minor variations are commonly observed, particularly during early stages of training, the selected gestures represent commonly taught and widely adopted handling strategies.

Interaction with the control head of the endoscope, where the camera is located, is performed using the contralateral hand, replicating real-world handling. In clinical practice, the control head is typically held in the left hand (see Fig. 2), and our interface preserves this bimanual manipulation paradigm.

The up/down angulation wheel is mapped to the vertical (Y) axis of the joystick, enabling control of vertical tip deflection. By press-

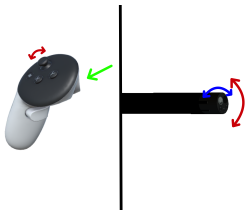

Figure 7: Mapping between the handheld left controller and endoscope tip angulation. Moving the joystick along its vertical axis (red arrow, left image) produces up/down tip deflection (red arrow, right image). When the index trigger (indicated by the green arrow) is pressed, the same joystick input is remapped to control left/right tip angulation (blue arrow, right image).

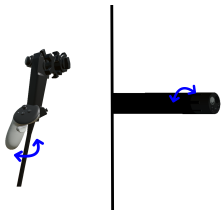

Figure 8: Torque control mapping. Rotating the control head around its longitudinal axis (blue arrow, left image) applies axial torque to the endoscope shaft, resulting in rotation of the distal tip (blue arrow, right image).

ing the index trigger button, the system switches control mode to the left/right angulation wheel (Fig. 7). In this mode, the same joystick input is reassigned to control lateral tip deflection.

Additionally, the user may rotate the virtual control head about its longitudinal axis to apply fine torque to the endoscope shaft (Fig. 8). This rotational input is transmitted incrementally to the simulator, allowing subtle orientation adjustments without interfering with axial insertion control.

This bimanual mapping preserves the functional separation between insertion control and tip articulation, closely reflecting real procedural ergonomics while maintaining an intuitive VR interaction model.

4.4 Compensating for the Absence of Haptic Feedback

A central challenge in this VR-based endoscope simulation is the absence of physical reaction forces. In real procedures, force feedback naturally constrains motion and prevents uncontrolled acceleration. In VR, however, hand motion can introduce unrealistic displacement if not explicitly regulated.

The virtual gesture design must explicitly account for the absence of force feedback, particularly during the grabbing and sliding states. In clinical practice, these actions are typically performed without visual attention to the external shaft of the instrument, relying primarily on tactile cues.

To compensate for the lack of haptic feedback in VR, the grabbing mechanism incorporates a spatial tolerance between the fingertips and the endoscope surface. Without this tolerance, users would be required to visually confirm precise finger placement, which deviates from real-world behavior where tactile feedback naturally guides contact.

The sliding state is even more affected by the absence of haptics. To minimize the discrepancy from clinical handling, virtual sliding is computed based solely on the forward and backward displacement of the user's hand along a linear axis, as if the endoscope were locally approximated by a straight line. However, the visual representation of the virtual hand follows the actual curved geometry of the endoscope shaft, even when it assumes a sinusoid-like configuration.

This controlled decoupling between interaction input and geometric conformity preserves perceptual realism while maintaining stable and intuitive manipulation.

Another critical consideration during the grabbing state concerns the magnitude of the applied force input. In VR, the user's hand moves without physical resistance, allowing unconstrained acceleration. Directly mapping this motion to the simulated endoscope can result in segment overlap, leading to infeasible constraint resolutions within the physics solver.

Moreover, excessive force inputs may reach magnitudes that, in a real clinical scenario, would cause colonic perforation, an event that would immediately terminate the procedure and require emergency surgical intervention. Preventing such unrealistic and pedagogically inappropriate behavior is therefore essential.

To address these issues, we introduce two complementary mechanisms: (1) visual force feedback and (2) force input interpolation. The visual feedback provides real-time information about the applied force magnitude, enabling users to self-regulate their interaction. In parallel, force interpolation smooths abrupt variations in input, limiting peak forces and preventing numerical instability within the solver. Together, these mechanisms constrain interaction dynamics while preserving responsiveness and maintaining simulation stability.

The visual force feedback is presented as a bar graph displayed in Fig. 2. Its value is computed as the 90th percentile of penetration depths measured across all active collision events at a given simulation step. This percentile-based formulation reduces sensitivity to transient spikes while remaining responsive to sustained excessive loading.

The maximum scale of the bar was heuristically defined as the highest force observed during controlled penetration stress tests conducted for each task-specific artificial anatomical model. This upper bound approximates the force magnitude required to perforate or structurally collapse the simulated tissue.

Although the simulation does not generate explicit visual perforation effects, excessive force may momentarily allow the endoscope tip to intersect or visually pass through tissue due to constraint resolution limits. The force bar therefore serves as a preventive feedback mechanism, encouraging users to regulate acceleration and avoid force levels that would be clinically unsafe in real procedures.

Through this mechanism, users are trained to maintain interaction within physiologically plausible limits, reinforcing safe manipulation strategies despite the absence of haptic resistance.

4.5 Force Stabilization and Numerical Control

During the **grabbing state**, the grasped segment is driven toward a target position defined by the user's hand pose. The positional difference between the hand and the segment is continuously evaluated, and a directional acceleration is applied to reduce this error. The magnitude of this acceleration is proportional to the distance to the target, enabling smooth convergence while preserving responsiveness.

To prevent overshooting and oscillatory behavior, a positional tolerance of 1,mm is defined. When the segment falls within this threshold relative to the target position, it is considered aligned and no additional corrective acceleration is applied. This tolerance improves numerical stability and avoids high-frequency micro-adjustments caused by tracking noise.

Rotational control follows a similar principle. At the moment of grasping, the current hand orientation is stored as a reference. Subsequent hand rotations are computed relative to this initial orientation and translated into torque applied to the grasped segment. By operating on relative rotation rather than absolute world orientation, the system ensures smooth torque transfer and avoids drift or unintended rotational artifacts.

Both the applied acceleration and torque are modeled as quadratic functions of their respective errors. This nonlinear formulation increases responsiveness for moderate inputs while progressively attenuating extreme deviations.

The gain parameters of these quadratic mappings were empirically calibrated to ensure numerical stability and to maintain the in-

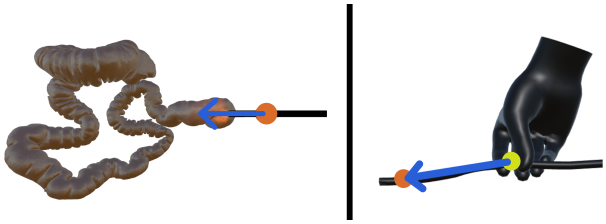

Figure 9: Mapping between the virtual interaction and the simulation domain. The orange points represent the frontier nodes linking the simulation space (left) to the virtual interaction space (right). The yellow point indicates the user's grabbing position on the endoscope. The blue vector represents the input direction computed from the displacement between the yellow point and the corresponding orange (frontier) point in the virtual space. This input is then transmitted to the corresponding orange frontier node in the simulation domain and applied a scale factor.

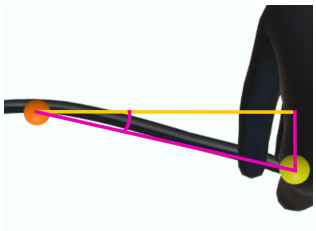

Figure 10: In the top view, the same two points are projected onto the XZ plane, forming a right triangle shown in pink. Again, the adjacent side is aligned with the Z axis (orange line), and the opposite side corresponds to the displacement along the X axis. The angle between the projected hypotenuse and the Z axis defines the horizontal component of the insertion angle

Figure 11: In the lateral view, the orange point represents the frontier reference and the yellow point represents the hand contact position. These two points are projected onto the YZ plane, forming a right triangle shown in cyan. The adjacent side of the triangle is aligned with the Z axis (orange line), while the opposite side corresponds to the displacement along the Y axis (vertical segment). The angle between the projected hypotenuse and the Z axis defines the vertical component of the insertion angle

sertion tube's behavior within the physically plausible limits of the simulation. Without such regulation, the absence of real-world resistance would allow abrupt hand accelerations to be directly transferred to the simulated structure, potentially causing segment overlap, excessive constraint violations, or unstable solver behavior.

By adopting a quadratic response profile, the system effectively dampens extreme user inputs while preserving precise and responsive control during typical manipulation. This approach balances stability and fidelity, ensuring robust interaction under unconstrained hand motion.

4.6 Synchronization Between Visual Interface and Internal Simulation

All previously described interface components operate exclusively on the portion of the endoscope that is visually exposed outside the patient's body. To propagate these interaction inputs to the internal physical simulation, we synchronize control based on the relative motion of the user's hand with respect to the virtual patient's anus at the moment the endoscope is grasped.

More specifically, when the instrument is grasped, the hand pose is recorded as a reference frame relative to the insertion point. Subsequent variations in hand position and orientation are computed as relative transformations, comprising both translational displacement and rotational deviation, with respect to this reference. These deltas represent insertion, withdrawal, and torque inputs.

The resulting relative transformation is then applied to the internal simulated endoscope model. A scaling factor is introduced to compensate for dimensional differences between the visual representation and the physical simulation domain (see Fig. 9), ensuring consistent behavior across coordinate systems.

This synchronization strategy decouples visual manipulation from internal constraint resolution while maintaining coherent and stable control transfer to the simulated structure.

Insertion and withdrawal of the instrument are computed from variations in the Euclidean distance between the virtual hand and a predefined frontier that represents the insertion reference. At each simulation frame, the current distance is compared with the distance measured in the previous frame. The resulting difference defines the incremental insertion or withdrawal displacement.

This frame-to-frame delta is interpreted as the user's axial manipulation input and is transmitted to the internal simulator through a dedicated function call. To ensure consistency between interaction space and simulation space, the displacement is multiplied by a scale factor that compensates for the larger spatial dimensions of the physical simulation environment. By relying on incremental distance variation rather than absolute position, the system achieves smooth and stable insertion control while preventing drift or discontinuities across coordinate frames.

The orientation transmitted to the simulator is represented as a quaternion constructed from two independent rotational components: the insertion angle and the axial rotation of the tube. The insertion angle is derived from the relative position between the virtual hand and the predefined frontier. Specifically, the angular deviation is computed from the hand's displacement along the X and Y axes with respect to the primary insertion axis (Z), as illustrated in Fig. 10 and Fig. 11. This component encodes the bending direction of the endoscope tip.

The axial rotation is obtained from the rotational motion of the user's hand and the control head, representing torque applied along the longitudinal axis of the tube. These two rotational components are combined to construct a quaternion that defines the full three-dimensional orientation input. The resulting quaternion is continuously transmitted to the simulator while the instrument remains in the grabbing state, ensuring smooth and consistent propagation of rotational control.

The virtual endoscope is driven by the same positional and rotational inputs used for the internal simulation, but without applying the spatial scale factor introduced in the simulator. Rather than interacting with internal physics constraints, the first segment of the virtual tube is directly updated according to the user's hand movement. To ensure controlled motion, this leading segment is constrained within four elongated rectangular colliders that define the walls of a guiding box. These boundaries restrict lateral displacement and provide immediate visual feedback regarding insertion alignment and angular deviation. The guiding box therefore acts as a spatial reference frame for external manipulation.

By separating visual control from physical simulation, the virtual endoscope serves as a stable and intuitive interaction proxy, while the physically simulated endoscope is responsible for handling internal deformation, collision response, and constraint resolution. This architectural decoupling improves perceptual stability and prevents solver corrections from propagating directly to the user interface.

We adopted this architecture instead of directly synchronizing each segment of the virtual endoscope with its physically simulated

counterpart to improve overall system stability. Preliminary implementations based on one-to-one segment synchronization produced undesirable rotational artifacts and abrupt positional discontinuities, primarily caused by transient inconsistencies in the physics solver during constraint resolution.

By decoupling the visual representation from the internal segment-level simulation, these high-frequency corrections are prevented from propagating to the user interface. In this configuration, the virtual endoscope functions primarily as a stable visual guide, while the physically simulated model governs the internal mechanical behavior of the procedure. This design choice prioritizes perceptual coherence and interaction stability over strict geometric mirroring between visual and simulated structures.

5 Preliminary Results

Preliminary evaluations were conducted with experienced colonoscopist preceptors and their residents at a hospital affiliated with the research group (see Fig. 12). Participants were encouraged to provide structured usability feedback while also contributing domain-specific terminology and procedural insights relevant to high-quality colonoscopy practice. The feedback revealed several important usability and modeling issues.

First, the angulation wheel of the control head was perceived as overly sensitive, requiring recalibration to achieve more precise and controllable tip deflection. Second, the mini-map displayed the colon in a non-standard orientation unfamiliar to practitioners. Participants recommended repositioning the colon to match its conventional anatomical orientation, thereby improving cognitive alignment with clinical workflows.

A supplementary screen positioned to the left of the intra-colonic view had been implemented to display the user's hand and the external portion of the virtual endoscope. This feature was originally introduced to reduce the need for users to look downward at the virtual hand and to compensate for the absence of haptic feedback. However, participants reported that the additional visual feed introduced attentional fragmentation and disrupted their natural focus. In clinical settings, colonoscopists primarily attend to the internal endoscopic view and glance at the external instrument only when necessary. Consequently, this auxiliary display was removed in subsequent iterations to better preserve authentic visual workflow.

The evaluation also identified a modeling inconsistency in the angulation mechanics of the simulated endoscope tip along the left/right axis. The initial implementation permitted up to 180° of lateral angulation in each direction, exceeding the mechanical constraints of standard clinical endoscopes, which typically allow approximately 135° per side. This discrepancy enabled unrealistic maneuverability and was corrected to better reflect real instrument limitations.

Finally, an issue was identified in the mapping between joystick inputs and angulation controls. The initial configuration assigned up/down angulation to the joystick's Y-axis and left/right angulation to the X-axis. Although this mapping appeared intuitive from an interface design perspective, it inadvertently promoted excessive reliance on lateral angulation.

In clinical practice, colonoscopists generally rely more heavily on shaft torque combined with vertical tip deflection rather than extensive left/right angulation. The original control scheme therefore encouraged interaction patterns that deviated from authentic procedural technique. The mapping was subsequently revised to better align with real-world colonoscopy maneuvering strategies.

6 Conclusion and Future Work

This work presented a low-cost virtual reality interface for colonoscope manipulation designed to provide stable and realistic interaction using off-the-shelf hardware. The primary contributions lie in the interaction architecture: (1) a three-state hand model (grabbing, sliding, and release) that reflects procedural handling strategies; (2) a frontier-based synchronization mechanism that maps user input to the internal simulator through relative positional variation and quaternion-based orientation control; and (3) a decoupled visual endoscope representation that functions as a stable interaction proxy while the internal physical model governs deformation and constraint resolution.

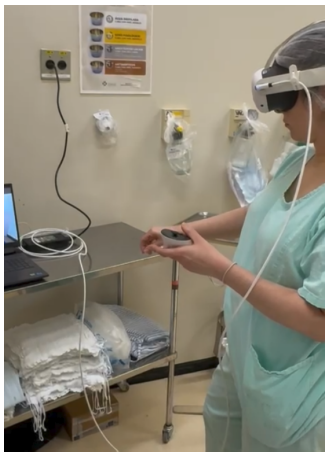

Figure 12: A practicing colonoscopist evaluated the interface and provided feedback at a hospital affiliated with the research group.

Together, these design decisions address key challenges in VR-based endoscopic simulation, including numerical instability, unconstrained user acceleration, and perceptual inconsistencies arising from the absence of haptic feedback. The proposed architecture balances interaction responsiveness with physical plausibility while remaining computationally lightweight and accessible.

Despite these advances, several limitations remain. The interface does not provide true force feedback, and multiple control parameters were empirically tuned to achieve stable behavior. Ongoing work focuses on formal validation, specifically face, content, and construct validity, along with adaptive parameter calibration, enhanced multimodal feedback strategies, and further robustness improvements. These efforts aim to support structured integration of the system into early-stage colonoscopy training programs.

References

[1] Omitted for blind review. 2

[2] C. Deul, T. Kugelstadt, M. Weiler, and J. Bender. Direct position-based solver for stiff rods. *Computer Graphics Forum*, 37:313–324, 9 2018. doi: 10.1111/cgf.13326 2

[3] V. K. Dik, L. M. Moons, and P. D. Siersema. Endoscopic innovations to increase the adenoma detection rate during colonoscopy. *World journal of gastroenterology: WJG*, 20(9):2200, 2014. 1

[4] C. A. Doubeni, D. Corley, V. Quinn, C. D. Jensen, A. Zauber, M. Goodman, J. R. Johnson, S. Mehta, T. A. Becerra, W. K. Zhao, J. Schottinger, V. P. Doria-Rose, T. Levin, N. Weiss, and R. Fletcher. Effectiveness of screening colonoscopy in reducing the risk of death from right and left colon cancer: a large community-based study. *Gut*, 67:291–298, 2016. doi: 10.1136/gutjnl-2016-312712 1

[5] M. Finocchiaro, C. Zabban, Y. Huan, A. D. Mazzotta, S. Schostek, A. Casals, A. Hernansanz, A. Menciassi, A. Arezzo, and G. Ciuti. Physical simulator for colonoscopy: A modular design approach and validation. *IEEE Access*, 11:36945–36960, 2023. doi: 10.1109/ACCESS.2023.3266087 1

[6] D. Hong, W. Tavanapong, J. Wong, J. Oh, and P. C. De Groen. 3d reconstruction of virtual colon structures from colonoscopy images. *Computerized Medical Imaging and Graphics*, 38(1):22–33, 2014. 1

[7] A. D. Koch, J. Haringsma, E. J. Schoon, R. A. De Man, and E. J. Kuipers. Competence measurement during colonoscopy training: the

use of self-assessment of performance measures. *Official journal of the American College of Gastroenterology (ACG)*, 107(7):971–975, 2012. 1

[8] M. G. Martins, L. Z. Morais, R. P. Torchelsen, A. Maciel, and L. P. Nedel. Efficient position-based deformable colon modeling for endoscopic procedures simulation. *SIGGRAPH*, 24(91), July 2024. doi: 10.1145/3641519.3657454 1

[9] A. Parra-Blanco1, N. González, R. González1, J. Ortiz-Fernández-Sordo, and C. Ordieres. Animal models for endoscopic training: do we really need them? *Endoscopy*, 45(6), June 2013. doi: 10.1055/s-0033-1344153 1

[10] S. F. Paulo, N. Figueiredo, J. A. Jorge, and D. S. Lopes. 3d reconstruction of ct colonography models for vr/ar applications using free software tools. 2018. 1

[11] K. Triantafyllou, L. D. Lazaridis, and G. D. Dimitriadis. Virtual reality simulators for gastrointestinal endoscopy training. *World Journal Gastrointest Endoscopy*, 6, Jan. 2014. doi: 10.4253/wjge.v6.i1.6 1

[12] S. Van der Wiel, R. K. Magalhães, C. R. R. Gonçalves, M. Dinis-Ribeiro, M. Bruno, and A. Koch. Simulator training in gastrointestinal endoscopy–from basic training to advanced endoscopic procedures. *Best Practice & Research Clinical Gastroenterology*, 30(3):375–387, 2016. 1

[13] C. M. Walsh, S. B. Umar, S. Ghassemi, H. Aihara, G. S. Anand, L. Cassani, P. Chahal, S. Dacha, A. Duloy, C. Huang, T. E. Kowalski, V. Kushnir, E. Qayed, S. G. Sheth, C. R. Simons-Linares, J. R. Taylor, S. A. F. Vela, R. L. Williams, and M. S. Wagh. Colonoscopy core curriculum. *Core Curriculum*, 93(2), Feb. 2021. doi: 10.1038/ajg.2011.481 1

[14] W. J. Woodall, E. H. Chang, S. Toy, D. R. Lee, and J. H. Sherman. Does extended reality simulation improve surgical/procedural learning and patient outcomes when compared with standard training methods? a systematic review. *Simulation in Healthcare*, 19(1S), 2024. 1

[15] W. Zhang, X. Liu, and B. Zheng. Virtual reality simulation in training endoscopic skills: A systematic review. *Laparoscopic, Endoscopic and Robotic Surgery*, 4:97–104, 2021. doi: 10.1016/j.lers.2021.09.002 1