题目 Title

Title: Adaptive Visuotactile-Guided Control for Dexterous In-Hand Rolling of Slender Objects

摘要 Abstract

The robotic manipulation of slender objects is fundamentally limited by a trade-off between precision and versatility. Specialized grippers provide sub-millimeter accuracy but are rigid by design, failing in dynamic, multi-object workflows. Conversely, general-purpose grippers offer adaptability but lack stable contact and fine force control essential for delicate tasks. This leaves a critical class of dexterous, multi-stage manipulations beyond the reach of automation. We resolve this dilemma with an Adaptive Visuotactile-Guided Control framework. Our system features a multi-DoF, sensorized gripper that performs force-position coupled in-hand rolling. It robustly manipulates a continuous range of slender objects, from sub-millimeter fibers to 5mm rods, with a single end-effector. Experimental results demonstrate that our approach successfully combines precision and versatility, enabling the automation of complex, multi-object manipulation tasks.

介绍 Introduction

In-Hand Rolling

Dexterous Fingertip Manipulation: Discuss prior research on robotic finger/gripper design for fine manipulation. What have others done? What are their limitations

Dexterous manipulation of slender objects—like fine wires, optical fibers, or surgical sutures—poses a grand challenge in robotics. Mastering this class of objects is a critical step toward the next generation of automation in high-stakes domains, from microsurgery to the assembly of intricate electronics. Unlike rigid bodies, the dynamics of slender objects are complex and unpredictable, demanding a level of finesse that lies at the frontier of modern robotic perception and control. Achieving human-like dexterity in these tasks promises to unlock new surgical procedures, accelerate scientific discovery, and create truly adaptable manufacturing lines.

Yet, current robotic systems present a stark trade-off between precision and versatility. On one hand, specialized mechanisms can achieve sub-millimeter accuracy for a single, well-defined task. However, these systems are inherently brittle, often failing at the slightest deviation from their programmed function. On the other hand, general-purpose grippers offer adaptability but typically lack the stable contact and delicate force control needed for fine-scale work. They treat a fragile fiber like a blunt instrument, leading to unstable grasps and rendering them clumsy and unsafe for procedures where dexterity is paramount.

To bridge this gap, an ideal system must unite three capabilities that are currently separate. First, it requires stable, continuous contact to execute controlled motions like rolling without slippage. Second, it needs high-resolution tactile feedback to perceive and regulate minute interaction forces, preventing damage to the delicate object or its surroundings. Third, it must possess intrinsic in-hand dexterity to adapt to various object sizes and perform complex motions on the fly. The effective integration of these three capabilities remains an open and critical research gap.

Visuotactile Sensing

Visuotactile Sensing for Control: Survey the different types of visuotactile sensors (like GelSight, TacTip, etc.) and how they have been used in control loops. What information can they provide? What challenges remain in using them for this specific task?

Main Contributions

To fill this gap, we present an Adaptive Visuotactile-Guided Control framework. Our approach uses a novel multi-DOF gripper with visuotactile sensors to gain fine-grained control over an object's in-hand pose while regulating interaction forces. This tight coupling of rotational displacement and applied force is the key to dexterous in-hand rolling. In this paper, we make three primary contributions:

A novel, dual-DOF dexterous fingertip mechanism, integrated into a parallel gripper and designed specifically for in-hand rolling.

A coupled force-pose control algorithm that leverages visuotactile feedback for robust manipulation of sub-millimeter objects.

Experimental validation on a simulated surgical platform, demonstrating the system's effectiveness in a complex, multi-object task.

图1 论文大图

Robot Structures

The experimental platform consists of a 6-DOF Universal Robots UR7e arm for macro-positioning, equipped with a custom-designed dexterous end-effector for in-hand rolling manipulation. The system architecture integrates custom mechanics, commercial sensing, and hierarchical control.

A. Dexterous Rolling End-Effector Mechanics

The core novelty of the end-effector lies in its integration of independent grasping and rolling degrees of freedom (DOF), enabling decoupled control of object clamping and manipulation.

Grasping Actuation: The primary grasping DOF is a parallel mechanism actuated by a Damiao DM-J4310 brushless servo motor via a rotational-to-linear transmission. This provides a precise, high-stiffness clamping force essential for stable manipulation.

Rolling Actuation: The in-hand rolling capability is achieved through two independent linear stages, one integrated into each finger assembly. These stages are driven by 28mm linear stepper motors which translate the fingertips tangentially relative to the grasped object. Coordinated differential motion of these stages generates precise rotational torque. The steppers are controlled by a Wheeltec D36A dual-path driver module interfaced with an Arduino Nano.

Critical structural components are fabricated from CNC-machined steel to maximize rigidity. This design choice prevents torsional flexion and non-parallel alignment of the fingers when gripping forces are applied. Minimizing this unmodeled mechanical compliance is essential for maintaining consistent contact geometry and ensuring sensor data fidelity.

B. Visuotactile Sensing Subsystem

Each fingertip integrates a commercial Daimon Robotics DM Tac W visuotactile sensor. The sensors provide high-resolution (320x240 pixels) imagery of the contact interface over a 24mm x 18mm area. While the sensor hardware captures data at 120Hz, the final application-level processing rate for extracting feedback varies with computational load. The system typically processes feedback at 45 FPS, dropping to around 15 FPS during more complex control loop iterations. This feedback stream provides real-time data on contact geometry, shear force distribution, and incipient slip for the control algorithm detailed in proceeding section.

图2 结构设计

Methodology

In-Hand Rolling of Slender Objects

Before detailing our model, it is important to frame its contribution in a practical context. The operation of slender, flexible objects often relies on an operator's intuition, guided by visual feedback from the object's distal tip. Due to the object's compliance, rotational manipulation applied locally by a gripper does not necessarily translate directly to the pose of this tip.

Our work, therefore, does not attempt to model the complex dynamics of the entire flexible body. Instead, we focus on a foundational prerequisite: ensuring the robustness and consistency of the manipulation action itself. The goal of our control algorithm is to precisely model and control the local rotational displacement at the point of contact. By doing so, we provide a predictable and repeatable "building block" for manipulation. The broader challenge of modeling how this local rotation affects the final tip pose is considered beyond the scope of this paper.

Regime 1: Kinematic Rolling of Thicker Objects (d≥2 mm)

For objects with a diameter of 2 mm or greater, the gripper fingers maintain a separation gap even after a secure grasp is established. In this configuration, we can model the system as the non-slip rolling of a rod between two parallel racks, similar to a rack-and-pinion mechanism.

The theoretical rotational displacement is derived from the linear travel of the fingertips. Given a linear travel per motor step, s, the total effective displacement generated by the two opposing fingers for n steps is 2sn. The ideal angular displacement in degrees is thus:

However, this ideal model must be corrected for two real-world factors. First, any angular misalignment, ϕ, between the gripper and the object's axis means that only the perpendicular component of the finger motion contributes to the rotational torque. This is accounted for by a cos(ϕ) term. Second, to compensate for unmodeled effects such as material compliance and mechanical friction, we introduce an empirically determined correction factor, K rod.

This leads to the final model for this regime:

Regime 2: Deformation-Based Rolling of Thinner Objects (d≤1 mm)

When manipulating sub-millimeter objects, the dynamic model changes significantly. The gripper fingers close completely, making contact with each other and enclosing the slender object within their soft, deformable pads. The physics of this state are far more complex, dominated by material compression and frictional forces.

To create a tractable model, we narrow the scope to a non-slip scenario for a limited range of motion (n≤30 steps). In this case, we observed experimentally that the object's rotation is primarily caused by the shearing deformation of the soft finger material, rather than kinematic rolling. Our data revealed a strong linear relationship between the number of motor steps and the resulting rotation, with the object's specific diameter having a statistically insignificant effect.

This relationship is captured by the following linear model:

where m wire is the empirically derived gain (deg/step) and C wire is a constant offset that accounts for system backlash or initial material slack. While this empirical model is effective for control, its relationship to the system's underlying kinematics is non-obvious. We analyze this behavior in detail in our discussion.

Unified Rotational Pose Model

Combining these two models yields the final piecewise formulation for our pose control algorithm:

where the constants are empirically determined as {provide specific value}. This hybrid model allows our controller to adapt its strategy based on the object being manipulated, forming the foundation of our system's versatility.

图. 实验测试数据

仿真分析

图. 仿真

视触觉力标定

实验 Experiments

针灸实验（证明的是在指尖完全闭合状态下，对极细物体的强力抓握能力）

绞线实验（证明是在指尖完全闭合状态下，极细物体-多股线不同角度-在指尖的共同旋转能力）

血管介入实验（证明夹爪的功能性和实用性）

To demonstrate the versatility and safety of our system in a clinically-relevant context, we designed a manipulation task using a vascular phantom. The objective is to show that a single gripper can seamlessly manipulate multiple collaborative instruments—a catheter and a guidewire—that are often used in tandem. The task is performed in two variations: (1) with the catheter and guidewire located in separate entry points, and (2) with the guidewire pre-inserted into the catheter.

Manipulation Workflow

The manipulation strategy is executed entirely via keyboard-based teleoperation. To simplify control for the operator, the robot arm's movement is constrained to forwards and backwards motion along pre-defined paths. The gripper's discrete actions, such as grasping, releasing, and in-hand rolling, are also mapped to keyboard commands. The generalized workflow is as follows:

Approach Catheter: The operator moves the gripper along a set path until it reaches the catheter.

Grasp and Roll Catheter: The operator initiates a grasp. Once a stable grip is confirmed by the visuotactile sensors, the operator commands an in-hand rolling motion to demonstrate successful manipulation.

Release and Transition: The operator releases the catheter and then pilots the arm along another pre-defined path to the guidewire's location.

Grasp and Roll Guidewire: The operator grasps the guidewire and performs another rolling manipulation.

This sequence demonstrates the core capability of our system: using a single, versatile end-effector to perform dexterous in-hand manipulation on multiple objects of varying size and properties within the same workspace, all under the control of a human operator.

Visuotactile-Based Safety Monitoring

Safety is ensured by leveraging the rich shear data from the visuotactile sensors. The sensor provides a real-time shear-force vector field, visualized as a mesh of arrows representing the magnitude and direction of deformation on the gripper's soft surface.

This feedback is integrated into a safety protocol:

Force Calibration: Through prior offline calibration, we created a mapping from the shear vector magnitudes to a specific force value (in Newtons).

Real-Time Monitoring: During manipulation, the system continuously processes this data to monitor the interaction forces.

Threshold-Based E-Stop: We define a maximum force-limit threshold. If the measured force exceeds this threshold—indicating a potential snag, excessive compression, or unsafe condition—a controlled emergency stop is triggered, immediately halting the manipulation to prevent damage to the objects or the phantom environment. This closed-loop monitoring is active during all phases of the task.

讨论 Discussion

结论 Conclusion