Towards Assistive Teleoperation for Knot Untangling

Berk Güler[1,2], Kay Pompetzki[1], Simon Manschitz[2], Jan Peters[1,3,4,5]

***Abstract*— Manipulating deformable linear objects (DLOs) such as ropes is challenging due to their complex dynamics. To address these issues, we present a novel assistive teleoperation framework that combines human expertise with autonomous assistance. Our approach integrates a vision-based module to identify grasp poses, a shared autonomy mechanism that balances human input with autonomous guidance, and an optimization-based inverse kinematic solver for smooth, collision-free manipulation. Additionally, a virtual reality (VR) interface provides intuitive control and real-time feedback to the operator. A user study on knot untangling under time-delayed and non-delayed conditions shows that shared autonomy enhances task performance under delay while reducing the operator's physical and mental workload. These findings highlight the potential of shared autonomy to improve teleoperation systems for complex DLO manipulation, particularly in environments affected by communication delays or uncertainties.**

I. Introduction

Manipulating deformable objects remains one of the most critical challenges in robotics. In particular, deformable linear objects (DLOs) are a specific subset characterized by their one-dimensional structures, infinite degrees of freedom, and tendency to form self-intersections. These properties make DLOs particularly challenging for perception, planning, and control, as their configurations are highly sensitive to small perturbations. Existing manipulation strategies for DLO manipulation tasks such as knot untangling typically rely on a set of predefined motions and heavily depend on their perception pipelines, making them susceptible to variations in configuration, material, and texture [1], [2], [3], [4]. These challenges indicate that complex manipulation tasks like knot untangling can be addressed by integrating human's intelligence and decision-making capability, particularly when preplanned strategies are prone to failure or when operating in critical environments.

Teleoperation offers a promising solution by leveraging human situational awareness and planning capabilities. However, teleoperated robotic systems face their own set of challenges, including limited scene visibility, embodiment mismatch, and communication delays, which can degrade the performance of the human operator [5]. To address these issues, our work proposes a context-aware shared autonomy

[1]TU Darmstadt, Institute for Intelligent Autonomous Systems, Computer Science Department, 64289 Darmstadt, Germany berk@robot-learning.de

[2]Honda Research Institute Europe GmbH, Carl-Legien-Straße 30, 63073 Offenbach/Main, Germany

[3]German Research Center for AI (DFKI), Research Department: Systems AI for Robot Learning

[4]Hessian.AI

[5]Centre for Cognitive Science, Hochschulstr. 10, 64293 Darmstadt

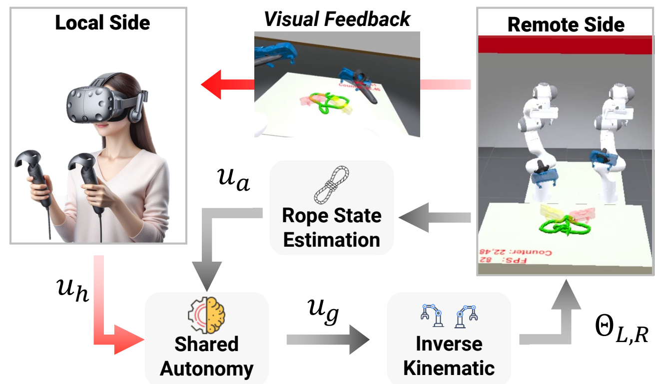

Fig. 1. System architecture integrating perception, optimization, and shared autonomy modules. The arbitration mechanism adjusts autonomy levels by blending human input $\boldsymbol{u}_h$ and vision-based autonomy input $\boldsymbol{u}_a$ to compute the target pose $\boldsymbol{u}_g$. This pose, including position and orientation in task space, is then processed by a non-linear optimization-based inverse kinematics solver to generate joint angles $\Theta_{L,R}$ for the left and right robots.

framework, depicted in Figure 1, that preserves human intentionality while compensating for latency-induced performance degradation. To assess the efficacy of the proposed framework, we conducted experiments with five participants and measured how both time delay and shared autonomy influence teleoperation performance during a knot-untangling task.

The key contributions of our paper include 1) Real-time estimation of grasping positions for DLOs using a vision-based state estimator, 2) Generation of collision-free and smooth motions through a constrained optimization-based inverse kinematics solver, and 3) Adaptive blending of human inputs with autonomous grasping suggestions to achieve shared autonomy.

II. Related Work

Teleoperation for complex manipulation tasks has attracted considerable attention, particularly in scenarios where fully autonomous strategies fall short. In this section, we briefly review prior work on shared autonomy in teleoperation and vision-based state estimation for DLOs.

Shared Autonomy in Teleoperation Shared autonomy (or shared control) frameworks balance human input with autonomous assistance to enhance task performance while preserving the operator's sense of control [6]. Several approaches have focused on estimating human intention and blending it with robotic control. For example, research leveraged Gaussian Hidden Markov Models with gaze input [7] to infer user intent and employed behavioral state machines in virtual reality (VR) setups for remote bimanual operations [8]. Additionally, methods based on Partially

Observable Markov Decision Processes (POMDPs) have been used to optimize teleoperation under uncertainty [9]. Although some teleoperation frameworks have addressed DLOs manipulation in structured industrial settings [10], they did not address possible unpredictable configurations and non-linear dynamics as in knot untangling.

Vision-Based State Estimation for DLOs Accurate state estimation is critical for manipulating DLOs. Existing algorithms such as HANDLOOM [3], FastDLO [11], RT-DLO [12], and mBest [13] typically rely on fixed or simplified camera setups and make assumptions about object geometry. While effective in controlled environments, these methods often struggle with complex scenarios, such as knots, where self-intersections (occlusion) and dynamic deformations occur. In contrast, our approach employs wrist-mounted cameras to capture dynamic viewpoints, allowing robust vision-based estimation even in the presence of occlusions and significant deformations.

Together, these studies highlight the challenges of relying solely on autonomous systems for complex DLO manipulation, and they motivate our integrated teleoperation framework that combines shared autonomy with robust vision-based state estimation.

III. Knot Untangling For Assistive Teleoperation

In this section, we introduce our assisted teleoperation framework, which consists of an inverse kinematic solver, rope state estimator, and shared autonomy blending mechanism.

Optimization-based Inverse Kinematics: Building on the RelaxedIK solver proposed by Rakita et al. [14], our system employs a constrained optimization formulation to compute collision-free, smooth motions for 7-DoF robotic arms. This optimization minimizes a weighted sum of task-specific objectives, such as motion smoothness, self-collision and singularity avoidance, while enforcing joint limits and ensuring sufficient separation between the arms.

Rope State Estimation: Our system uses wrist-mounted cameras to capture images of the rope, which are segmented using FastSAM with a text prompt (e.g., "rope") [15]. The initial segmentation is refined through morphological processing. Contours are then extracted from the segmented masks, and a Voronoi-based graph is constructed to identify key structural points (intersections and terminal nodes). Rope segments are defined between these key points; the center node of each segment provides the grasping pixel, while the tangent at that pixel determines the grasping angle. Finally, by converting from pixel space to task space, we obtain the grasp suggestions for rope segments as an autonomous input $\boldsymbol{u}_a$.

Shared Autonomy Blending: In our framework, the human input $\boldsymbol{u}_h$ and the autonomous suggestion $\boldsymbol{u}_a$ consist of both position and orientation components, i.e., $\boldsymbol{u}_h = (\boldsymbol{p}_h, \boldsymbol{o}_h)$ and $\boldsymbol{u}_a = (\boldsymbol{p}_a, \boldsymbol{o}_a)$. However, the arbitration process is applied only to the positional component. We dynamically blend the two position inputs using the arbitration equation

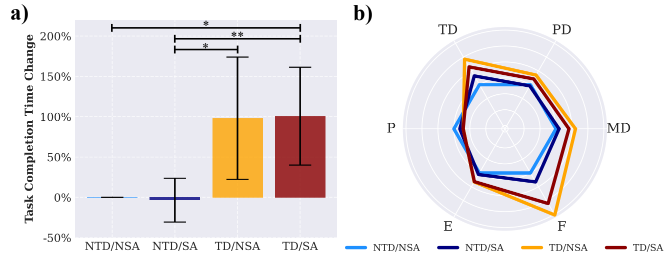

Fig. 2. Performance metrics in teleoperation with conditions: No Time Delay/No Shared Autonomy (NTD/NSA), No Time Delay/Shared Autonomy (NTD/SA), Time Delay/No Shared Autonomy (TD/NSA), and Time Delay/Shared Autonomy (TD/SA). (a) Percent change in Task Completion Time relative to baseline (NTD/NSA), higher values indicate longer completion times. (b) NASA Task Load Index relative changes from baseline across Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (P), Effort (E), and Frustration (F).

$$\boldsymbol{p}_g = \alpha\, \boldsymbol{p}_h + (1-\alpha)\, \boldsymbol{p}_a,$$

where $\boldsymbol{p}_h$ is the human-specified target position and $\boldsymbol{p}_a$ is the suggested position derived from the vision pipeline. The arbitration factor $\alpha \in [0, 1]$, computed via a sigmoid function based on the proximity between $\boldsymbol{p}_h$ and $\boldsymbol{p}_a$, determines the balance of authority [6]. The final target pose is then given by $\boldsymbol{u}_g = (\boldsymbol{p}_g, \boldsymbol{o}_h)$, meaning that the orientation remains entirely under human control.

IV. Experiments

We evaluated our teleoperation framework for rope untangling using two Franka Emika Panda arms in a simulated environment. Five participants performed the knot untangling task under four conditions, varying the presence of time delay (2 seconds bidirectional communication delay) and shared autonomy, while task completion time (TCT) and NASA TLX workload metrics were recorded. The main task required the participants to untangle the rope by manipulating the robot arms until the rope was free of self-intersections.

The results indicate that communication delay increases TCT and mental, physical, and temporal demands, leading to higher effort and frustration and reducing perceived task success, as shown in Figure 2. Shared autonomy helped mitigate these negative effects under time delay, although its benefits varied among participants and sometimes increased complexity during real-time operations. Task completion times varied based on individual differences in how participants approached and solved the task.

V. Conclusion

We introduced an assistive teleoperation and simulation framework integrating a vision-based rope state estimator, an optimization-based inverse kinematics solver, and a shared autonomy module to support DLO manipulation. The study demonstrated that while shared autonomy can compensate for the adverse effects of communication delays, its impact depends on the user's interaction style.

Future work will focus on refining the vision-based rope state estimator, enhancing the shared autonomy interface by integrating human intention estimation over multiple goals and conducting comprehensive real-world robot experiments.

REFERENCES

[1] J. Grannen, P. Sundaresan, B. Thananjeyan, J. Ichnowski, A. Balakrishna, V. Viswanath, M. Laskey, J. E. Gonzalez, and K. Goldberg, "Learning Robot Policies for Untangling Dense Knots in Linear Deformable Structures," 2020.

[2] K. Shivakumar, V. Viswanath, A. Gu, Y. Avigal, J. Kerr, J. Ichnowski, R. Cheng, T. Kollar, and K. Goldberg, "SGTM 2.0: Autonomously Untangling Long Cables using Interactive Perception," Sep. 2022.

[3] V. Viswanath, K. Shivakumar, M. Parulekar, J. Ajmera, J. Kerr, J. Ichnowski, R. Cheng, T. Kollar, and K. Goldberg, "HANDLOOM: Learned Tracing of One-Dimensional Objects for Inspection and Manipulation," 2023.

[4] A. Caporali, P. Kicki, K. Galassi, R. Zanella, K. Walas, and G. Palli, "Deformable linear objects manipulation with online model parameters estimation," *IEEE Robotics and Automation Letters*, vol. 9, no. 3, pp. 2598–2605, 2024.

[5] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, no. 6, pp. 1231–1245, 2007.

[6] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *The International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013.

[7] S. Fuchs and A. Belardinelli, "Gaze-Based Intention Estimation for Shared Autonomy in Pick-and-Place Tasks," *Frontiers in Neurorobotics*, vol. 15, p. 647930, Apr. 2021.

[8] A. D. Dragan, S. Siddhartha Srinivasa, and K. Kenton Lee, "Teleoperation with Intelligent and Customizable Interfaces," *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 33–79, Jun. 2013.

[9] S. Javdani, H. Admoni, S. Pellegrinelli, S. S. Srinivasa, and J. A. Bagnell, "Shared Autonomy via Hindsight Optimization for Teleoperation and Teaming," May 2017.

[10] D. Chiaravalli, A. Caporali, A. Friz, R. Meattini, and G. Palli, "A vision-based shared autonomy framework for deformable linear objects manipulation," in *2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2023, pp. 733–738.

[11] A. Caporali, K. Galassi, R. Zanella, and G. Palli, "FASTDLO: Fast Deformable Linear Objects Instance Segmentation," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 9075–9082, Oct. 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9830852/

[12] A. Caporali, K. Galassi, B. L. Žagar, R. Zanella, G. Palli, and A. C. Knoll, "Rt-dlo: Real-time deformable linear objects instance segmentation," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 11, pp. 11 333–11 342, 2023.

[13] A. Choi, D. Tong, B. Park, D. Terzopoulos, J. Joo, and M. K. Jawed, "mBEST: Realtime Deformable Linear Object Detection Through Minimal Bending Energy Skeleton Pixel Traversals," *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 4863–4870, Aug. 2023, arXiv:2302.09444 [cs]. [Online]. Available: http://arxiv.org/abs/2302.09444

[14] D. Rakita, B. Mutlu, and M. Gleicher, "RelaxedIK: Real-time Synthesis of Accurate and Feasible Robot Arm Motion," in *Robotics: Science and Systems XIV*. Robotics: Science and Systems Foundation, Jun. 2018.

[15] X. Zhao, W. Ding, Y. An, Y. Du, T. Yu, M. Li, M. Tang, and J. Wang, "Fast Segment Anything," Jun. 2023, arXiv:2306.12156 [cs]. [Online]. Available: http://arxiv.org/abs/2306.12156