Interacting with a Soft Continuum Manipulator Using Mixed Reality

Samuel Pilch[1,*], Michael Pabst[1,2,*], Annalena Gagsch[1], Dieter Schmalstieg[1,3] and Oliver Sawodny[1]

***Abstract*— Soft continuum manipulators are well suited for human–robot interaction due to their compliance and flexibility, but their continuous deformation poses challenges for modeling, visualization, and intuitive control. This work presents human interaction with a closed-loop controlled, pneumatically actuated soft continuum manipulator using a mixed-reality interface. The interface provides a spatially integrated visualization of the physical system and enables intuitive control. The physical soft continuum manipulator is modeled using the Euler-Lagrange formalism and controlled in a closed loop using input-output linearization, whereas the virtual soft continuum manipulator is modeled by constant-curvature joint chains. Its deformable surface is represented by a discretized cylindrical triangle mesh with vertices that are smoothly deformed by linear blend skinning. This provides continuous motion with low computational complexity. Inverse kinematics is formulated as a nonlinear optimization problem and is solved with IPOPT. The interface is implemented in Unity on a Meta Quest 3, using ROS for communication to the robot, which is actuated and controlled via a dSPACE real-time system. Users can visualize system states, manipulate the virtual tool center point to control the real soft continuum manipulator, and plan trajectories via simulation before execution. A qualitative expert study comparing virtual-only and coupled virtual-real interaction indicates that mixed reality control was perceived as intuitive and supportive for accurate 3D positioning. The results demonstrate the potential of mixed reality for the intuitive operation of compliant robotic systems and support the integration of immersive technologies into soft robot control.**

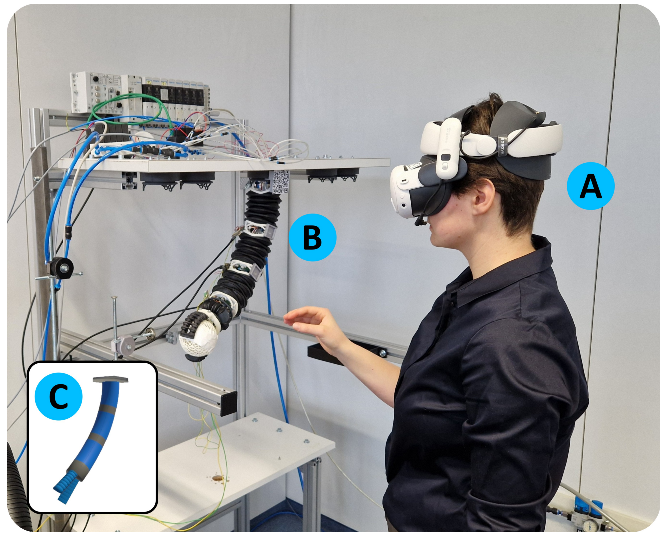

Fig. 1. We present a mixed reality interface for interacting with a soft continuum manipulator. The user wears a mixed reality device (A), which enables interaction with a pneumatically actuated, fully controlled soft continuum manipulator, i.e., the soft arm (B). To facilitate this interaction, we developed a virtual model of the soft arm (C), allowing the user to see a virtual representation of the real soft manipulator. This virtual model serves as a bridge between the real and virtual worlds, enabling control through the mixed reality interface.

I. INTRODUCTION

In recent decades, robotics has changed its focus from rigid industrial machines operating behind safety fences to systems that share their workspace with humans and enable safe human–robot interaction [1]. In industrial and service-oriented domains, robots are widely deployed for demanding or safety critical tasks [2], but their use in close proximity to humans still requires designs that are inherently safe and, at the same time, intuitive to operate. This need aligns with the vision of a future production environment in which humans and robots collaborate and combine robotic precision and endurance with human creativity [3]. To realize such a collaboration, new design paradigms have emerged, ranging from lightweight rigid structures, such as the DLR lightweight robot III [4], to intrinsically compliant actuation concepts such as series elastic actuators used in Baxter [5].

* Equal contribution.
[1] University of Stuttgart, Germany.
[2] IMPRS-IS, Germany.
[3] Graz University of Technology, Austria.
samuel.pilch@isys.uni-stuttgart.de,
michael.pabst@visus.uni-stuttgart.de
Project website:
https://m-pabst.github.io/xr-soft-manipulator/

Beyond these approaches, soft robotic systems offer a particularly promising basis for collaboration, because their flexible and compliant structures can adapt to unknown or dynamic environments with reduced risk of injury [6]. Continuum manipulators are especially attractive in this context due to their continuously deformable bodies, and pneumatic actuation is well suited because it is lightweight, inexpensive, and easy to implement [7]. Moreover, the compressibility of air and direct-drive characteristics provide crucial safety benefits for collaboration [8]. Consequently, the soft arm (Fig. 1), a soft continuum manipulator actuated pneumatically by soft bellows, offers favorable prerequisites for close human–robot collaboration. However, the very properties that make continuum robots safe and versatile also complicate their operation, and intuitive control remains a central challenge [9]. This problem motivates interaction methods that make robot states spatially understandable, and, their control more natural. In particular, mixed reality (MR) can support users by providing intuitive three dimensional visualizations and coupling of virtual and real representations for motion preview and spatially meaningful interaction feedback.

This work presents the design, implementation, and evaluation of an MR system for control and visualization of a soft

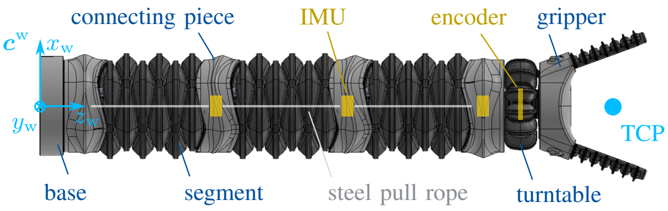

Fig. 2. Schematic illustration of the soft arm.

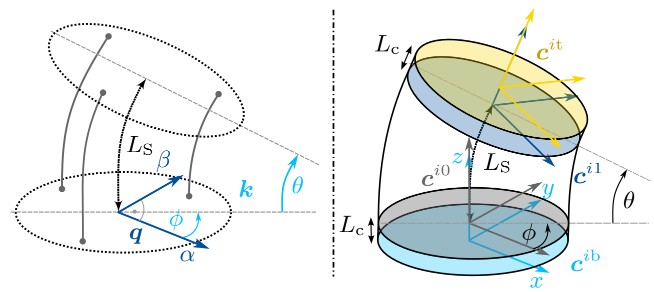

Fig. 3. Kinematics of the segment i.

continuum manipulator. A virtual model of the soft arm, incorporating both forward and inverse kinematics, is coupled in real time with a physical pneumatically actuated system, enabling motion of the real robot through manipulation of its virtual counterpart. In addition, a trajectory planning module allows movements to be simulated in the virtual environment before being executed on the physical system.

To assess the suitability of the MR interface, a qualitative expert study is conducted. We evaluate how MR interaction supports intuitive operation, spatial understanding, and controllability of the soft robotic system, and identify strengths and limitations of the approach.

In summary, we make the following contributions:

- Human interaction with a pneumatically actuated, closed-loop-controlled soft continuum manipulator using a real-time MR interface.
- Visualization and spatially integrated control of a pneumatically actuated soft continuum manipulator, coupling a virtual kinematic model with the physical system.
- A qualitative expert study with soft robotics researchers comparing virtual-only and coupled virtual-real interaction in terms of intuitiveness, spatial understanding, and controllability.

II. RELATED WORK

A. Control of Soft Manipulators for Interaction

Despite their compliance and safety, soft continuum manipulators remain difficult to model and control due to nonlinear deformation and complex actuation-motion mappings [10], [11]. Nevertheless, collaborative control has been achieved with inverse kinematics methods [12], including adaptive dynamic control with redundancy resolution for pneumatic arms to improve task-space tracking under uncertainty [13], geometrically exact Jacobian IK enabling accurate large-deformation control with reduced DOF [14] and efficient variable-separation IK for two-segment robots providing full solution coverage, faster computation, and workspace/singularity characterization [15]. For human-robot collaboration, prior work includes interaction-aware offline trajectory optimization considering stiffness and external force detectability [16], and immersive inverse-kinematics interfaces enabling intuitive real-time control of soft robots [17].

B. Interacting with Robots using Mixed Reality

Conventional robot operation via a desktop user interface (using keyboard, mouse or joystick) splits attention between screen and robot and requires mapping 2D displays to a 3D workspace, impairing spatial understanding and precision. Mixed reality (MR) embeds virtual content in situ [18], aligning robot models and controls with the physical workspace via a head-mounted display to improve situational awareness and reduce the input-execution gap [19]–[21]. Prior work demonstrates diverse MR-based robot interaction concepts. Gaze-based AR control enables hands-free robot commands [22]. VR-based target manipulation supports intuitive inverse kinematics control of soft robots, although decoupled from the real workspace [17], while teleoperation transfers demonstrated configurations directly to robots [23]. However, MR interfaces for soft robot control remain underexplored, which is addressed in this work.

III. MODELING AND CONTROL OF SOFT ARM

A. Soft Arm

The soft arm considered in this article (Fig. 2) is a pneumatically driven, spatially moving soft continuum manipulator based on bellows [24], [25]. It consists of a base, three segments, a turntable, and a gripper, which are interconnected by 3D-printed connecting pieces. The segments, turntable, and gripper are equipped with bellows that can be actuated independently. Each segment contains three circumferentially arranged bellows with equal angular spacing, while the turntable is actuated by two antagonistic bellows [24]. A steel pull rope along the central axis prevents elongation. For sensing, an inertial measurement unit (IMU) with an internal sensor-fusion algorithm, the Mti-630 AHRS from Xsens, is integrated into each segment, and the angle of the turntable is measured by a rotary encoder.

The kinematics of the soft arm can be decomposed into segment kinematics and turntable kinematics. The gripper is modeled with a binary state, i.e., either open or closed. The motion of the turntable is described by the rotation angle $\psi \in [-40°, 40°]$. The kinematics of the i^{th} segment can be parameterized either by the configuration coordinates $\boldsymbol{k}_i = [\phi_i, \theta_i]^\intercal$, which provide an intuitive description of the segment orientation and curvature, or by the generalized planar coordinates $\boldsymbol{q}_i = [\alpha_i\ \beta_i]^\intercal$ (Fig. 3). The latter representation is particularly suitable for control, as it is continuously differentiable and enables smooth point-to-point trajectories in Cartesian space $\boldsymbol{c} = [x\, y\, z]^\intercal$. The orientation angle is defined as $\phi_i \in [0°, 360°]$, and the curvature angle as $\theta_i \in [0°, 30°]$. The relationship between the generalized coordinates in polar form as

$$\begin{bmatrix}\alpha_i \\ \beta_i\end{bmatrix} = \begin{bmatrix}\cos\phi_i \\ \sin\phi_i\end{bmatrix} \theta_i, \tag{1}$$

where ϕ_i denotes the polar angle, and θ_i, the radial magnitude [24]. Conversely, the configuration coordinates follow from the generalized components via

$$\phi_i = \operatorname{atan2}(\beta_i, \alpha_i), \qquad \theta_i = \sqrt{\alpha_i^2 + \beta_i^2}. \tag{2}$$

It is assumed that each segment deforms with constant curvature [25].

The forward kinematics of the soft arm are expressed using homogeneous transformation matrices. The homogeneous transformation of the bottom to the top segment [25] is given by

$$
\begin{aligned}
&{}^{\mathrm{b}}\mathbf{H}_{\mathrm{S}i}^{\mathrm{t}} = {}^{\mathrm{b}}\mathbf{H}_{\mathrm{S}i}^{0}\, {}^{0}\mathbf{H}_{\mathrm{S}i}^{1}\, {}^{1}\mathbf{H}_{\mathrm{S}i}^{\mathrm{t}},\ i = \{1, 2, 3\},\ \text{with} \\
&{}^{\mathrm{b}}\mathbf{H}_{\mathrm{S}i}^{0} = {}^{1}\mathbf{H}_{\mathrm{S}i}^{\mathrm{t}} = \begin{bmatrix} & 0 \\ \mathbf{I}_{3\times3} & 0 \\ & L_{\mathrm{C}} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix},\ {}^{0}\mathbf{H}_{\mathrm{S}i}^{1} = \begin{bmatrix} {}^{0}\mathbf{R}_{\mathrm{S}i}^{1} & \mathbf{o}_i^1 \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix}, \\
&{}^{0}\mathbf{R}_{\mathrm{S}i}^{1} = \begin{bmatrix} c_\phi^2(c_\theta - 1) + 1 & s_\phi c_\phi (c_\theta - 1) & c_\phi s_\theta \\ s_\phi c_\phi (c_\theta - 1) & c_\phi^2(1 - c_\theta) + c_\theta & s_\phi s_\theta \\ -c_\phi s_\theta & -s_\phi s_\theta & c_\theta \end{bmatrix}, \\
&\mathbf{o}_{\mathrm{S}i}^{1} = \begin{cases} \begin{bmatrix} \frac{c_\phi(1-c_\theta)L_{\mathrm{S}}}{\theta} & \frac{s_\phi(1-c_\theta)L_{\mathrm{S}}}{\theta} & \frac{s_\theta L_{\mathrm{S}}}{\theta} \end{bmatrix}^{\mathsf{T}}, & \theta \neq 0, \\ \begin{bmatrix} 0 & 0 & L_{\mathrm{S}} \end{bmatrix}^{\mathsf{T}}, & \theta = 0, \end{cases}
\end{aligned} \tag{3}
$$

where L_{C} denotes half the length of the connecting pieces, and L_{S}, the constant segment length due to lack of elongation caused by the steel pull rope. For brevity, the segment index i is omitted in ϕ and θ, and the shorthand $s_{(\cdot)} = \sin(\cdot)$, $c_{(\cdot)} = \cos(\cdot)$ is used. The turntable rotates the soft arm about its central axis at an angle ψ and is described by the homogeneous transformation ${}^{\mathrm{b}}\mathbf{H}_{\mathrm{T}}^{\mathrm{t}}$. Together with the three bending segments, the robot can manipulate the tool center point (TCP) in all six degrees of freedom. The origin of the right-handed world coordinate system $\boldsymbol{c}^{\mathrm{w}}$ is located at the bottom of the base (Fig. 2). The rigid kinematic offsets between the base and the first bellow segment, and between the turntable and the TCP within the gripper, are modeled by pure translations of length L_* with $* \in \{\mathrm{B}, \mathrm{G}\}$ as the following

$$
{}^{\mathrm{b}}\mathbf{H}_{\mathrm{T}}^{\mathrm{t}} = \begin{bmatrix} \mathrm{c}_\psi & -\mathrm{s}_\psi & 0 & 0 \\ \mathrm{s}_\psi & \mathrm{c}_\psi & 0 & 0 \\ 0 & 0 & 1 & L_{\mathrm{T}} \\ 0 & 0 & 0 & 1 \end{bmatrix},\ {}^{\mathrm{b}}\mathbf{H}_{*}^{\mathrm{t}} = \begin{bmatrix} & 0 \\ \mathbf{I}_{3\times3} & 0 \\ & L_* \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix}. \tag{4}
$$

Consequently, the overall forward kinematics from the world frame to the TCP are obtained by concatenating the individual component transformations

$$
{}^{\mathrm{w}}\mathbf{H}^{\mathrm{TCP}} = {}^{\mathrm{b}}\mathbf{H}_{\mathrm{B}}^{\mathrm{t}} \prod_{i=1}^{3} {}^{\mathrm{b}}\mathbf{H}_{\mathrm{S}i}^{\mathrm{t}}\ {}^{\mathrm{b}}\mathbf{H}_{\mathrm{T}}^{\mathrm{t}}\, {}^{\mathrm{b}}\mathbf{H}_{\mathrm{G}}^{\mathrm{t}}. \tag{5}
$$

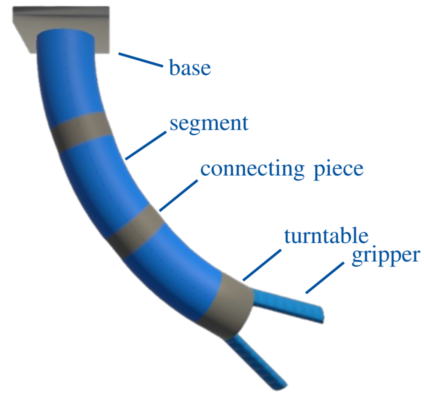

Fig. 4. Virtual soft arm.

B. Model-Based Control in Configuration Space

A dynamic model of the soft arm describes inter-segment coupling and supports controller design [25]. The bellows are modeled as massless, and each segment is represented by a lumped mass located on the central axis at the tip of the segment. The equations of motion are derived via the Euler-Lagrange formalism with respect to the generalized coordinates $\boldsymbol{q}$.

Based on this model, a centralized model-based controller is designed in generalized coordinates $\boldsymbol{q}$ and combined with a decoupled pneumatic control layer employing medium-pressure adaptations [25]. The mechanical controller uses input-output linearization and applies proportional-integral-derivative (PID) feedback to the resulting linearized error dynamics. The desired actuator moment vector is computed according to the control law

$$\boldsymbol{\tau}_{\mathrm{act,d}}(\boldsymbol{p}_{\mathrm{d}}) = \boldsymbol{M}_{\mathrm{d}}(\boldsymbol{q}_{\mathrm{d}})\,\boldsymbol{\nu}_{\mathrm{FF}} + \boldsymbol{N}_{\mathrm{d}}(\boldsymbol{q}_{\mathrm{d}}, \dot{\boldsymbol{q}}_{\mathrm{d}}, \boldsymbol{p}_{\mathrm{d}}) + \boldsymbol{\nu}_{\mathrm{FB}}. \tag{6}$$

where $\boldsymbol{M}(\boldsymbol{q}) \in \mathbb{R}^{2n\times2n}$ denotes the mass matrix, $\boldsymbol{N}(\boldsymbol{q}, \dot{\boldsymbol{q}}, \boldsymbol{p}) \in \mathbb{R}^{2n}$ collects the generalized gravity, spring, and damping forces and moments, and $\boldsymbol{\tau}_{\mathrm{act}}(\boldsymbol{p}) \in \mathbb{R}^{2n}$ denotes the actuator moment vector. The virtual feedforward input is selected as $\boldsymbol{\nu}_{\mathrm{FF}} = \ddot{\boldsymbol{q}}_{\mathrm{d}}$. The feedback term $\boldsymbol{\nu}_{\mathrm{FB}}$ is implemented as a PID law,

$$\boldsymbol{\nu}_{\mathrm{FB}} = K_{\mathrm{P}}\,\boldsymbol{e}(t) + K_{\mathrm{I}} \int_0^t \boldsymbol{e}(\tau)\,\mathrm{d}\tau + K_{\mathrm{D}} \frac{\mathrm{d}\boldsymbol{e}(t)}{\mathrm{d}t}, \tag{7}$$

with the control error $\boldsymbol{e} = \boldsymbol{q}_{\mathrm{d}} - \boldsymbol{q}$ and K_{P}, K_{I}, and K_{D} being the proportional, integral, and derivative gains, respectively.

The rotation angle of the turntable ψ is controlled in a quasi-static manner using a feedforward term obtained from an experimentally identified characteristic curve and an additional PID feedback term.

C. Model-Based Control in Task Space

For manipulation tasks, the desired motion is naturally specified in the task space as Cartesian target points of the TCP. However, the soft arm is controlled in the joint space. Therefore, each desired TCP position $\boldsymbol{x}_{\mathrm{TCP,d}}$ is converted into a corresponding set of joint coordinates by solving an

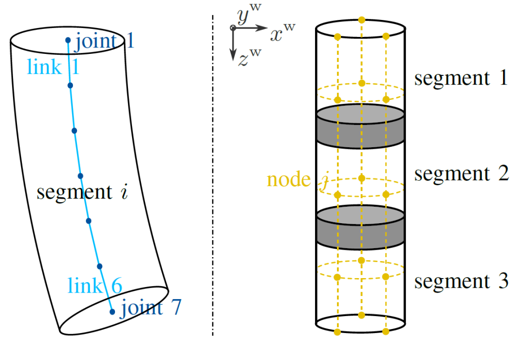

Fig. 5. Kinematic joint chain of segment i (left) and mesh representation of the segments (right).

inverse kinematics (IK) problem. This mapping is formulated as the following constrained nonlinear program

$$\begin{aligned} \min_{\boldsymbol{q}} \quad & \| \boldsymbol{x}_{\mathrm{TCP}}(\boldsymbol{q}) - \boldsymbol{x}_{\mathrm{TCP,d}}(\boldsymbol{q}_{\mathrm{d}}) \|_2^2 , \\ \text{s.t.} \quad & \alpha_i^2 + \beta_i^2 \leq \theta_{\max}^2, \qquad i = 1,2,3. \end{aligned} \tag{8}$$

The decision variables are the generalized coordinates $\boldsymbol{q}$. The objective minimizes the squared Cartesian position error at the TCP. Feasibility is ensured by constraining the bending magnitude of each segment, which is $\theta_{\max} = 30°$. The IK problem is solved online as a nonlinear program using IPOPT via CasADi [26]. To ensure continuity in motion, the IK is warm-started with the previous solution. On a standard laptop, the solve time typically ranges between 5 ms and 15 ms. Given a MR update period of 20 ms, this solution time is sufficient for online operation. Due to the high update rate and the comparatively low-dynamic user-robot interaction in MR (Sec. VI), the intermediate dynamics between successive IK solutions can be neglected. The given TCP path points are connected using second-order trajectories that ensure continuous differentiability of position and velocity. The rotation angle ψ is controlled in feedforward mode to improve usability in the MR interface.

IV. Modeling and Control of Virtual Soft Arm

A. Virtual Soft Arm

The virtual soft arm (Fig. 4) represents the physical system for visualization and serves as the interaction body in MR. The bending components are shown in blue, while rigid structures, including the turntable, are shown in gray. Geometric consistency between the virtual and real setups is ensured by matching the virtual manipulator's length and circumference to those of the physical system (Fig. 7). The geometry is simplified to a set of cylindrical primitives, where each cylinder represents a component of the soft arm. The visualization of the bellows folds are intentionally omitted, as they provide no benefit for the intended application.

B. Kinematic Modeling

Due to the bending deformation of the bellows segments, they cannot be adequately modeled by rigid cylinders. Therefore, the surface of each of the three segment cylinders is represented by a triangular mesh. The length of the cylinder is discretized into 100 nodes, and the circumference, into 64 nodes. The segment centerline is modeled as a joint chain (Fig. 5). Each segment consists of seven joints connected by six links. The segment curvature is projected onto the joint chain by subdividing it into six increments of equidistant curvature. Joint positions are obtained through successive coordinate transformations along the chain using an incremental bending angle. Under the constant-curvature assumption, the segment orientation evolves consistently along the centerline. Finally, each mesh node is positioned relative to the centerline by applying the linear blend skinning algorithm, resulting in a smooth deformation of the cylindrical surface that follows the underlying joint chain.

Linear blend skinning deforms the mesh of the segments (Fig. 5) by computing each node as a weighted blend of multiple link transformations with respect to the unbent pose. Each link pose in the world coordinates is represented by the homogeneous transformation

$$ {}^{\mathrm{w}}\mathbf{H}^{\mathrm{S}i,k+1} = {}^{\mathrm{b}}\mathbf{H}^{\mathrm{t}}_{\mathrm{B}} \prod_{i=1}^{3} \prod_{k=1}^{6} {}^{\mathrm{k}}\mathbf{H}^{k+1}_{\mathrm{S}i}, \tag{9}$$

where ${}^{\mathrm{k}}\mathbf{H}^{k+1}_{\mathrm{S}i}$ corresponds to the homogeneous transformation from joint k to joint $k+1$ within segment i. Here, ${}^{\mathrm{k}}\mathbf{H}^{k+1}_{\mathrm{S}i}$ corresponds to ${}^{0}\mathbf{H}^{1}_{\mathrm{S}i}$, which describes the homogeneous transformation from the beginning to the end of segment i (Eq. 3) but with $\theta/6$ instead of θ and $L_{\mathrm{S}}/6$ instead of L_{S}. The unbent link pose in world coordinates is denoted by ${}^{\mathrm{w}}\mathbf{H}^{\mathrm{S}i,k+1}_{\theta=0}$, since no curvature applies. Each node is influenced by exactly two adjacent links with indices k and $k+1$. The corresponding weights are obtained by linear interpolation of the node coordinate z_j between two links z_k and z_{k+1} in a segment

$$\begin{aligned} t_j &= \frac{z_j - z_k}{z_{k+1} - z_k}, \quad t_j \in [0,1], \\ w_{j,k} &= 1 - t_j, \quad w_{j,k+1} = t_j. \end{aligned} \tag{10}$$

The node position in world coordinates results in

$$\mathbf{v}_j^{\mathrm{w}} = \sum_{k}^{k+1} w_{j,k} \, {}^{\mathrm{w}}\mathbf{H}^{\mathrm{S}i,k+1} \left({}^{\mathrm{w}}\mathbf{H}^{\mathrm{S}i,k+1}_{\theta=0} \right)^{-1} \mathbf{v}^{\mathrm{w}}_{\theta=0,j} \tag{11}$$

where $\mathbf{v}^{\mathrm{w}}_{\theta=0,j}$ is the node in world coordinates in the unbent pose. In addition to the segments, the visualization of the soft arm (Fig. 4) includes the base, the turntable, and the gripper. The base is modeled as a rigid cuboid mesh, while the turntable is represented by a rigid cylinder mesh attached to the last arm joint. The gripper fingers are modeled as the segments of a kinematic joint chains with individual mesh segments to visualize bending motions while opening and closing. For interaction, the finger segments are equipped with colliders to enable collision-based grasping.

C. Kinematic Control and Trajectory Generation

For interactive use in MR, smooth motion between two arm configurations is required. To avoid discontinuities in the

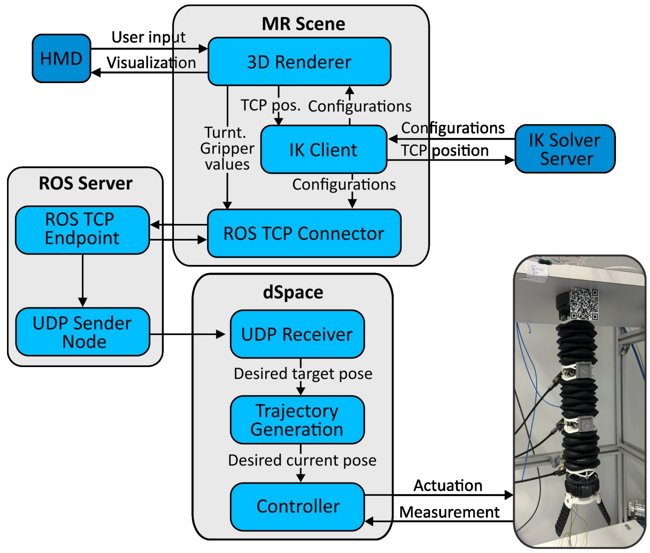

Fig. 6. System overview.

Fig. 7. Comparison of the model visualization (top) with the real soft arm (bottom) in four exemplary poses.

orientation angle ϕ_i caused by directly interpolating configuration coordinates $\boldsymbol{k}$ when crossing its origin, interpolation is performed in the generalized coordinates

$$\boldsymbol{q}(t) = (1-t)\,\boldsymbol{q}_0 + t\,\boldsymbol{q}_1, \tag{12}$$

where $t \in [0,1]$ is the normalized time parameter, and the indices 0 and 1 denote the start and target configuration. The interpolated values are mapped back to the configuration coordinates $\boldsymbol{k}$ and passed to the forward kinematics. The turntable rotation is interpolated in the same way and applied synchronously.

V. MIXED REALITY INTERFACE

Our system connects an MR user interface with a soft robotic arm through an external computation and communication pipeline (Fig. 6). The architecture consists of three main components, an MR application, an IK service, and a middleware layer connected to a real-time controller. An application developed in Unity serves as the user interface, which visualizes target poses for the robot. IK is computed externally on a separate IK solver server. Unity and the IK service communicate via a persistent WebSocket connection. The computed robot configuration is then transmitted from the head-mounted display (HMD) via the Robot Operating System (ROS). Motion commands are forwarded to the downstream real-time control system (dSPACE) for physical robot actuation.

A. Mixed Reality Soft Manipulator Control Interface

The MR interface enables users to communicate and interact with the soft arm. It defines which interactions are possible, how they are performed, and how user actions affect the virtual model and the physical robot. In the following, we summarize the main interface components.

The MR Interface is operated exclusively via gesture recognition and hand tracking. This enables fine-grained manipulation without controllers and is particularly beneficial in live-teaching scenarios where the virtual and physical systems share the same workspace. The interface comprises mode-specific panels (forward kinematics, inverse kinematics/live teaching, trajectory planning). All panels can be freely placed in 3D space to support an organized layout.

B. Mixed Reality Soft Manipulator System

a) Virtual soft arm: The soft arm visualization is present across all modes, while its control depends on the active menu (Fig. 4). In every mode, the turntable can be rotated via an interaction handle; an arc indicator shows the remaining rotation range. A gripper control allows for opening and closing the gripper in all modes. We used a QR code attached to the robot to align the virtual soft arm with the real soft arm.

b) Forward kinematics (virtual): The forward-kinematics panel allows specifying orientation and curvature angles in degrees per section using sliders or numeric input fields; turntable rotation can be set in the same way (E in Fig. 8). Slider changes update the pose immediately, whereas numeric input is applied explicitly (optionally over a specified duration). Together with the continuously displayed TCP position, this approach supports planning and simulation. The mode also includes grasp interaction with virtual objects.

c) Inverse kinematics (live) and live teaching: In inverse kinematics, users control the arm by grasping and placing a TCP target (B,D in Fig. 8). A workspace indicator visualizes the reachable region; if a target is placed outside, the nearest reachable point is used. The panel displays current values and the TCP position (display-only). Targets can optionally be specified as $[x^{\mathrm{w}}, y^{\mathrm{w}}, z^{\mathrm{w}}]$ coordinates. In live teaching, the virtual arm is aligned with the real system and the commanded TCP motion, turntable rotation, and gripper state are executed on the physical robot. Alternatively, teleoperation can be performed without co-location of the virtual and real systems.

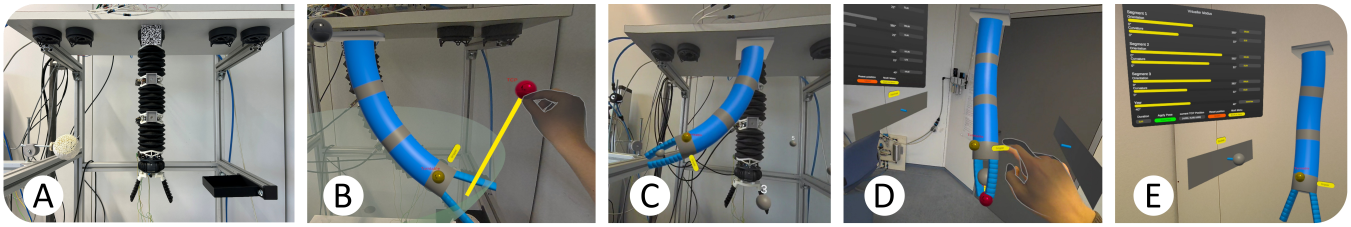

Fig. 8. A: Study setup task 2.2; B: Task 2.2; C: Task 2.3; D: Task 1.3; E: Task 1.1 and 1.2.

d) Trajectory planning: Trajectory planning allows placing multiple TCP targets and simulating the resulting motion (C in Fig. 8). Saved targets are listed with coordinates, per-point timing (move or hold time), and gripper state. The virtual arm can replay the trajectory for verification before sending it to the real system.

VI. EXPERT STUDY

To qualitatively evaluate usability and overall user experience of the proposed MR interface, we conducted an expert study involving a comprehensive task that covered all implemented functions.

A. Study Setup

A total of $N = 6$ roboticists participated in the expert study (four female, two male), aged 20–30 years. All participants reported little to no prior experience with MR applications or immersive technologies. All participants have a background in soft robotics and are familiar with a desktop interface to interact with a soft robot. All participants completed all tasks and completed the study. The study was divided into two phases in the real environment: (i) a virtual-soft-arm phase to assess the usability of forward- and inverse-kinematics interaction, and (ii) a coupled virtual-real phase to evaluate control transfer to the physical system and trajectory execution. After each phase, participants completed a questionnaire. For most tasks, participants received only target goals rather than step-by-step instructions to assess the intuitiveness and self-descriptiveness of the application. After each phase, participants completed a questionnaire with items rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). All participants also participated in a semi-structured interview to capture impressions, difficulties, and improvement suggestions in greater depth.

1) Phase 1 (Virtual): This phase consisted of three tasks using the virtual soft arm. Task 1.1 evaluated the forward-kinematics control using sliders (E in Fig. 8). Participants were shown the forward-kinematics interface and asked to set prescribed orientation and curvature angles for each arm section using sliders. The goal was to match all target values with an accuracy of $\pm 2°$. After reaching the target configuration, the TCP position was read out, and the arm was reset. Task 1.2 evaluated the forward-kinematics control using numeric input (E in Fig. 8). Participants entered prescribed orientation and curvature values through input fields, in order to reach the target pose within $5\,\mathrm{s}$. After reaching the target pose, the TCP position was read and the arm was reset. Task 1.3 evaluated inverse-kinematics TCP control and virtual grasping (D in Fig. 8). Participants were shown the inverse-kinematics interface and asked to move the arm via TCP control. They then grasped a virtual ball from a virtual wall on the left side and placed it onto a virtual wall on the right side.

2) Phase 2 (Real): Phase 2 comprised three tasks using the coupled virtual-real system. Task 2.1 focused on exploration and control of the target pose. Participants were introduced to the real coupled system and asked to freely explore its behavior and motion limits. They were then instructed to command a prescribed TCP position using inverse-kinematics control. Task 2.2 evaluated grasp-and-place performance (A,B in Fig. 8). Participants were shown the grasping setup and asked to grasp a 3D-printed ball from a magnetic holder using the physical soft manipulator and place it in a designated box, guided through the MR interface. Task 2.3 evaluated the creation and execution of the trajectory (C in Fig. 8). Participants were shown the trajectory-mode interface and asked to create and execute a trajectory on the real robot. The trajectory required the arm to bend toward the left frame post at an intermediate time point and end bent toward the right frame post, while avoiding a rear obstacle. The goal was to reach the final configuration with a closed gripper and complete the motion within $15\,\mathrm{s}$.

B. Study Results

Phase 1 results: After phase 1, participants rated the system feedback (Fig. 9, a)), focusing on the transparency and traceability of their settings, i.e., whether the chosen parameters were clearly reflected in the state and visualization of the system. In general, the system feedback was perceived as clear and easy to understand, indicating that the participants were well-informed about the state of the system throughout the tasks (P1.Q1). Participants also reported high satisfaction with the IK-based interaction, specifically with handling of the grasped TCP target sphere used for control (P1.Q3). Some participants initially struggled with sliders or grasping the TCP target due to limited MR experience, but completed the tasks without major issues once the interaction concept was understood (P1.Q4). Ratings for the ease of positioning the arm using IK and the grasp interaction were among the highest, highlighting the intuitiveness of the virtual application (P1.Q5). All participants agreed that entering orientation and curvature via numeric input fields was easier than using sliders (P1.Q2).

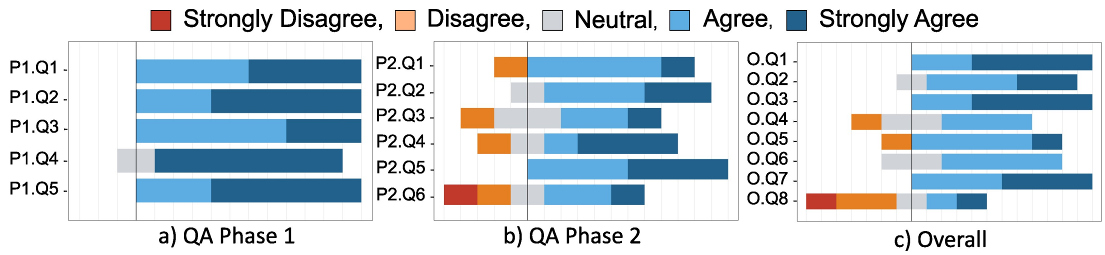

Fig. 9. Results of phase 1, 2, and overall impression of expert study.

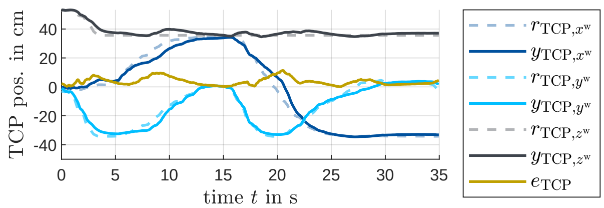

Fig. 10. Interactive IK control of the soft arm's TCP in task 2.2.

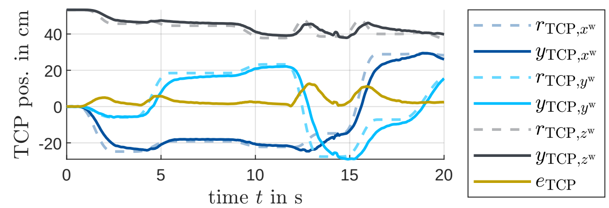

Fig. 11. Point-to-point IK control of the soft arm's TCP in task 2.3.

Phase 2 results: After phase 2, participants evaluated the coupled virtual-real operation (Fig. 9, b)). They rated the simultaneity and overall quality of executed motions positively, indicating good correspondence between virtual commands and real-system behavior (P2.Q1, P2.Q2). Grasping was perceived as more demanding than purely virtual interaction, mainly due to increased depth-perception and targeting requirements (P2.Q3). Nevertheless, phase 1 was considered a useful preparation, as learning the control scheme in isolation facilitated subsequent real-system operation (P2.Q4). Observing and guiding the real arm through its virtual counterpart was frequently described as intuitive and supportive for understanding motion execution, and trajectory planning was generally perceived as clear and easy to use (P2.Q3, P2.Q5). A recurring suggestion concerned the lack of explicit reachability feedback: despite workspace visualization, participants wanted confirmation whether a selected target would be executed as specified and which alternative pose would be chosen otherwise. The confined physical workspace occasionally limited hand movement near the mounting rig of the soft arm, although participants noted that the soft arm could still reach many of these regions from alternative approach directions (P2.Q6).

Overall user feedback: Participants rated the overall experience positively, including the clarity of the interface, the comprehensibility of the menu, and the labeling of the interaction elements (Fig. 9, c)), (O.Q1, O.Q3). Most of the participants reported improvements in performance with more practice (O.Q2). The reported below-average mental workload indicates that they did not have to concentrate hard to avoid errors (O.Q4). Participants reported that MR helped them better understand 3D arm motion and grasp interactions. Participants generally felt in control of the arm and understood the capabilities and limits of the system (O.Q6, O.Q7). Participants also reported little concern about unsafe motions, although some were concerned about triggering unintended movements (O.Q8). All participants were able to complete the tasks successfully. In the interviews, panel structure and navigation were described as logical, and perceived MR benefits included improved spatial understanding and intuitive direct 3D interaction with the arm. The live-teaching mode and pre-execution trajectory simulation were highlighted as the most valuable features. Overlay offsets and virtual-real latency were mainly noticed during strong bending or fast motions (O.Q5). Sometimes these artifacts were perceived as reducing control, but never considered highly disruptive. Participants preferred input fields over sliders for curvature and orientation and suggested user interface changes such as hiding the virtual arm during teaching and highlighting unreachable targets. Overall, participants reported a positive experience with the system.

Sim-to-real transfer: The model-based control of the soft arm in the configuration space has been validated in previous work [25]. In this work, we evaluate the sim-to-real transfer of model-based task-space control for the soft arm. Fig. 10 shows a representative excerpt of the controller response during interactive IK control of the soft arm's TCP in task 2.2, and Fig. 11 illustrates a representative excerpt of the closed-loop trajectory tracking behavior during point-to-point IK control of the TCP in task 2.3. Here, r_{TCP} denotes the TCP reference signal; y_{TCP}, the measured TCP signal obtained by forward kinematics, and e_{TCP}, the resulting Euclidean task-space error. Overall, the trajectories follow well in both cases. The root mean square error (RMSE) is $4.5\,\mathrm{cm}$ in task 2.2, and, $6.4\,\mathrm{cm}$ in task 2.3, with higher RMSE in task 2.3 caused by the more dynamic point-to-point motions. In steady state, the RMSE decreases to $1.4\,\mathrm{cm}$ for task 2.2, and $2.4\,\mathrm{cm}$ for task 2.3, which is sufficiently small to complete the tasks successfully.

VII. CONCLUSION AND FUTURE WORK

In this work, we presented a novel MR system for intuitive interaction with a soft continuum manipulator via hand gestures. The system couples a discretized kinematic model and inverse-kinematics-based TCP control with a pneumatically actuated robot through a Unity-ROS pipeline, enabling spatial manipulation and trajectory planning with prior virtual simulation.

A qualitative expert study indicates good usability and spatial understanding, particularly for 3D control. Overall, MR interaction shows strong potential for intuitive control of soft manipulators and future hybrid or teleoperation scenarios. Future work can extend the current system to a multi-robot soft arm setting. In addition, integrating multimodal interaction with the soft manipulator using a large language model for natural-language communication can be explored.

APPENDIX

For phase 1 we asked the following questions: **P1.Q1**: *I was satisfied with the system feedback.* **P1.Q2**: *The input field was easier to use than the input slider.* **P1.Q3**: *Handling the IK target sphere worked well.* **P1.Q4**: *I quickly understood how to position the arm in order to grasp.* **P1.Q5**: *Positioning the arm using IK was easy for me.*

For phase 2 we asked the following questions: **P2.Q1**: *The alignment of the virtual and real arm was precise.* **P2.Q2**: *The movements of the real arm matched those of the virtual arm.* **P2.Q3**: *Grasping via the virtual arm was intuitive and handling was easy.* **P2.Q4**: *The prior virtual task helped me grasp the ball faster.* **P2.Q5**: *Trajectory planning was easy to perform.* **P2.Q6**: *The shared workspace of the virtual and real arm did not cause any issues.*

For the post-study questionnaire, we asked the following questions: **O.Q1**: *I could quickly understand the modes/menus and knew what was happening.* **O.Q2**: *I felt that my performance improved significantly with more practice.* **O.Q3**: *The labeling and size of the buttons were clear.* **O.Q4**: *I did not have to concentrate hard to avoid errors.* **O.Q5**: *The latency between the virtual and real arm was not disturbing.* **O.Q6**: *I always felt in control of the arm.* **O.Q7**: *The system's limits and capabilities were clearly understandable to me.* **O.Q8**: *I was not afraid of triggering dangerous or unintended movements.*

ACKNOWLEDGMENTS

This work was supported by Festo SE & Co. KG, ebert zobel industrial design gbr. and by the German Research Foundation (DFG) under Germany's Excellence Strategy - EXC 2120/2 - 390831618. The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) and the Alexander von Humboldt Foundation funded by the German Federal Ministry of Research, Science and Space.

REFERENCES

[1] S. Haddadin and E. Croft, "Physical human–robot interaction," in *Springer handbook of robotics*. Springer, 2016, pp. 1835–1874.

[2] A.-H. Chiang and S. Trimi, "Impacts of service robots on service quality," *Serv. Bus.*, vol. 14, no. 3, pp. 439–459, 2020.

[3] A. Baratta, A. Cimino, M. G. Gnoni, and F. Longo, "Human robot collaboration in industry 4.0: a literature review," *Procedia Comput. Sci.*, vol. 217, pp. 1887–1895, 2023.

[4] A. Albu-Schäffer, S. Haddadin, C. Ott, A. Stemmer, T. Wimböck, and G. Hirzinger, "The dlr lightweight robot: design and control concepts for robots in human environments," *Industrial Robot: an international journal*, vol. 34, no. 5, pp. 376–385, 2007.

[5] Z. Ju, C. Yang, and H. Ma, "Kinematics modeling and experimental verification of baxter robot," in *Proceedings of the 33rd Chinese Control Conference*, 2014, pp. 8518–8523.

[6] D. J. Sut and P. Sethuramalingam, "Soft manipulator for soft robotic applications: a review," *J. Int. Robot. Sys.*, vol. 108, no. 1, p. 10, 2023.

[7] W. Dou, G. Zhong, J. Cao, Z. Shi, B. Peng, and L. Jiang, "Soft robotic manipulators: Designs, actuation, stiffness tuning, and sensing," *Adv. Mater. Technol.*, vol. 6, no. 9, p. 2100018, 2021.

[8] K. Hoffmann, D. Müller, R. Simon, and O. Sawodny, "On trajectory tracking control of fluid-driven actuators," *at-Automatisierungstechnik*, vol. 69, no. 11, pp. 970–980, 2021.

[9] A. Das and M. Nabi, "A review on soft robotics: Modeling, control and applications in human-robot interaction," in *ICCCIS*, 2019, pp. 306–311.

[10] M. S. Xavier, C. D. Tawk, A. Fleming, A. Zolfagharian, J. Pinskier, D. Howard, T. Young, J. Lai, S. M. Harrison, and Y. K. Yong, "Soft pneumatic actuators: A review of design, fabrication, modeling, sensing, control and applications," *Open Research Newcastle.*, 2022.

[11] S. Neppalli, M. A. Csencsits, B. A. Jones, and I. D. Walker, "Closed-form inverse kinematics for continuum manipulators," *Advanced robotics*, vol. 23, no. 15, pp. 2077–2091, 2009.

[12] C. Della Santina, C. Duriez, and D. Rus, "Model-based control of soft robots: A survey of the state of the art and open challenges," *IEEE Control Syst.*, vol. 43, no. 3, pp. 30–65, 2023.

[13] M. Azizkhani, S. Kousik, and Y. Chen, "Dynamic task space control of redundant pneumatically actuated soft robot," *IEEE Robot. Autom. Lett.*, vol. 10, no. 6, pp. 6408–6415, 2025.

[14] F. Renda, C. Armanini, A. Mathew, and F. Boyer, "Geometrically-exact inverse kinematic control of soft manipulators with general threadlike actuators' routing," *IEEE Robot. Autom. Lett.*, vol. 7, no. 3, pp. 7311–7318, 2022.

[15] Y. Wang, Z. Wu, L. Wang, B. Feng, and K. Xu, "Inverse kinematics and dexterous workspace formulation for 2-segment continuum robots with inextensible segments," *IEEE Robot. Autom. Lett.*, vol. 7, no. 1, pp. 510–517, 2022.

[16] S. Pilch, D. Klug, and O. Sawodny, "Optimal interaction-friendly trajectory generation of a pneumatic soft continuum manipulator," in *AIM*, 2024, pp. 1189–1196.

[17] J. M. Bern, W. C. May, A. Osborn, F. Stella, S. Zargarzadeh, and J. Hughes, "A soft robot inverse kinematics for virtual reality," in *IEEE ICRA*. IEEE, 2024, pp. 14 957–14 963.

[18] H. Liu, D. Qu, F. Xu, Z. Du, K. Jia, and M. Liu, "An efficient online trajectory generation method based on kinodynamic path search and trajectory optimization for human-robot interaction safety," *Entropy*, vol. 24, no. 5, p. 653, 2022.

[19] M. Rettinger, L. Hacker, P. Wolters, and G. Rigoll, "Optimizing robot programming: Mixed reality gripper control," in *ICRA*. IEEE, 2025, pp. 4768–4774.

[20] L. Morando and G. Loianno, "Spatial assisted human-drone collaborative navigation and interaction through immersive mixed reality," in *ICRA*. IEEE, 2024, pp. 8707–8713.

[21] F. Kennel-Maushart, R. Poranne, and S. Coros, "Interacting with multi-robot systems via mixed reality," in *ICRA*, 2023, pp. 11 633–11 639.

[22] J. Lee, T. Lim, and W. Kim, "Investigating the usability of collaborative robot control through hands-free operation using eye gaze and augmented reality," in *IROS*. IEEE, 2023, pp. 4101–4106.

[23] V. Patil and M. Hutter, "Radiance fields for robotic teleoperation," in *IROS*. IEEE, 2024, pp. 13 861–13 868.

[24] S. Pilch, A. Beger, and O. Sawodny, "Kinematic modeling and control of a pneumatic 3d-printed spatial soft manipulator using configuration and actuator models," *IFAC-Pap.*, vol. 59, no. 30, pp. 144–149, 2025, 5th Conference on Modeling, Estimation and Control MECC 2025.

[25] S. Pilch, E. Menrad, A. Beger, and O. Sawodny, "Dynamic modeling and control of an inextensible pneumatically actuated soft continuum manipulator," *Mechatronics*, vol. 115, p. 103462, 2026.

[26] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.