Geometrically Consistent Tactile Servoing via Hybrid Force–Position Control at the Center of Pressure

Sébastien Kleff, Lucas Joseph, Vincent Padois

Geometrically Consistent Tactile Servoing via Hybrid Force–Position Control at the Center of Pressure

Sébastien Kleff[1], Lucas Joseph[1], Vincent Padois[1]

***Abstract*— In robotics, traditional force control lacks local contact information. Tactile sensors provide rich feedback on physical interaction, but remain difficult to integrate consistently into real-time control loops. In this paper, we show that hybrid force–position control, when expressed at the Center of Pressure (CoP), becomes geometrically consistent with tactile sensing. Based on this insight, we formulate a tactile servoing approach that allows explicit control of contact pose and force at the CoP. Unlike conventional tactile servoing techniques, which are tightly coupled to a specific sensor and often treat the contact wrench as a disturbance to be rejected, our approach relies on a generic and physically grounded feature space. We derive a hybrid force-position control law based on the Jacobian at the CoP that naturally decouples force and motion subspaces, ensuring geometric consistency during contact interaction. Real-world experiments on a robotic manipulator demonstrate robust contact maintenance and improved force tracking compared to standard image-based and pose-based controllers in a physical interaction regulation task.**

I. Introduction

Robots operating in unstructured environments must regulate both contact forces and contact geometry while interacting with uncertain or changing surfaces (e.g., polishing curved surfaces, exploring object contours, or dis/assembling tight-fitting components). Humans achieve this seamlessly through tactile perception, continuously adjusting motion and force at the contact interface [1]. While endowing robots with this capability remains an open problem, recent progress in tactile sensing offers a promising way to provide robots with rich geometric and force cues during physical interaction [2].

In particular, vision-based tactile sensors (VBTS) [3]–[7] provide high-resolution local information at the contact interface, including contact location, surface orientation and pressure distribution. By contrast, wrist-mounted force/torque (F/T) sensors used in classical force control provide only a resultant wrench at the wrist, which makes it difficult to infer the local contact state, especially when different contact geometries produce similar wrench measurements [8].

More generally, regulating contact geometry alone does not guarantee regulation of the contact force: for compliant tactile interfaces, changes in contact area or local stiffness can modify the normal force even when the nominal contact pose is maintained. VBTS are therefore especially attractive for tasks in which contact geometry must be controlled together with force, such as surface following and exploration. But despite their increasing availability, integrating rich tactile measurements consistently and effectively into real-time control remains a fundamental challenge. In recent years, tactile control methods have enabled improved contact regulation by exploiting information available directly at the contact interface, often by drawing inspiration from visual servoing. This is done either by regulating low-level tactile image features [9], [10] or by estimating the contact pose from tactile data [11]. These approaches, however, typically lack physical grounding, which leads to control laws that are either purely kinematic or tightly coupled to a specific sensor and task. In practice, this induces parasitic couplings between force and motion, restricting stability margins and limiting the ability to perform precise, dynamic interaction.

In this paper, we address these challenges by revisiting tactile servoing through the lens of classical hybrid force–position control. Our central insight is that hybrid force-position control should be expressed at the Center of Pressure (CoP) to remain geometrically consistent with tactile sensing – an insight that, to our knowledge, has not been formalized in prior tactile servoing literature. Based on this insight, we propose a CoP-centered tactile servoing formulation that integrates tactile and force feedback within a single, physically grounded control structure. Our main contributions are:

- **The geometric insight** that expressing the control task at the CoP removes parasitic coupling terms and yields a geometrically consistent formulation of tactile servoing.
- **A practical sensor-agnostic instantiation** of this insight, combining CoP kinematics with explicit force–position regulation and modular perception pipelines compatible with both learned and analytical estimators.
- **Real-world experimental validation** in a translational contact-following task, demonstrating improved force control, contact stability, and robustness to perturbations compared to existing methods.

By defining the interaction at the CoP, our formulation decouples control from perception and naturally accommodates heterogeneous sensor fusion. We show this by combining local pose information from a VBTS with high-bandwidth force feedback from an F/T sensor.

II. Related Work

Despite the substantial body of research on tactile sensing, the *systematic* integration of tactile feedback into model-based control frameworks has received disproportionately less attention. This gap is largely driven by hardware landscape fragmentation: the vast diversity of sensing technologies (capacitive arrays, optical elastomers, etc.) produces heterogeneous tactile information that lacks the standardized

[1] Inria, AUCTUS team, Talence, France

representations of conventional robotics modalities like vision or proprioception.

A. Force control

Classical hybrid force–position control [12] and operational-space formulations [13] provide the theoretical foundation for regulating motion and wrench simultaneously. These methods are highly effective for tasks where the global contact wrench at the end-effector is solely sufficient (e.g., peg-in-hole with rigid tools, contour following with known geometry). However, they lack critical local information about the contact interface (contact point location, contact patch, pressure distribution). Recent efforts have integrated tactile sensors to improve the estimation of the contact state. [14] fuses tactile array signals with wrist F/T measurements to recover distributed contact locations and forces, improving estimation accuracy for whole-body control. Similarly, [9] and [15] regulate specific contact features (e.g., pressure centroids) using task-dependent projection matrices. Most recently, [16] achieved 3D force regulation using magnetic tactile sensors, demonstrating that accurate wrist-independent force feedback is possible.

Overall, while these methods leverage local tactile information to improve global force estimation, they do not co-locate the control task with the tactile measurements. Consequently, they fail to decouple contact force and motion at the local level, preventing their simultaneous, explicit regulation.

B. Tactile Servoing

Tactile servoing [9], [17] is a class of closed-loop control strategies that regulate robot motion based on tactile feedback obtained during physical interaction. Analogous to visual servoing, tactile servoing usually extracts a set of tactile features, ranging from raw pressure distributions to estimated contact poses, and computes a control action that drives these features toward a desired tactile goal.

1) Image-Based Tactile Servoing (IBTS): IBTS frameworks [9], [10] map low-level image features (pressure centroids, marker displacements) to the robot end-effector motions through carefully designed interaction matrices. While effective for maintaining contact, these methods treat interaction forces as implicit consequences of image features regulation, which prevents the independent specification of force and motion targets, and relies on task- and contact-specific projections.

2) Pose-Based Tactile Servoing (PBTS): Recent work shifted toward PBTS [11], leveraging vision-based tactile sensors and data-driven models (CNNs) to estimate the contact pose, and even tangential shear [18]. These estimates feed standard Cartesian controllers for surface-following and object pushing tasks. However, PBTS typically regulates contact pose rather than contact force explicitly. As a result, tangential shear and normal force are handled only indirectly through the pose controller and the compliance of the tactile interface, which can be sufficient in some settings but limits direct specification of force targets.

Beyond these limitations, existing tactile servoing methods are tightly coupled to the tactile sensor.

C. Learning-Based Tactile Policies

A parallel body of work leverages data-driven methods to bypass analytical modeling. Approaches ranging from tactile MPC [19]–[23] to representation learning [24]–[26] successfully enable robust manipulation by learning interaction dynamics implicitly. Recent works like SaTA [27] have attempted to recover geometric consistency by spatially anchoring these features. However, these methods generally operate in abstract feature spaces optimized for task success and lack physical interpretability. Specifically, they do not expose contact wrench required for explicit force regulation, which prevents their application in safety-critical applications (e.g. human-robot collaboration).

While learned policies excel at robustness, we argue that explicit physical grounding offers a complementary layer of precision. We propose to bridge this gap by defining a structured state space (CoP and wrench) that allows to decouple estimation from control: data-driven estimation pipelines [28]–[30] can be used to inform our controller.

III. Methodology

This section formalizes the key limitation of related works (Sec. III-A), defines the problem geometry in the CoP frame (Sec. III-B), derives the hybrid control law that exploits this representation (Sec. III-C), and describes a perception pipeline instantiating the required contact state (Sec. III-D)

A. Motivation: Decoupling Interaction Dynamics

The limitations identified in Sec. II-A,B stem from a fundamental geometric inconsistency: the mismatch between the control frame (typically the sensor origin) and the physical interaction interface. We argue that local hybrid force-position control at the contact interface is inconsistent due to kinematic couplings between force and motion when the contact point is offset from the sensor origin.

Let us define the interaction between the robot and its environment using two primary reference frames: the *sensor frame* $\{S\}$ and the *contact frame* $\{C\}$. The sensor frame is rigidly attached to the tactile sensor housing with its origin $\mathbf{o}$ at the geometric center of the sensing surface and its z-axis aligned with the outward-pointing surface normal. The contact frame is a floating frame whose origin coincides with the effective contact point $\mathbf{p}$ to be specified, while its orientation remains aligned with $\{S\}$[1] (Fig. 1).

Notation. A left superscript indicates the frame of expression. For twists and wrenches, the right subscript denotes the point of action, e.g. ${}^{S}\boldsymbol{\nu}_{\mathbf{o}}$ is the body twist taken at $\mathbf{o}$ and expressed in $\{S\}$ and ${}^{C}\mathbf{w}_{\mathbf{p}}$ is the wrench reduced at p and expressed in $\{C\}$. Unless otherwise stated, omitted left superscripts should be understood as referring to $\{S\}$.

[1]The validity and implications of this assumption are discussed in Sec. V.

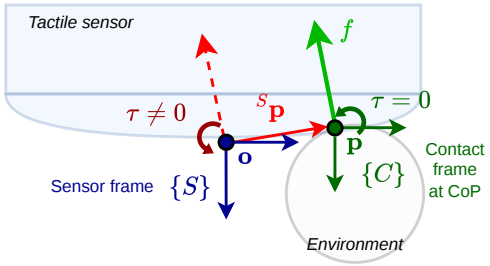

Fig. 1: When force is commanded at the sensor origin (a), the offset ${}^S\mathbf{p}$ induces a parasitic moment τ. Controlling force at the CoP (b) eliminates this coupling by formulating the task where forces actually act.

When commanding a twist ${}^S\boldsymbol{\nu}_\mathbf{o} = [{}^S\mathbf{v}_\mathbf{o}^\top, {}^S\boldsymbol{\omega}^\top]^\top$ at the sensor origin, the instantaneous velocity at the contact point ${}^S\mathbf{p} \triangleq {}^S\overrightarrow{\mathbf{op}}$ is given by the rigid-body kinematic relationship:

$$ {}^S\mathbf{v}_\mathbf{p} = {}^S\mathbf{v}_\mathbf{o} + {}^S\boldsymbol{\omega} \times {}^S\mathbf{p} \tag{1} $$

A pure rotational command (${}^S\mathbf{v}_o = \mathbf{0}$, ${}^S\boldsymbol{\omega} \neq \mathbf{0}$) intended to adjust sensor alignment inadvertently induces tangential motion ${}^S\boldsymbol{\omega} \times {}^S\mathbf{p}$ at the contact. This tangential motion can generate tangential shear and force spikes at the interface, even though the control intent was purely rotational adjustment of the sensor frame. Whether shear actually leads to slip depends on the friction coefficient, the sensor compliance, and the contact load. This motivates expressing the control task at the contact point rather than at the sensor origin. Conversely, consider a contact force ${}^S\mathbf{f}$ acting at the contact point ${}^S\mathbf{p}$. The wrench measured at the sensor origin is:

$$ {}^S\mathbf{w}_\mathbf{o} = \begin{bmatrix} {}^S\mathbf{f} \\ {}^S\mathbf{p} \times {}^S\mathbf{f} \end{bmatrix} \tag{2} $$

The cross-product term ${}^S\mathbf{p} \times {}^S\mathbf{f}$ is a parasitic moment arising purely from the measurement location: it does not reflect the actual contact phenomena but rather a geometric artifact of measuring away from the contact point. Attempting to regulate the contact wrench using ${}^S\mathbf{w}_\mathbf{o}$ means we are implicitly regulating a coupled wrench that conflates true contact forces with geometric offset effects (Fig. 1).

To resolve this, we must identify a control frame where these coupling terms vanish. In contact mechanics, this unique point $\mathbf{p}$ is known as the Center of Pressure (CoP). Defined as the centroid of the contact pressure distribution, the CoP is the specific point on the contact interface where the resultant force acts with zero tangential moment [31]. By shifting the control frame to the CoP, we align the mathematical control objective with the physical interaction, ensuring that force regulation commands do not generate tangential moments (Fig. 1). This property is extensively exploited in locomotion stability [32] and more recently in planar contacts modeling [33].

B. CoP-based Tactile Feature Space

Following the geometric consistency requirement established in Sec. III-A, we define the origin of the contact frame $\{C\}$ at the Center of Pressure (CoP). Physically, the CoP corresponds to the centroid of the contact patch weighted by the normal pressure distribution P:

$$ {}^S\mathbf{p}_{\text{CoP}} = \frac{\int_\mathcal{C} \mathbf{r}\, P(\mathbf{r})\, dA}{\int_\mathcal{C} P(\mathbf{r})\, dA} \tag{3} $$

where dA is the differential area element on the contact surface $\mathcal{C}$ and $\mathbf{r} = {}^S\overrightarrow{\mathbf{or}}$ is the vector from the sensor origin. Hence, in the remainder, the origin of the contact frame $\{C\}$ coincides with the CoP, i.e. ${}^S\mathbf{p} \equiv {}^S\mathbf{p}_{\text{CoP}}$.

By definition, evaluating the wrench at the CoP eliminates the tangential moments. We further adopt the standard *soft-finger contact* model [34] which retains torsional friction about the contact normal. Consequently, the *admissible interaction wrench* expressed in $\{C\}$ is restricted to the 4 degrees of freedom relevant for stable contact regulation:

$$ {}^C\mathbf{w}_\mathbf{p} = [f_x, f_y, f_z, 0, 0, \tau_z]^\top \in \mathbb{R}^6 \tag{4} $$

We define the tactile feature state $\mathcal{S}$ on $\mathbb{SE}(3) \times \mathbb{R}^6$:

$$ \mathcal{S} = \begin{bmatrix} \mathbf{T} \\ \mathbf{w} \end{bmatrix}, \quad \mathcal{S}^d = \begin{bmatrix} \mathbf{T}^d \\ \mathbf{w}^d \end{bmatrix} \tag{5} $$

where, for compactness $\mathbf{T} = {}^S\mathbf{T}_C \in \mathbb{SE}(3)$ and $\mathbf{w} := {}^C\mathbf{w}_p$ so that $\mathbf{T}$ denotes the pose of the contact frame $\{C\}$ relative to the sensor frame $\{S\}$, $\mathbf{w}$ is the interaction wrench acting at the CoP and expressed in $\{C\}$, and $\mathcal{S}^d$ is the desired feature state. The error vector $\mathbf{s} \in \mathbb{R}^{12}$, which drives the hybrid controller, is composed of the pose error $\mathbf{s}_\mathbf{T}$ and the wrench error $\mathbf{s}_\mathbf{w}$:

$$ \mathbf{s}_\mathbf{T} = \log_{\mathbb{SE}(3)}(\mathbf{T}^{-1}\mathbf{T}^d)^\vee \tag{6} $$

$$ \mathbf{s}_\mathbf{w} = \mathbf{w}^d - \mathbf{w} \tag{7} $$

where $(\cdot)^\vee$ denotes the mapping from the Lie algebra $\mathfrak{se}(3)$ to $\mathbb{R}^6$, $\mathbf{T}^d, \mathbf{w}^d$ are the desired pose and wrench respectively. Although the contact orientation is fixed by definition of $\{C\}$, we retain an $\mathbb{SE}(3)$ pose representation to ensure geometric consistency and support future extensions to rolling or surface-aligned contacts without altering the control formulation (see discussion in Sec. V).

C. Hybrid Tactile Servoing (HTS) Framework

We now derive the hybrid force-position control law that regulates the tactile feature state $\mathcal{S}$ defined in Sec. III-B.

1) CoP-Shifted Jacobian: To control the contact frame $\{C\}$, we must address the kinematic coupling identified in Eq. (1). Standard tactile servoing typically uses a fixed sensor-frame Jacobian $\mathbf{J}_\mathbf{o} \in \mathbb{R}^{6\times n}$, which maps joint velocities to the body twist taken at the sensor origin $\mathbf{o}$. However, as shown in Sec. III-A, a rotation command generated via $\mathbf{J}_\mathbf{o}$ induces an unwanted tangential velocity at the contact point $\mathbf{p}$, leading to shear. To ensure kinematic consistency, we employ the *CoP-shifted Jacobian*, i.e., the Jacobian taken at the contact point ${}^S\mathbf{p}_{\text{CoP}}$, obtained via the adjoint transformation:

$$ \mathbf{J}_\mathbf{p} = \text{Ad}_{{}^S\mathbf{T}_C}^{-1}\mathbf{J}_\mathbf{o} = \begin{bmatrix} \mathbf{I}_3 & -[{}^S\mathbf{p}_{\text{CoP}}]_\times \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix} \mathbf{J}_\mathbf{o} \tag{8} $$

TABLE I: Comparison of tactile servoing methodologies

	IBTS [10]	PBTS [18]	HTS (Proposed)
Feature Space (s)	**2D Image Features** ($\in \mathbb{R}^k$) (Centroids, Principal Components)	**Contact Pose** ($\in \mathbb{SE}(3)$) (Relative Pose in Sensor Frame)	**CoP State** ($\in \mathbb{SE}(3) \times \mathbb{R}^6$) (CoP Pose & Contact Wrench)
Force Regulation	**Indirect** (via Feature Ref.) Tracks tactile image features	**Implicit** (via Fixed Depth) Tracks reference indentation z_{ref}	**Explicit** (via Wrench Error) Tracks normal force f_y directly
Control Law	**Interaction Matrix** Maps $\dot{\mathbf{s}}$ to sensor twist ${}^S\boldsymbol{\nu}_\mathbf{o}$	**Proportional Twist** Maps pose error to sensor twist ${}^S\boldsymbol{\nu}_\mathbf{o}$	**Decoupled Jacobian** ($\mathbf{J_p}$) Maps desired CoP twist to joint rates via $\mathbf{J}_p$
Decoupling Strategy	**None** (Coupled) Force/Motion coupled in Interaction Matrix	**Learned** (Implicit) Implicitly encoded in NN weights	**Analytical** (Explicit) Decoupled by CoP frame mechanics

where $[{}^S\mathbf{p}_{\text{CoP}}]_\times$ is the skew-symmetric matrix of the CoP vector. This transformation provides the exact mapping between joint velocities and the twist at the moving CoP. Crucially, it ensures that a pure rotation command about the CoP results in a rolling motion without linear shear, while a pure force regulation command moves the robot along the normal without inducing parasitic lever-arm moments.

2) Hybrid Control Law: To simultaneously regulate contact geometry and interaction forces, we adopt a hybrid control structure [12]. Let $\mathbf{S_T}, \mathbf{S_w} \in \mathbb{R}^{6\times6}$ be diagonal selection matrices with binary entries defining orthogonal subspaces for motion control and force regulation, respectively [35][2]. The control input is the desired CoP twist $\boldsymbol{\nu}^*$

$$\boldsymbol{\nu}^* = \mathbf{S_T}\underbrace{\mathbf{C}_{\text{pose}}(\mathbf{s_T})}_{\text{Pose Control}} + \mathbf{S_w}\underbrace{\mathbf{C}_{\text{force}}(\mathbf{s_w})}_{\text{Force Control}} \tag{9}$$

where $\mathbf{C}_{\text{pose}}$ computes the corrective velocity to minimize pose error (e.g., via proportional gain), and $\mathbf{C}_{\text{force}}$ is an admittance controller [36], [37] mapping wrench errors to compliance velocities. As discussed in Sec. V, the task-space command $\boldsymbol{\nu}^*$ is agnostic to the low-level robot control interface. It can be tracked via an impedance law on torque-controlled actuators or via differential kinematics on velocity-controlled joints.

3) Comparison to IBTS and PBTS: Table I summarizes the fundamental differences between the methods. IBTS [10], [38] attempts to regulate contact features (proxies for force and pose) via a coupled interaction matrix. PBTS [18], conversely, abstracts the tactile image into an $\mathbb{SE}(3)$ contact pose. However, it typically treats force regulation implicitly by tracking a fixed indentation depth, effectively behaving as a position controller with fixed stiffness. In contrast to IBTS and PBTS, the proposed HTS framework explicitly decouples the contact force and pose objectives. By defining the state in the CoP frame, HTS affords their simultaneous regulation via separate motion/force subspaces.

The versatility of the HTS framework lies in the design of the selection matrices $\mathbf{S_T}, \mathbf{S_w}$. In particular, the proposed formulation can be seen as a generalization of PBTS [11], [18] and force control: PBTS is recovered by setting $\mathbf{S_w} = \mathbf{0}_{6\times6}$ and $\mathbf{S_T} = \mathbf{I}_6$ (force is regulated implicitly via pose commands). By setting $\mathbf{S_w} = \mathbf{I}_6$ and $\mathbf{S_T} = \mathbf{0}_{6\times6}$, the proposed law becomes equivalent to explicit force control.

[2]Because not all wrenches are admissible at the CoP, $\mathbf{S_T}, \mathbf{S_w}$ are not complementary projectors (i.e. their sum is not $\mathbf{I}_6$).

D. Tactile processing and perception pipeline

The execution of (9) relies on the real-time estimation of the contact pose $\mathbf{T}$ and the wrench $\mathbf{w}$ at the CoP. We explain here the VBTS perception pipeline used in the experiments.

1) Contact geometry: The VBTS provides a high-resolution RGB image $I : \Omega \subset \mathbb{Z}^2 \to \mathbb{R}^3$ of the deformed elastomer surface, where Ω denotes the image domain. Following standard practice [28], [39], we estimate the surface gradients $(\partial z/\partial u, \partial z/\partial v)$ to reconstruct the height map $H : \Omega \to \mathbb{R}$, which represents the gel indentation. Surface gradients are estimated by a neural network trained on known indenters, and the height map H is obtained by integrating the normal map. The local surface normal vector $\mathbf{n}$ at image coordinate $\mathbf{u} = (u, v)$ is then computed as

$$\mathbf{n}(\mathbf{u}) = \frac{1}{\sqrt{1 + \left(\frac{\partial z}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2}} \begin{bmatrix} -\frac{\partial z}{\partial u} & -\frac{\partial z}{\partial v} & 1 \end{bmatrix}^\top \tag{10}$$

This yields a dense normal map $\mathcal{N} : \Omega \to \mathbb{S}^2$, providing geometric information compatible with both analytical and data-driven pipelines. Using the pressure proxy $P(\mathbf{u}) \approx H(\mathbf{u})$, valid under small deformations [3], we first compute the CoP in pixel coordinates, $\mathbf{u}_c = (u_c, v_c)$, as the pressure-weighted centroid:

$$\mathbf{u}_c = \frac{\sum_{\mathbf{u}\in\mathcal{C}} P(\mathbf{u})\,\mathbf{u}}{\sum_{\mathbf{u}\in\mathcal{C}} P(\mathbf{u})}, \tag{11}$$

where $\mathcal{C}$ denotes the set of pixels belonging to the contact patch. We then convert this image-plane location into a metric vector by scaling the pixel coordinates with the sensor resolution δ (mm/pixel):

$$\mathbf{p}_c^{\text{img}} = \delta \begin{bmatrix} u_c & v_c & d_c \end{bmatrix}^\top, \tag{12}$$

where $d_c = H(\mathbf{u}_c)$ is the indentation depth sampled at the CoP. Finally, this metric vector is mapped into the sensor frame $\{S\}$ using the sensor dimensions yielding ${}^S\mathbf{p}_{\text{CoP}}$.

2) Force estimation: Importantly, our control law is agnostic to the source of the contact pose and wrench estimates: they can come from *any* sensing or estimation pipeline. We demonstrate the benefit of this modularity in Sec. IV by implementing two distinct force estimation pipelines (Fig 2):

a) Extrinsic Sensor Fusion (HTS1): We use an F/T sensor mounted at the robot's wrist. The raw wrench is gravity-compensated to obtain ${}^S\mathbf{w}_\mathbf{o}$ at the sensor origin.

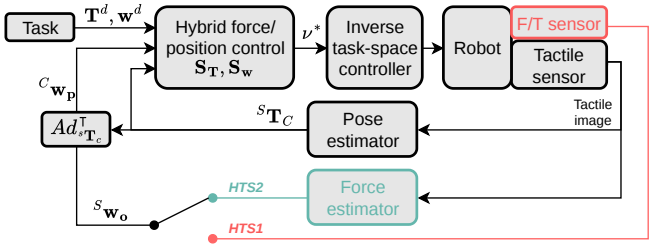

Fig. 2: Diagram of the proposed tactile control architecture: The tactile image is processed as described in Sec. III-D to extract the contact pose ${}^S\mathbf{T}_C$ and the contact wrench ${}^C\mathbf{w_p}$ at the CoP. Those estimates are then fed to the hybrid force-position controller, which computes a task space twist command ν^* sent to the robot through an inverse task space controller. The proposed variants HTS1/HTS2 only differ in the source of the contact wrench estimate.

b) Intrinsic Estimation (HTS2): . The $3D$ force is inferred from the tactile image via a data-driven approach [30]. The raw estimate is padded to form the sensor-frame wrench:

$$ {}^S\mathbf{w_o} = \begin{bmatrix} \hat{\mathbf{f}}^\top & \mathbf{0}_{1\times 3} \end{bmatrix}^\top \tag{13} $$

In both cases, ${}^S\mathbf{w_o} = [{}^S\mathbf{f}^\top, {}^S\boldsymbol{\tau}_\mathbf{o}^\top]^\top$, is transformed to the contact frame $\{C\}$ via the adjoint map:

$$ {}^C\mathbf{w_p} = \mathrm{Ad}^\top_{{}^S\mathbf{T}_C}\, {}^S\mathbf{w_o} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_3 \\ -[{}^S\mathbf{p}_{\mathrm{CoP}}]_\times & \mathbf{I}_3 \end{bmatrix} \begin{bmatrix} {}^S\mathbf{f} \\ {}^S\boldsymbol{\tau}_\mathbf{o} \end{bmatrix} \tag{14} $$

which effectively subtracts the lever-arm moment $\mathbf{p}_{\mathrm{CoP}} \times \mathbf{f}$ from the measured torque and removes the coupling (2). Note that in HTS1 the wrist F/T signal reflects the wrench transmitted through the sensor structure rather than the instantaneous contact wrench itself. Viscoelastic effects in the elastomer and intermediate structures may introduce lag between the true contact force and the measured wrist wrench. In the quasi-static interaction regime considered, this discrepancy is limited, although a residual low-pass filtering effect is still present. HTS1 should thus be interpreted as a high-bandwidth proxy for force feedback rather than a direct measurement of the instantaneous contact wrench.

IV. Experiments

This section describes our experimental validation of the proposed framework, which includes the protocol (Sec. IV-A), the setup (Sec. IV-B), the calibration (Sec. IV-C) and the results (Sec. IV-D and IV-E).

A. Experimental protocol

In order to evaluate our framework, we conduct repeatable contact-following experiments. More specifically we test the following scientific hypotheses:

- **(H1) : Explicit force regulation matters for contact stability**: We hypothesize that the proposed controller outperforms architectures that treat force implicitly as a position offset or an image feature, particularly in maintaining stable contact.
- **(H2) : Geometric consistency improves force tracking**: We hypothesize that the CoP-shifted Jacobian $\mathbf{J_p}$ improves force tracking by decoupling pose and force dynamics, whereas sensor-frame Jacobian $\mathbf{J_o}$ induces parasitic coupling between orientation changes and normal force, leading to unsafe force overshoots.
- **(H3) : Physical features allow improved performance through fusion**: We hypothesize that the performance of VBTS is currently limited by sensor bandwidth, not the control framework. We predict that fusing high-bandwidth external force data with tactile geometry improves tracking, validating the framework's capability to decouple the control law from the perception pipeline.

To isolate the impact of our contributions, we compare four distinct implementations of the controllers described in Sec. III-C and Table I: IBTS, PBTS, HTS1 ("extrinsic" variant of HTS fusing tactile measurement of CoP with high-frequency F/T sensing) and HTS2 ("intrinsic" variant of HTS using data-driven force estimation from tactile images). Crucially, to ensure a fair comparison of control strategies, we apply the CoP-shifted Jacobian $\mathbf{J_p}$ introduced in Sec. III-C.1) to *all* methods (including PBTS and IBTS), unless otherwise noted for ablation in Sec. IV-E. This implementation choice allows to isolate the effect of the tactile servoing scheme from the geometric coupling described in Sec. III-A.

For the comparative study, gains for each controller are tuned empirically to achieve the maximum stable performance (highest performance without oscillations). For the ablation study (Sec. IV-E, **H2**), gains are kept identical to isolate the effect of CoP-shifted Jacobian.

B. Experimental Setup

The experimental validation is conducted on a dual-arm setup depicted in Figure 3.

1) Robotic Platform: The main platform is a 7-DoF *Franka Emika Panda* manipulator controlled via the `franka_ros` interface, which accepts joint velocity commands at 1 kHz. To generate dynamic contact scenarios, we use a second manipulator, a 7-DoF *Flexiv Rizon*, acting as a "leader." The Flexiv executes pre-planned trajectories to disturb the contact state, while the Panda (the "follower") actively regulates the contact interaction.

2) Tactile and Force Sensing: The end-effector is equipped with a *GelSight Mini* VBTS [3], rigidly mounted to the robot flange via a custom 3D-printed bracket. A 6-axis *ATI Gamma* F/T sensor is mounted between the wrist and the tactile sensor, which is used to provide a force signal in the case of HTS1, but serves only as ground-truth validation for IBTS, PBTS and HTS2. The contact interface consists of a deformable ball placed between the tactile sensor and a rigid flat board mounted on the Flexiv's end-effector.

3) Software Architecture: The system adopts a decoupled ROS-based architecture. A dedicated *perception server* (Python) handles the tactile image stream, performing neural networks inference and features extraction on an NVIDIA GeForce RTX GPU. It publishes the high-level tactile state (CoP, force) at 25 Hz. The *control client* (C++) implements the proposed hybrid controller and runs in real-time at 1 kHz. To handle the multi-rate synchronization between the slow

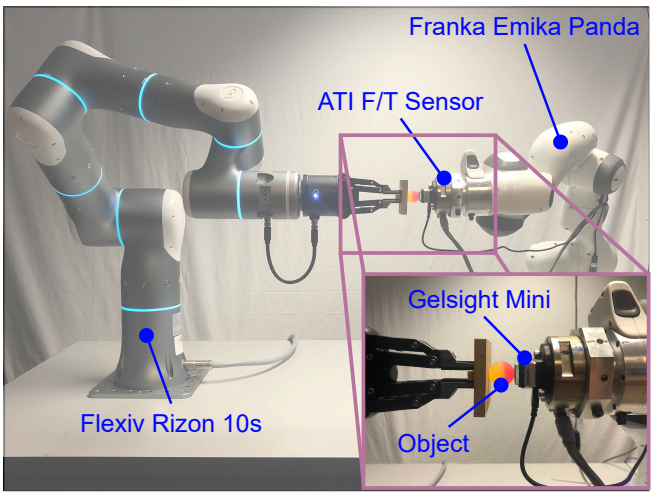

Fig. 3: Hardware setup: a deformable ball object is placed between the leader robot (Flexiv) and the follower (Panda).

vision thread and the fast control loop, the tactile feedback is up-sampled via linear extrapolation. Detailed implementation of the perception pipeline is provided in Appendix.

C. Automatic calibration procedure for IBTS and PBTS

Since baselines IBTS and PBTS do not regulate the contact force explicitly, comparing them against force-controlled methods requires a consistent initial state. To address the variability in elastomer conditioning across experiments, we implement an automatic calibration phase. The procedure performs two force-controlled probes to estimate the local normal stiffness K, then computes the indentation depth d^* corresponding to the desired target force f_z^*. A closed-loop verification step ensures the starting conditions are identical (within $\pm 0.3\,\mathrm{N}$) for every trial, enabling a fair comparison between all methods.

D. Contact following task

The task is to maintain contact with the ball by exerting a constant normal force $f_z^d = 5\,\mathrm{N}$ while keeping the CoP at the center of the sensor surface. This corresponds to the desired reference states $\mathbf{w}^d = [0, 0, f_z^d, 0, 0, 0]^\top$ and $\mathbf{T}^d = (\mathbf{I}_3, \mathbf{0}_3)$. Although the control framework defines the contact pose on $\mathbb{SE}(3)$, the experiments reported here are deliberately restricted to a translational, fixed-orientation contact-following task (see Sec. V). In this setting, active regulation of the contact-frame orientation is not required to maintain stable contact, and we therefore mask the rotational error terms via the selection matrices $\mathbf{S_w} = \mathrm{diag}(0, 0, 1, 0, 0, 0)$ and $\mathbf{S_T} = \mathrm{diag}(1, 1, 0, 0, 0, 0)$. Our experiments thus validate normal-force regulation and CoP centering only; tangential shear effects are not explicitly regulated in the present study.

The admittance gain is selected as a trade-off between responsiveness and stability: increasing it improves force-tracking speed, but requires sufficiently high-bandwidth and low-noise force feedback to avoid oscillations. In practice, gains were tuned empirically by increasing them until oscillatory behavior was observed, then slightly reducing them to remain in a stable regime. HTS1 tolerates higher admittance gains than HTS2 because the F/T force signal is available at $1\,\mathrm{kHz}$, whereas the tactile force estimate is updated at $25\,\mathrm{Hz}$.

TABLE II: Tracking performance (RMSE $\pm$ std) at low speed ($3\,\mathrm{mm\,s^{-1}}$)

Baseline	N	CoP X RMSE [mm]	CoP Y RMSE [mm]	Force Z RMSE [N]
HTS1 (ours)	5	**0.46 ± 0.00**	**0.59 ± 0.00**	**0.32 ± 0.01**
HTS2 (ours)	5	**0.46 ± 0.00**	**0.59 ± 0.00**	1.33 ± 0.01
IBTS	5	0.52 ± 0.05	0.66 ± 0.01	1.52 ± 0.24
PBTS	5	**0.46 ± 0.00**	**0.59 ± 0.00**	2.13 ± 0.16

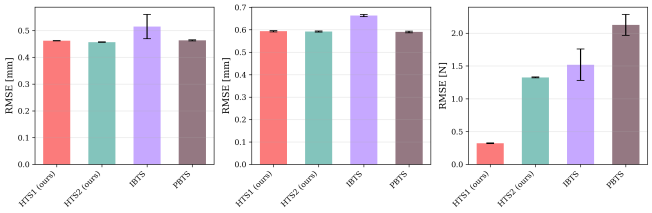

Fig. 4: RMSE of CoP and normal force tracking errors for the contact following task (slow speed, $3\,\mathrm{mm\,s^{-1}}$). HTS has similar CoP tracking performance as IBTS/PBTS but higher force tracking performance thanks to direct force feedback.

The leader robot trajectory is split into two phases: Phase 1 is a circle of radius $5\,\mathrm{cm}$ in the global (x, z)-plane, and Phase 2 is a backward–forward line motion along the global y-axis. The contact-following experiment is repeated for all methods at low speed ($3\,\mathrm{mm\,s^{-1}}$) and then at higher speed ($20\,\mathrm{mm\,s^{-1}}$). The CoP and normal force tracking RMSE reported in Table II show that the proposed approach (both HTS1 and HTS2) achieves higher force tracking performance than the baselines while maintaining similar CoP tracking accuracy, which supports **H1**. The force-tracking advantage of HTS over PBTS is expected, since PBTS regulates force only indirectly through a fixed indentation depth and is not designed for explicit force regulation. IBTS therefore provides the fairer baseline for force-tracking comparison, since the tactile features it regulates in our implementation can still covary with normal force and provide some implicit force awareness. At higher speed ($20\,\mathrm{mm\,s^{-1}}$), only HTS1 is able to maintain contact, with CoP tracking RMSEs of $1.92\,\mathrm{mm}$ and $2.32\,\mathrm{mm}$ along the x- and z-axes, respectively, and a normal-force RMSE of $2.10\,\mathrm{N}$. This is enabled by its high-frequency force feedback, which allows a higher admittance gain than HTS2, whose force estimate is updated only at the GelSight sampling frequency ($25\,\mathrm{Hz}$). This confirms the advantage of using physical features that are sensor-agnostic, since HTS1 essentially fuses geometric contact information (Jacobian at the CoP updated at $25\,\mathrm{Hz}$) with contact force information (ATI measurement updated at $1\,\mathrm{kHz}$), which validates **H3**.

E. Ablation study on $\mathbf{J}_{CoP}$ (*H2*)

In order to assess the benefit of the Jacobian transformation (**H2**) we run IBTS with and without the CoP-shift Jacobian. As shown in Table III, the baseline applied 70% excess normal force ($2.40\,\mathrm{N}$ vs $1.41\,\mathrm{N}$). This shows that controlling the velocity at the sensor origin induces unmodeled

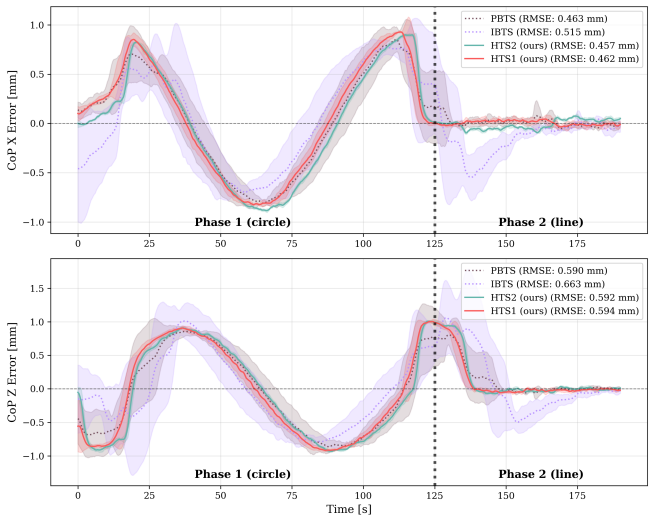

Fig. 5: Mean CoP trajectories for contact-following task at slow speed ($3\,\mathrm{mm\,s^{-1}}$).

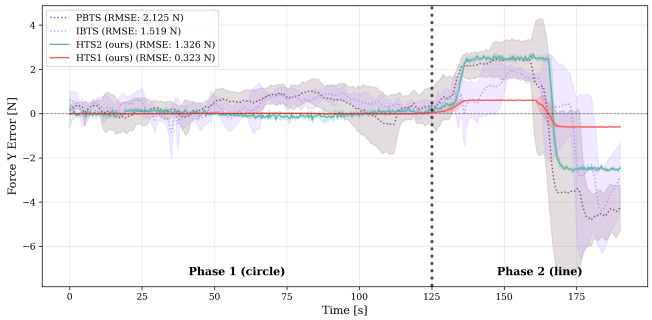

Fig. 6: Mean normal force trajectories for contact-following task at slow speed ($3\,\mathrm{mm\,s^{-1}}$).

TABLE III: $\mathbf{J}_{\mathrm{CoP}}$ ablation study (mean $\pm$ std)

Baseline	CoP X [mm]	CoP Z [mm]	Force Y [N]
IBTS-CoP	0.47	**0.65**	**1.41**
IBTS-NoCoP	0.46	1.42	2.40

compression when the contact is offset (${}^{S}\mathbf{p}_{\mathrm{CoP}} \neq 0$). This is due to parasitic moments in (1): IBTS-NoCoP commands ${}^{S}\mathbf{v_o}$ and ignores the lever arm effect ${}^{S}\omega \times {}^{S}\mathbf{p}_{\mathrm{CoP}}$, while IBTS-CoP controls directly the CoP velocity. This lever arm is explicitly measured through the "DZMP" feature in [38], but our ablation shows that this is insufficient for dynamic tracking, where the CoP is constantly offset from the sensor center. We also observed that IBTS-CoP remained stable with proportional gains $3.5\times$ higher ($K_p = 0.7$ vs 0.2) than the IBTS-NoCoP.

V. Discussion

a) Validation of the proposed formulation: The contact-following experiments validate the main design principles of the proposed formulation in the restricted translational setting studied here. First, using a CoP-centered feature state enables simultaneous regulation of normal force and CoP position, while the CoP-shifted Jacobian decouples motion and force subspaces. The ablation study in Sec. IV-E shows that controlling at the sensor origin produces large excess normal forces when the contact is offset, confirming the geometric inconsistency of sensor-origin formulations. Second, the controller operates with different perception pipelines, although performance depends on perception bandwidth, as illustrated by the gap between HTS1 and HTS2. Together, these results support the claim that the proposed CoP-centered formulation provides a physically grounded way to regulate contact interaction with modern tactile sensors.

b) Assumptions and limitations: The present formulation assumes a single connected contact patch and a soft-finger contact model. Under multiple simultaneous contact patches, the pressure-weighted centroid may lie between physically distinct interaction points and lose its mechanical meaning as a control frame. Likewise, the soft-finger assumption is appropriate under moderate shear, where tangential friction torques remain negligible. Under large shear or torsional friction, the admissible wrench space would need to be extended and the selection matrices revised accordingly. These effects are not explicitly regulated or experimentally validated in the present work and should therefore be kept in mind when interpreting the reported experiments.

c) Contact frame definition and orientation control: The modeling assumption that the contact frame $\{C\}$ is aligned with the sensor frame $\{S\}$ (i.e., ${}^{S}\mathbf{R}_{C} = \mathbf{I}_3$) is justified here by the planar geometry of the sensor and its compliant elastomer. Under quasi-static conditions with moderate indentation and bounded shear, the sensor conforms to the object geometry and the local surface normal at the CoP remains predominantly aligned with the sensor normal. Although the formulation naturally supports general contact-frame orientation updates through ${}^{S}\mathbf{R}_{C}$, this capability is not experimentally validated in the present work. The experiments are deliberately restricted to a translational, fixed-orientation setting because the current tactile estimate of the contact normal is substantially noisier and slower than the CoP position estimate, making stable rotational feedback difficult at the required bandwidth. This limitation reflects the present trade-off between spatial resolution and temporal bandwidth in vision-based tactile sensing. We therefore present rotational motion and surface-following validation as future work, likely requiring tactile sensing with a more favorable spatial/temporal compromise.

As a simple conditioning argument for this restricted translational setting, one may also consider estimating the CoP directly from the wrist F/T measurement. Under the single-contact, fixed-orientation assumption, and for a nominally normal contact force, first-order propagation of the manufacturer-quoted ATI force/torque resolutions yields a sub-millimetric steady-state CoP uncertainty at the nominal operating force of $5\,\mathrm{N}$. This estimate degrades approximately as $1/f_z$ as the normal force decreases and does not account for dynamic effects. While this does not replace an F/T-only baseline, it supports the narrower claim that wrist F/T

sensing can in principle provide a useful CoP estimate in the restricted setting studied here.

d) Torque control: The proposed formulation is agnostic to the robot control space. In our implementation, the hybrid law is realized through joint-velocity commands, leveraging the robust inner loops of an industrial manipulator and effectively treating force regulation as an outer-loop correction. The same formulation could, in principle, be combined with a friction compensator for deployment on torque-controlled robots.

VI. Conclusion

This paper showed that expressing hybrid force–position control at the Center of Pressure yields a geometrically consistent reformulation of tactile servoing. Based on this insight, we developed a CoP-centered tactile servoing approach that enables explicit regulation of contact force and contact geometry within a physically grounded feature space. In the translational, fixed-orientation contact-following task studied here, the proposed formulation improved force regulation and contact stability relative to image-based and pose-based tactile servoing baselines. Extending the experimental validation to rotational motion, shear regulation, and surface-following tasks remains an important direction for future work.

References

[1] R. S. Johansson, "19 - sensory control of dexterous manipulation in humans," in *Hand and Brain*, A. M. Wing, P. Haggard, and J. R. Flanagan, Eds. San Diego: Academic Press, 1996.

[2] S. Luo, N. F. Lepora *et al.*, "Tactile robotics: An outlook," *IEEE Transactions on Robotics*, 2025.

[3] W. Yuan, S. Dong *et al.*, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, no. 12, 2017.

[4] D. F. Gomes, Z. Lin *et al.*, "Geltip: A finger-shaped optical tactile sensor for robotic manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

[5] A. Padmanabha, F. Ebert *et al.*, "Omnitact: A multi-directional high-resolution touch sensor," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2020.

[6] N. F. Lepora, Y. Lin *et al.*, "Digitac: A digit-tactip hybrid tactile sensor for comparing low-cost high-resolution robot touch," *IEEE Robotics and Automation Letters*, 2022.

[7] H. Li, Y. Lin *et al.*, "Classification of vision-based tactile sensors: A review," *IEEE Sensors Journal*, 2025.

[8] S. Haddadin, A. De Luca *et al.*, "Robot collisions: A survey on detection, isolation, and identification," *IEEE Transactions on Robotics*, 2017.

[9] Q. Li, R. Haschke *et al.*, "A control framework for tactile servoing," in *Robotics: Science and Systems (RSS)*, 2013.

[10] Z. Kappassov, J.-A. Corrales *et al.*, "Touch driven controller and tactile features for physical interactions," *Robotics and Autonomous Systems*, 2020.

[11] N. F. Lepora and J. Lloyd, "Pose-based tactile servoing: Controlled soft touch using deep learning," *IEEE Robotics and Automation Magazine*, 2020.

[12] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, 1981.

[13] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, 1987.

[14] A. Del Prete, L. Natale *et al.*, "Contact force estimations using tactile sensors and force/torque sensors," *IEEE-RAS International Conference on Humanoid Robots*, 2012.

[15] C.-T. Wen, S. Arai *et al.*, "Tactile servoing based pressure distribution control of a manipulator using a convolutional neural network," *IEEE Access*, 2021.

[16] E. Chelly, A. Cherubini *et al.*, "Tactile-based force estimation for interaction control with robot fingers," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2025.

[17] H. Zhang and N. N. Chen, "Control of contact via tactile sensing," *IEEE Transactions on Robotics and Automation*, 2000.

[18] J. Lloyd and N. F. Lepora, "Pose-and-shear-based tactile servoing," *The International Journal of Robotics Research*, 2024.

[19] S. Tian, F. Ebert *et al.*, "Manipulation by feel: Touch-based control with deep predictive models," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2019.

[20] Z. Xu and Y. She, "Letac-mpc: Learning model predictive control for tactile-reactive grasping," *IEEE Transactions on Robotics*, 2024.

[21] S. Wang, Y. Huang *et al.*, "A robust model predictive controller for tactile servoing," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[22] H. Qi, A. Kumar *et al.*, "In-Hand Object Rotation via Rapid Motor Adaptation," in *Conference on Robot Learning (CoRL)*, 2022.

[23] Y. Wang *et al.*, "Phy-tac: Toward human-like grasping via physics-conditioned tactile goals," 2025. [Online]. Available: https://arxiv.org/abs/2511.01520

[24] Z. Xu, R. Uppuluri *et al.*, "UniT: Data efficient tactile representation with generalization to unseen objects," 2025. [Online]. Available: https://arxiv.org/abs/2408.06481

[25] I. Guzey, B. Evans *et al.*, "Dexterity from touch: Self-supervised pre-training of tactile representations with robotic play," 2023.

[26] C. Higuera, A. Sharma *et al.*, "Sparsh: Self-supervised touch representations for vision-based tactile sensing," in *8th Annual Conference on Robot Learning*, 2024.

[27] J. Huang, Y. Ye, Y. Gong, X. Zhu, Y. Gao, and K. Zhang, "Spatially anchored tactile awareness for robust dexterous manipulation," 2026. [Online]. Available: https://arxiv.org/abs/2510.14647

[28] S. Dong, W. Yuan *et al.*, "Improved gelsight tactile sensor for measuring geometry and slip," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017.

[29] M. Bauza, A. Bronars *et al.*, "Tac2pose: Tactile object pose estimation from the first touch," *The International Journal of Robotics Research*, 2023.

[30] A.-H. Shahidzadeh, G. Caddeo *et al.*, "Feelanyforce: Estimating contact force feedback from tactile sensation for vision-based tactile sensors," 2024. [Online]. Available: https://arxiv.org/abs/2410.02048

[31] A. Bicchi, J. K. Salisbury *et al.*, "Contact sensing from force measurements," *The International Journal of Robotics Research*, 1993.

[32] K. Bouyarmane, S. Caron *et al.*, *Multi-contact Motion Planning and Control*. Dordrecht: Springer Netherlands, Jan. 2018.

[33] Y. de Mont-Marin, L. Montaut *et al.*, "On the conic complementarity of planar contacts," 2025. [Online]. Available: https://arxiv.org/abs/2509.25999

[34] R. M. Murray, S. Sastry *et al.*, "A mathematical introduction to robotic manipulation," 1994. [Online]. Available: https://api.semanticscholar.org/CorpusID:108605633

[35] M. T. Mason, "Compliance and force control for computer controlled manipulators," *IEEE Transactions on Systems, Man, and Cybernetics*, 1981.

[36] D. E. Whitney, "Force feedback control of manipulator fine motions," *Journal of Dynamic Systems, Measurement, and Control*, no. 2, 1977.

[37] N. Hogan, "Impedance control: An approach to manipulation," in *American Control Conference*, 1984.

[38] Z. Kappassov, J. A. C. Ramon *et al.*, "Tactile-based task definition through edge contact formation setpoints for object exploration and manipulation," *IEEE Robotics and Automation Letters*, 2022.

[39] S. Wang, Y. She *et al.*, "Gelsight wedge: Measuring high-resolution 3d contact geometry with a compact robot finger," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021.

Appendix

The perception pipeline is implemented as a modular ROS system composed of three stages:

a) Driver and Streaming: We integrated the official `gs_sdk` library into a custom ROS publisher node. This node interfaces directly with the GelSight Mini hardware and publishes the raw RGB tactile stream to the local network.

b) Contact Geometry Inference: A subscriber node processes the incoming images to estimate the surface geometry. We utilize the pre-trained neural networks provided in the `gs_sdk` to infer the pixel-wise normal map and height map from the raw tactile images. Inference is accelerated using CUDA on an NVIDIA RTX GPU to maintain a rate of $25\,\mathrm{Hz}$.

c) Tactile Feature Extraction: A custom feature extraction node implements the logic described above to compute the CoP and the contact normal. First, the contact mask is segmented by thresholding the height map. The CoP is computed as the centroid of the contact mask weighted by the height values. In the case of HTS2, a separate node estimates the 3D contact force via the FeelAnyForce [30] network.