Extrinsic Contact Sensing with Relative-Motion Tracking from Distributed Tactile Measurements

Daolin Ma, Siyuan Dong, Alberto Rodriguez
Mechanical Engineering Department — Massachusetts Institute of Technology
<daolinma,sydong,albertor>@mit.edu

***Abstract*— This paper addresses the localization of contacts of an unknown grasped rigid object with its environment, i.e., extrinsic to the robot.**

We explore the key role that distributed tactile sensing plays in localizing contacts external to the robot, in contrast to the role that aggregated force/torque measurements play in localizing contacts on the robot. When in contact with the environment, an object will move in accordance with the kinematic and possibly frictional constraints imposed by that contact. Small motions of the object, which are observable with tactile sensors, indirectly encode those constraints, and the geometry that defines them.

We formulate the extrinsic contact sensing problem as a constraint-based estimation framework subject to the kinematic constraints imposed by the tactile measurements of object motion, and subject to the kinematic (e.g., non-penetration) and possibly frictional (e.g., sticking) constraints imposed by rigid-body mechanics.

We validate the approach with simulation and real experiments on the case studies of fixed point contact and line contact.

This paper discusses the theoretical basis for the value of distributed tactile sensing in contrast to aggregated force/torque measurements. It also provides an estimation framework for contact localization with potential impact in contact-rich manipulation scenarios such as assembling or packing.

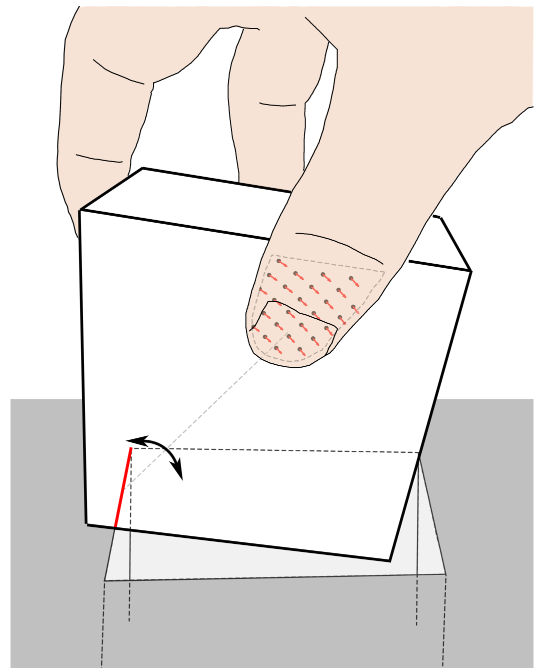


Fig. 1: Localizing extrinsic contact for pegging a box.

I. Introduction

In this work we address the problem of localizing contacts between a grasped object of unknown geometry and its environment, which we refer to as *extrinsic contact sensing*. The ability to infer the location of external contacts, which in the absence of environment instrumentation cannot be directly sensed, is of essence in insertion tasks as in Fig. 1, and is relevant to many other contact-rich manipulation tasks that involve using grasped objects, e.g., tools.

This is in contrast to *intrinsic contact sensing*, introduced by Bicchi et al. [1], which addresses the problem of localizing contacts on the surface of a robot with known geometry. A key aspect of this work is that we focus on the role that distributed tactile measurements play in localizing external contacts, in contrast to the role that aggregated force/torque measurements play in localizing internal contacts. Figure 2 illustrates the difference between the two approaches.

Intrinsic Contact Sensing. The goal is to build a map between resultant forces captured by a F/T sensor and the location of the contact force responsible for them on the surface of the robot body attached to that sensor. If the geometry of the body is known, the key challenge is to characterize when the map between contact forces and sensed wrist forces can be inverted. Initial works were based on the realization that if the body attached to the sensor has a calibrated geometry, and is convex, and the force is at a point contact, that map is invertible. In practice, in the scenario in Fig. 2 this would require precise knowledge of the geometry of the wrist→gripper→fingers chain. This idea can be extended to more complex estimation frameworks, for example to estimate contacts on a kinematic chain of known geometry [2], or when there is compliance in between the body and the sensor [3].

Extrinsic Contact Sensing. If we add the grasped object to the manipulation chain wrist→gripper→fingers→object, the assumption of known geometry becomes more impractical. We often do not have precise knowledge of the shape of the manipulated object, and the pose in the grasp is easily subject to change. Instead, we address extrinsic contact sensing by exploiting the ability of tactile sensors to observe object motion.

This work is supported by the HKUST-MIT Alliance.

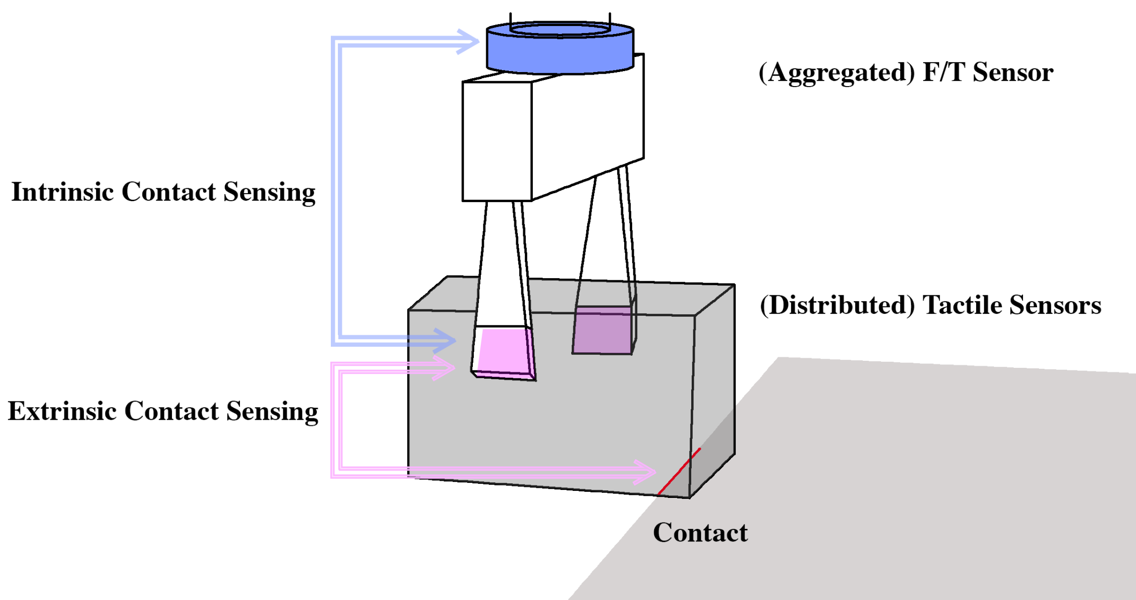


Fig. 2: **Extrinsic contact sensing** estimates the contact between the object and environment using distributed motion sensors on robot; **Intrinsic contact sensing** estimates the contact between finger and the object using aggregated Force/Torque sensor on robot.

When in contact with the environment, an object will move in accordance to the kinematic and possibly frictional constraints imposed by that contact. Small motions of the object, when observable with tactile sensors, indirectly encode those constraints, and the geometry that defines them. The approach we propose in this paper, infers contact location from observations of that motion, even without requiring knowledge of the geometry of the object. Object motion, in place, can be observed by tactile sensors that track contact deformation produced by the external contacts. We do this by exploiting tactile sensors with a new version of the high spatial resolution GelSlim [4], [5], [6] that have been shown to be able to recover the strain and slip fields at contact.

We formulate the extrinsic contact sensing problem in Section II as a constraint-based estimation framework subject to the kinematic constraints imposed by the tactile measurements of object motion, and subject to the kinematic (e.g., non-penetration) and possibly frictional (e.g., sticking) constraints imposed by rigid-body mechanics. A least-squares solver optimizes the linear or non-linear constraints to find the location of an external contact directly from tactile observations. We validate the approach with simulation and real experiments on the case study of point contact and an edge contact, detailed in Section III.

The proposed framework can be generalized to different distributed sensor setups beyond high-resolution tactile sensors, for example composed of sparser tactile sensors distributed through multiple finger contacts.

A. Related Work

Tactile sensing has been of interest to the robotics community for long. Here we review some recent work based on using high-resolution vision-based tactile sensors that can observe geometry, force, or slip.

Li *et al.* [7] presented the use of tactile sensor GelSight, to track pose of small parts held in a robot hand with high accuracy. The researchers demonstrated that the accuracy was enough to localize a USB connector and inserting it into the mating hole. Izatt *et al.* [8] extended the approach and presented a filtering algorithm fusing vision and touch to track pose of an object held by treating the output of GelSight as a pointcloud. they demonstrated that the accuracy was sufficient to grasp a screwdriver and insert it in a hole. Both works assumed known object geometry and focused on model-based pose tracking.

In recent work, instead, Dong *et al.* [9] used a self-supervised data-driven approach to estimate the contact formation and contact location of a grasped object against external features, directly from GelSlim [4] high-resolution tactile imprints. The success in localizing contact using GelSlim without direct knowledge of the geometry of the object is direct inspiration for the work in this paper. What is the mechanism by which tactile sensors can recover information about external contact location? What are the differences between using distributed tactile sensor measurements versus aggregated force/torque measurements?

II. Framework for Extrinsic Contact Sensing

The key idea is to recover the kinematic constraints that define the object's motion, which are captured by the tactile measurements.

A. Assumptions

We make the following assumptions to facilitate the formulation of the problem:

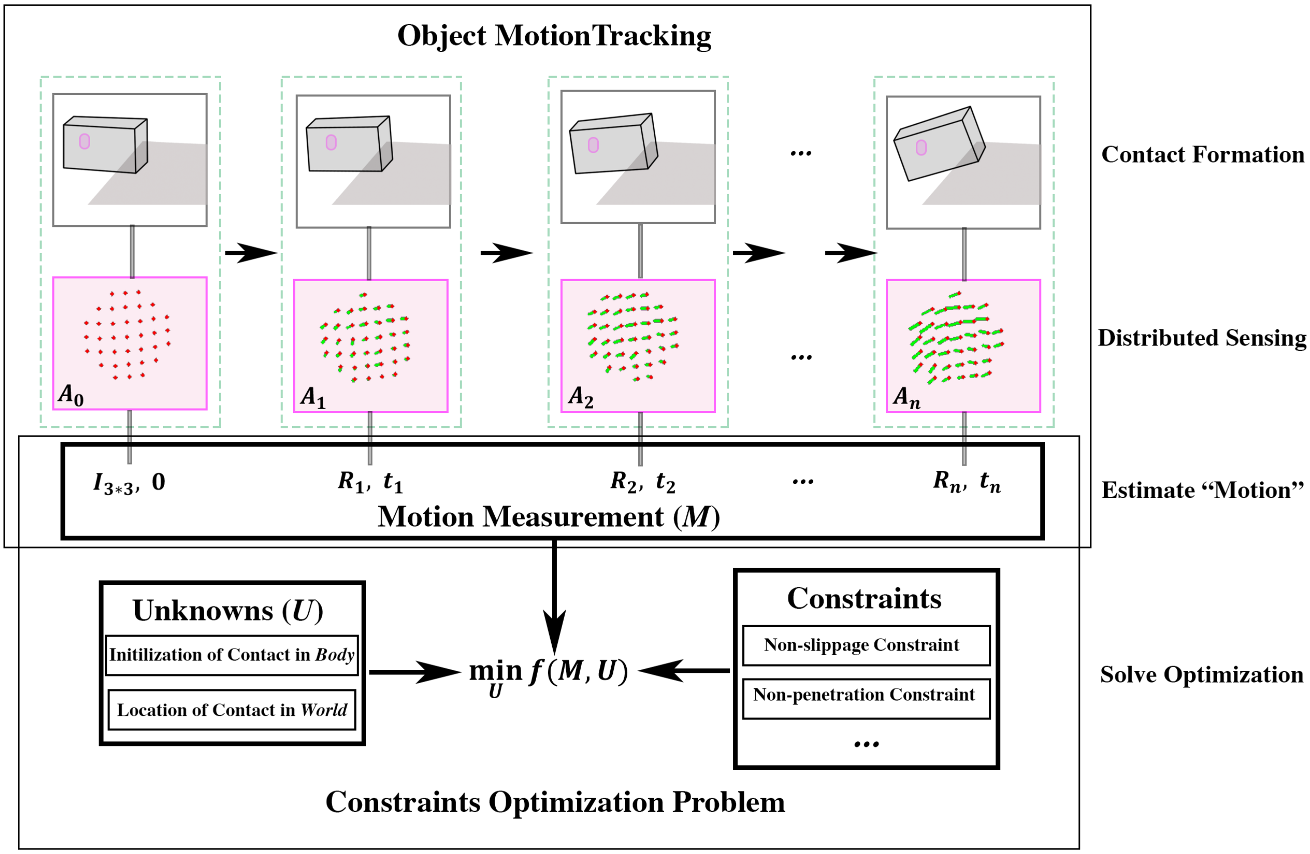


Fig. 3: Framework of *extrinsic contact localization* from distributed tactile measurements.

- The grasped object is rigid.
- Environment is rigid.
- The object remains in contact with environment.
- Grasp of the object is stable.

First, both the object and environment are rigid and they remain in contact with each other, so that stable constraint is valid and recoverable from object motion. Second, the grasp of object is stable so that no slip between object and the finger exists, which allows distributed tactile sensor to track the motion of the object.

B. Problem Formulation

We formulate the estimation problem as a constraint satisfaction optimization problem. Fig. 3 illustrates the framework of contact localization via Constraint Optimization Problem and Object Motion Tracking. The contstraints come in two forms: 1) Direct contact with the environment, and 2) observations of motion of the object, which we will describe in sections Section II-C and Section II-D.

By assumption, the object is rigid, it has 6 DOFs. Once in contact with the environment, its DOFs are reduced by the constrains formed by contact. Assuming that the motion of the grasped object is track-able, we will see that motion measurements can be used to localize extrinsic contacts by exploring what and how certain DOFs are constrained.

C. Contact Kinematics Constraints

We can add as many constraints as we have based on prior knowledge of the object and environment. These constraint equations can embed the unknown parameters for contact location in environment (or *the world frame*), such as position of a point $\boldsymbol{P_w}$, direction of a line $\boldsymbol{e_w}$, or normal direction of a plane $\boldsymbol{n_w}$. These constraints can also embed the unknown parameters for localizing the contact point on the object (or *the robot frame*), such as position of a point $\boldsymbol{P_b}$, direction of a line $\boldsymbol{e_b}$) , or normal direction of a plane $\boldsymbol{n_b}$. Denote the chosen unknowns as $\boldsymbol{U}$ based on the contact type, where

$$\boldsymbol{U} \subseteq \{\boldsymbol{P_w}, \boldsymbol{e_w}, \boldsymbol{n_w}, \boldsymbol{P_b}, \boldsymbol{e_b}, \boldsymbol{n_b}, \ldots\}.$$

For example, A fixed point contact, namely a point remains in contact with the environment all through a period of time,indicates that the position of this point $\boldsymbol{P_b}$ on object is constant in world frame even though the object is experiencing rigid transformation of rotation $\boldsymbol{R}(t)$ and translation $\boldsymbol{T}(t)$. This constraint can be expressed as

$$(\boldsymbol{R}(t)\boldsymbol{P_b} + \boldsymbol{T}(t)) - \boldsymbol{P_b} = 0 \qquad (1)$$

For another, A fixed direction contact, namely a direction remains unchanged all through a period of time, indicates that the direction $\boldsymbol{e_b}$ on the object is constant even though the object is experiencing rigid transformation of rotation $\boldsymbol{R(t)}$ and translation $\boldsymbol{T(t)}$. This constraint can be expressed as

$$(\boldsymbol{R}(t) - \boldsymbol{I}(3))\boldsymbol{e_b} = 0 \quad (2)$$

Optimization can be formulated based on satisfying these constraints. For example, by minimizing $\boldsymbol{J_1} = ||(\boldsymbol{R}(t)\boldsymbol{P_b} + \boldsymbol{T}(t)) - \boldsymbol{P_b}||$ or $\boldsymbol{J_2} = ||(\boldsymbol{R}(t) - \boldsymbol{I}(3))\boldsymbol{e_b}||$, where $\boldsymbol{I}(3)$ represents a 3*3 identity matrix.

Therefore, the tracked rigid body motion is used to regress the unknown localization of the contacts. Given the measurement of object's motion $\boldsymbol{M(t) = \{R(t), T(t)\}}$ in a time span, an optimization problem can be formulated as

$$\min_{\boldsymbol{U}} \boldsymbol{J(M,U)}, \quad (3)$$

where $\boldsymbol{J}$ corresponds to the residues of the implemented constraint equations.

The above-mentioned fix-point or fix-direction constraints don't require knowledge of the object's geometry information other than the relative rigid body motion. It means that non-slip point or line contact between unknown object and the environment can be localized using this method. On the other hand, to formulate slip-involved non-penetration constraints, information about local geometry around the contact pair is needed.

It's also worth noting that when assuming that the object is always in contact, the typical non-equality constraints for non-penetration are indeed equality constraints. Note also that the constraint equations can be expressed in either configuration space or velocity space. But measurement of object's motion in velocity space can sometimes be too noisy.

D. Relative-motion Tracking with Distributed Tactile Sensing

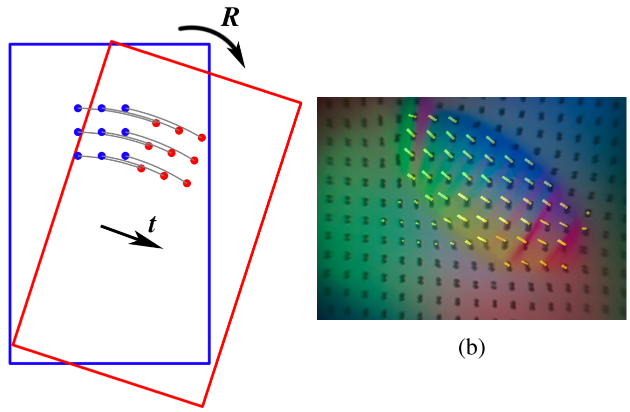
We track the relative motion $\boldsymbol{M}$ needed for the optimization problem, with the high-resolution vision-based tactile sensor GelSlim-mini, newer version of GelSlim [4], [10], [5], [6]. Key to such capability is that we can reconstruct the 3D positions of those markers in the contact patch and motions of these markers reflect motion of the grasped object.

As shown in Fig. 4, distributed tactile sensor can measure the 3D motion of markers on the sensing element that follows rigid-body motion of the grasped object. Suppose the 3D positions of m markers inside contact patch at the ZERO frame is $(x_0^i, y_0^i, z_0^i)^T$, where $i = 1, \ldots, m$. The 3D position of m markers inside contact patch from the k-th frame image is $(x_k^i, y_k^i, z_k^i)^T$. Rewrite the position list of markers in contact patch at k-th frame in a compact $m \times 3$ point list $\boldsymbol{A_k}$. Then the rigid-body transform from 0-th to k-th frame is rotation matrix $\boldsymbol{R_k}$ along with translation vector $\boldsymbol{t_k}$.

Then, the relative-motion of object from 0-th to k-th frame can be estimated by finding the optimal $\boldsymbol{R_k}$ and $\boldsymbol{t_k}$ so as to:

$$\min_{\boldsymbol{R_k},\boldsymbol{t_k}} [(\boldsymbol{R_k A_0} + \boldsymbol{t_k}) - \boldsymbol{A_k}]. \quad (4)$$

This is a typical Least-Squares Fitting problem between two 3-D point sets and can be solved efficiently using Singular Value Decomposition (SVD) [11].



(a) (b)

Fig. 4: (a) Simple 2D relative-motion tracking from motion field measurements. (b) Marker motions measured with GelSlim-mini.

III. Case Study: Point and Line Contact

We validate the proposed framework of localizing extrinsic contact with two case studies: 1) a fixed point contact, as shown in Fig. 5. 2) a line contact between the moving box and fixed an environment edge, as shown in Fig. 7.

A. Point Contact

1) Constraint Formulation: As illustrated in chapter II, given measurements of relative motion of the object $\{\boldsymbol{R_k}, \boldsymbol{t_k}\}$, we regress

$$\min_{\boldsymbol{P_b}} [||(\boldsymbol{R_k P_b} + \boldsymbol{t_k}) - \boldsymbol{P_b} = 0||] \quad (5)$$

to find location of the fixed point, P_b, in body frame.

This optimization can be seen as finding the best solution for the following equation set:

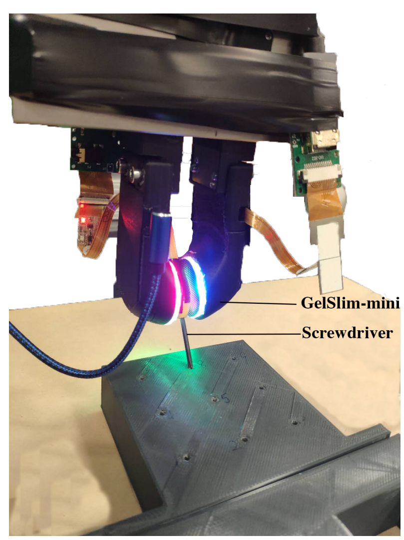


Fig. 5: Point contact: a screwdriver with tip trapped at a point.

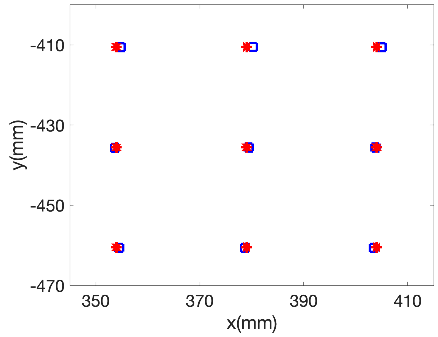


Fig. 6: Comparison of motion tracking accuracy or contact sensing accuracy. between VICON, Distributed Tactile Sensors, Single Tactile Sensor, Tactile without depth.

$$(\boldsymbol{I}(3) - \boldsymbol{R_k})\boldsymbol{P_b} = \boldsymbol{t_k} \tag{6}$$

The best estimation is $\boldsymbol{P_b} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{r})$, where $\boldsymbol{A}$ is a stacked matrix of $(\boldsymbol{I}(3) - \boldsymbol{R_k})$ and $\boldsymbol{r}$ is a stacked vector of $\boldsymbol{t_k}$ over the time span.

It can be known from eqn.6 that this equation is ill-posed and close to singularity when relative rotation matrix $\boldsymbol{R}(t)$ is close to $\boldsymbol{I}(3)$ (no rotation). It indicates that good accuracy of contact estimation relies on larger measured rotation.

2) Experiment validation: We carried out experiment to validate localization of a fixed point. The experiment setup is shown in Fig.5. A robot gripper, equipped with GelSlim-minis, is grasping a screwdriver whose tip is trapped in one of the holes on a flat plate. The robot grasp the screwdriver and move exert small motion to explore the motion space of the screwdriver and the tactile sensors measure relative motion of the screwdriver. We retrieve the point location using the proposed extrinsic contact sensing framework and the comparison of estimated results with groudtruth is shown in Fig.6

B. Line Contact

This case studies the localization of an edge in environment that is in contact with an moving box. If no slip is allowed, then this localization corresponds to localizing by solving both a fixed point constraint (eqn.1) and a fixed direction constraint (eqn.2), which is a linear regression that has a single optimal solution. We skip it in this paper and introduce the formulation of a more complex situation, where slip between the box and edge is considered.

1) Constraint Formulation: Suppose initially the object's bottom surface is a plane with equation

$$\boldsymbol{n_0} \cdot (\boldsymbol{x} - \boldsymbol{x_0}) = 0 \textbf{ or } \boldsymbol{n_0} \cdot \boldsymbol{x} = c_0, \tag{7}$$

where $\boldsymbol{n_0}$ is the normal direction of the bottom plane and $\boldsymbol{x_0}$ is a point in the plane. Then, with the tracked relative motion $\boldsymbol{R_k}$ and $\boldsymbol{t_k}$, we know that the normal direction of the bottom surface can be written as $\boldsymbol{n_k} = \boldsymbol{R_k}\boldsymbol{n}_0$ and the point on bottom surface is now at $\boldsymbol{x_k} = \boldsymbol{R_k}\boldsymbol{x_0} + \boldsymbol{t_k}$. The equation of bottom surface at the k-th frame is thus expressed as:

$$\boldsymbol{n_k} \cdot (\boldsymbol{x} - \boldsymbol{x_k}) = 0 \textbf{ or } \boldsymbol{n_k} \cdot \boldsymbol{x} = c_k, \tag{8}$$

where $c_k = \boldsymbol{n_k} \cdot \boldsymbol{x_k}$.

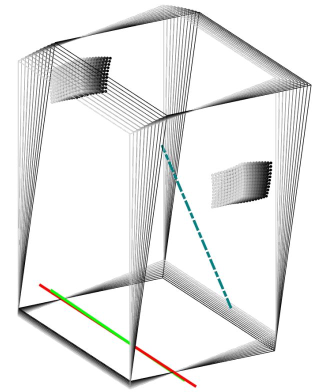


Fig. 7: Simulation of extrinsic contact sensing using tactile measurement. Red line represents the ground truth, green line is the estimation, while the cyan dashed line is the initialization input for optimization.

Because of the non-penetration constraints between the object (box) and the edge, the contact edge should be the intersection of the bottom surfaces from all time steps. Thus, the line should be perpendicular to the normal directions of the bottom surfaces in all time steps. Therefore, the unit direction vector, $\boldsymbol{l}$, of the edge can be regressed by solving the optimization problem in (9):

$$\min_{\boldsymbol{l},\boldsymbol{n_0}} \Sigma_{k=1}^{n} (\boldsymbol{n_k} \cdot \boldsymbol{l})^2 \tag{9}$$

while $\boldsymbol{l} \cdot \boldsymbol{l} = 1$. Note that the above is a nonlinear optimization problem now and the solution might be sensitive to initial guess.

To determine the position of the edge, a point on the edge should also be decided. Let's choose the point to be $\boldsymbol{x_0}$. Since the point $\boldsymbol{x_0}$ should be on the intersection of all bottom surfaces, it can be found be solving the optimization problem in (10):

$$\min_{\boldsymbol{x_0}} \Sigma_{k=1}^{n} (\boldsymbol{n_k} \cdot (\boldsymbol{R_k} \cdot \boldsymbol{x_0} + \boldsymbol{t_k} - \boldsymbol{x_0}))^2 \tag{10}$$

2) Simulation Validation: We first validate this case with synthetic simulation data.

A box is set to be in contact with an environment edge, illustrated as a red line in Fig.7. The box rotates and slides along the edge during the interaction. Virtual tactile sensor is placed on one side of the box and measures 3D motion of 11×11 points on the object with added random noise . We first solve the optimization problem in (4) to estimate the relative pose changes and then we solve the optimization problems in (9) and (10) to estimate the direction and location of the edge.

The simulation results of relative-motion tracking show that the estimated rotation matrices and translation vectors match ground truth well. The green line in Fig. 7 represents

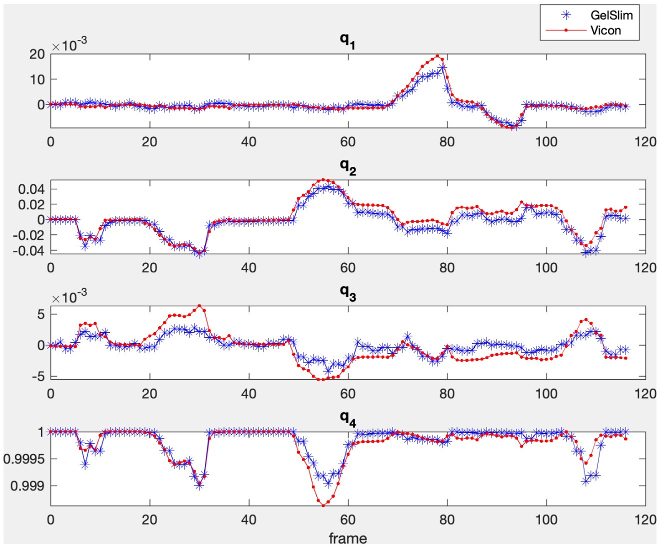


Fig. 8: Comparison of tracked relative rotation, represented as quaternion , between GelSlim-mini (blue) and Vicon (red, as groundtruth).

the estimated edge with estimated pose-change sequence shown with grey boxes. The red line represents the ground truth position of edge, which shows very good alignment with the ground truth.

3) Real Experimental Validation: We carried out experiments for the fixed-line contact case, in which tactile sensors GelSlim (and Vicon for ground truth comparison) are used to track the relative-motions of the grasped object. We achieve qualitatively good estimation results for those experiments where the line contact is maintained through the passive exploration. However, we observe that in most cases, in the absence of any controller that enforces the line contact, this is quite fragile, and easily degenerates into a point contact. As a result, many experiments for line contact estimation yield data that violates the assumptions baked into the estimation framework, and lead to poor results.

Still, we compared the tactile-measured relative motion with Vicon-tracked relative motion (serving as ground truth). Fig.8 shows a comparison of retrieved quaternion from both channels and the good matching indicates tactile tactile sensor can measure micro-scale relative-motion at a precision close to that of Vicon. Extrinsic contact estimation experiments require some low level active controller that enforces the line contact formation.

IV. Discussion

In this paper we introduced the notion of extrinsic contact sensing, by which we exploit distributed tactile measurements to estimation contacts on grasped objects without knowledge of their geometry.

A. Distributed Tactile Sensing v.s. Aggregated F-T Sensing

The manipulation of rigid objects is depends on the relationship between object motions and object forces. Force torque sensors measure force/torque resultants, while distributed tactile sensors can provide 3D force distributions which provide richer detail of the interaction. This paper shows that tactile sensors can also provide an important feature, local motion field at contact, critical for close-loop manipulation.

The strain field in tactile sensing contains all information about the interaction, and force field can be reconstructed from it [5]. At the same time, the measurement of the motion field from distributed tactile sensors presents a coordinated point cloud that that encodes information about the kinematic and friction constraints imposed by the external contacts. Such a relative finite motion, which cannot be perceived from FT sensor resultants, is key to extrinsic contact sensing.

B. Extrinsic Contact Sensing

An important feature of the proposed framework is that it can work on objects of unknown geometry. The motion tracking doesn't need any information about geometry of the object while the constraint formulation may only need knowledge of local geometry around the contact. Therefore, the proposed approach enables implementation on various unknown objects with unknown grasp pose. It's also worth noting that, the object relative-motion is described in the sensors' base frame that is not necessarily calibrated according to the world frame. Therefore, localization of the contact is also presented in this sensor frame, making it easier for deployment and closing the control loop.

Even though we assume the environment to be fixed in space. However, this ”environment” could also be another object that is moving and tracked by another robot, which relates to the scenario of assembling with a two-handed robot or multi-robot collaboration.

C. Contact Type Characterisation and Active Exploration

The key limitation to the presented work lies in having to assume a particular contact type to formulate the constraints. A clever system would be able to distinguish contact types based on continuous estimation.

Since the motion space from a more constrained contact type can be a subspace of that from a less constrained contact type, there can be ambiguity if exploration is not comprehensive even when estimation errors through different contact type assumption are compared. In order to remove the ambiguity, it's vital to guarantee that all constraints to object's motion are fully explored. Motions induced by passive exploration usually result in little rotation which also results in poor estimation accuracy. Therefore, a more general approach to extrinsic contact sensing would benefit from active exploration policies that explore all the allowed motion space dimensions with larger rotation near the current configuration by generating motions corresponding to the estimated both contact type and location, and keeping track of when contact formations are maintained or broken.

References

[1] A. Bicchi, J. K. Salisbury, and D. L. Brock, “Contact sensing from force measurements,” *The International Journal of Robotics Research*, vol. 12, no. 3, pp. 249–262, 1993.

[2] L. Manuelli and R. Tedrake, "Localizing external contact using proprioceptive sensors: The contact particle filter," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2016-Novem, pp. 5062–5069, 2016.

[3] K. T. Yu and A. Rodriguez, "Realtime State Estimation with Tactile and Visual Sensing for Inserting a Suction-held Object," *IEEE International Conference on Intelligent Robots and Systems*, pp. 1628–1635, 2018.

[4] E. Donlon, S. Dong, M. Liu, J. Li, E. Adelson, and A. Rodriguez, "Gelslim: A high-resolution, compact, robust, and calibrated tactile-sensing finger," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Oct 2018, pp. 1927–1934.

[5] D. Ma, E. Donlon, S. Dong, and A. Rodriguez, "Dense tactile force estimation using gelslim and inverse fem," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 5418–5424.

[6] S. Dong, D. Ma, E. Donlon, and A. Rodriguez, "Maintaining grasps within slipping bounds by monitoring incipient slip," in *2019 International Conference on Robotics and Automation (ICRA)*, May 2019, pp. 3818–3824.

[7] R. Li, R. Platt, W. Yuan, A. Ten Pas, N. Roscup, M. A. Srinivasan, and E. Adelson, "Localization and manipulation of small parts using GelSight tactile sensing," *IEEE International Conference on Intelligent Robots and Systems*, vol. 1, no. Iros, pp. 3988–3993, 2014.

[8] G. Izatt, G. Mirano, E. Adelson, and R. Tedrake, "Tracking objects with point clouds from vision and touch," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 4000–4007, 2017.

[9] S. Dong and A. Rodriguez, "Tactile-based insertion for dense box-packing," 2019.

[10] M. Bauza, O. Canal, and A. Rodriguez, "Tactile mapping and localization from high-resolution tactile imprints," *arXiv preprint arXiv:1904.10944*, 2019.

[11] K. S Arun, T. S Huang, and S. D Blostein, "Least-squares fitting of two 3-D point sets (computer vision)." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. vol.PAMI-9, no. no.5, pp. 698–700, 1987.