# Seeing Through Your Skin: A Novel Visuo-Tactile Sensor for Robotic Manipulation

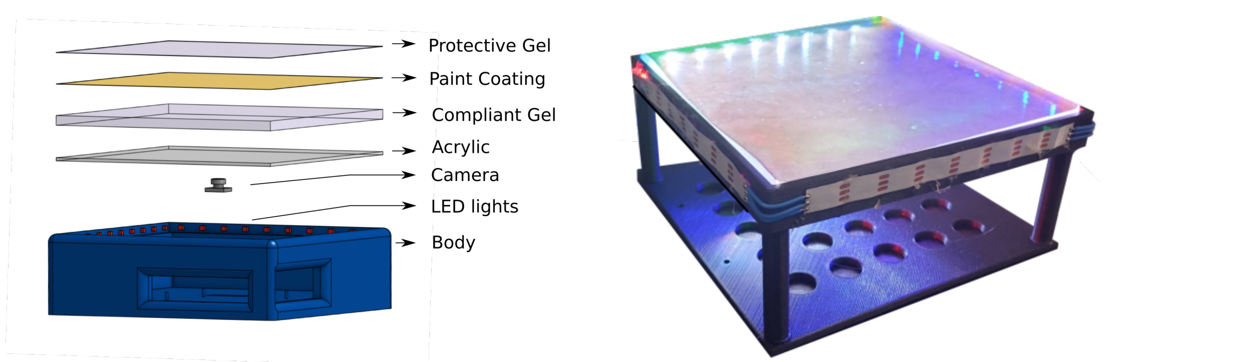
F. R. Hogan[1], S. Rezaei-Shoshtari[1,4], M. Jenkin[1,2], Y. Girdhar[1,3], D. Meger[1,4], G. Dudek[1,4]
[1]Samsung AI Center Montreal, [2]York University, [3]Woods Hole Oceanographic Institution, [4]McGill University

***Abstract*—This work describes the development of the novel tactile sensor, named Semitransparent Tactile Sensor (STS), designed to enable reactive and robust manipulation skills. The design, inspired from recent developments in optical tactile sensing technology, addresses a key missing features of these sensors: the ability to capture an "in the hand" perspective prior to and during the contact interaction. Whereas optical tactile sensors are typically opaque and obscure the view of the object at the critical moment prior to manipulator-object contact, we present a sensor that has the dual capabilities of acting as a tactile sensor and as a visual camera. This paper details the design and fabrication of the sensor, showcases its dual sensing capabilities, and introduces a simulated environment of the sensor within the PyBullet simulator.**

## I. Introduction

Tactile sensors have found wide application in teleoperation and autonomous robotic manipulation tasks where the tactile sensor informs and monitors the interaction between the robot's end-effector and the manipulated object. Optical tactile sensors use a combination of a light source and a detector, typically a camera, to capture the contact interface. Applied forces on the surface deform the surface in some way, which is captured by the sensor, often with the aid of specially positioned light sources to highlight this distortion. Existing optical tactile sensors, such as GelSight [5] are typically opaque and prevent the camera from visualizing the world external to the sensor. This is unfortunate, as the "in the hand" perspective holds valuable information about targeted contacts between the robot gripper and the environment at the critical moment prior to manipulator-object contact. Furthermore, during the manipulation process, the robot gripper will often occlude external visual cameras leaving the robot operating effectively blindly at the key moment preceding contact. This paper is motivated by the insight that optical tactile sensors afford a unique opportunity to exploit the presence of a camera in the robotic fingers to provide an additional viewpoint. We describe a novel tactile sensor, the *Semitransparent Tactile Surface (STS)* sensor which is sketched in Fig. 1. This sensor exploits a semitransparent interaction surface that provides both tactile as well as visual information

This sensor is inspired from FingerVision [4], which integrates an internal camera facing a stretched transparent elastomer embedded with printed dots. While FingerVision allows for multi-modal visual tracking feedback and tactile feedback, its tactile resolution is limited by the number of printed dots (approximately 30) on the elastomer. In contrast, the STS design introduced in this paper offers a high tactile resolution equivalent to that of the camera allowing for the recovery of richer feedback about thecontact geometry and texture.




(a) STS components

(b) STS prototype

Fig. 1. Semitransparent Tactile Sensor (STS). The sensor is designed to provide simultaneous and high resolution visual and tactile feedback by using a half-silvered paint coating over a compliant gel. When the interior of the STS is illuminated with an LED light source, the surface acts like a mirror, effectively operating as a tactile sensor. Conversely, when the interior of the STS is maintained dark, the surface acts as a translucent window, effectively operating as a camera. An early prototype is shown here which uses a Lego-based substrate, and which features a diverse set of alternative LED light sources, not all of which are needed in practice. (Lego was used as a construction material as a result of limited lab access during the pandemic.) A 3d printed version of the sensor is in development.

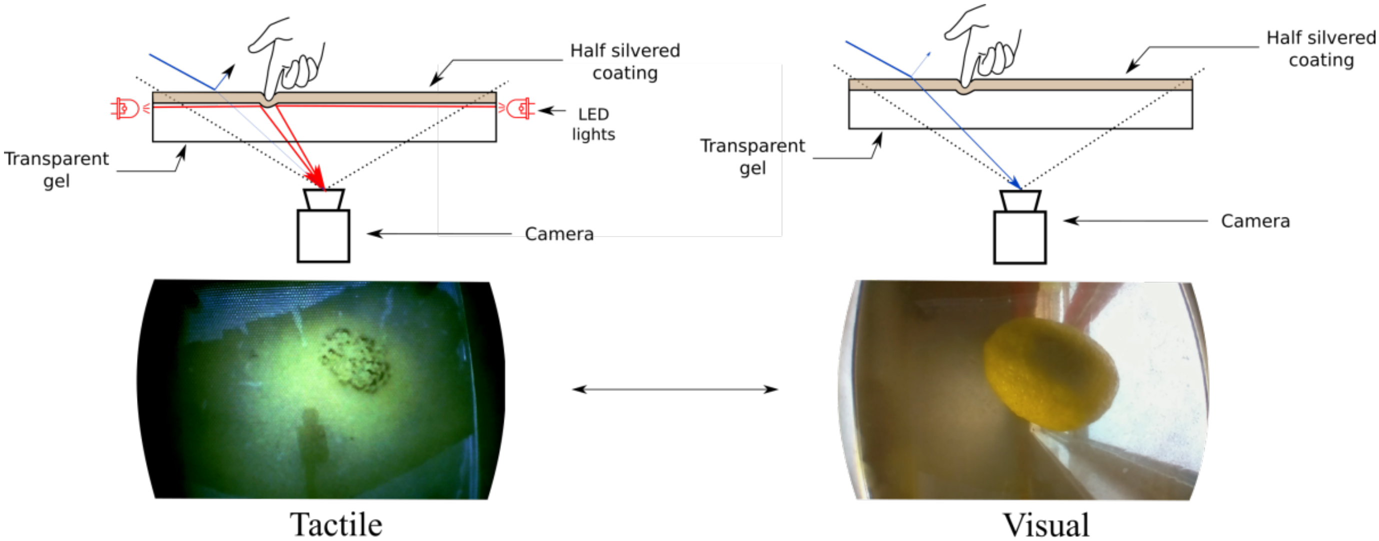



Fig. 2. The two modes of the the STS sensor. The coating of the gel is either primarily opaque, as on the left where it works much like existing sensors. Through proper illumination the surface of the gel can be made to be transparent as shown on the right, allowing the camera to view the outside world. The upper part of the figure provides a schematic of these operational modes. The lower part of the figure shows the view through the STS for the tactile and visual modes of the sensor the sensor. When the interior lighting levels are high relative to the outside world (left hand side) the surface becomes opaque and acts as a tactile sensor. When the interior lighting level is low relative to the outside of the sensor (right hand side) the camera can view through the gel and recover the outside world.

## II. SemiTransparent Tactile Sensor

The SemiTransparent Tactile Sensor (STS) uses a combination of directional light sources (LEDs) and a half-silvored paint coating to allow dual mode behavior that is dependent on the lighting conditions. This setup, depicted in Fig. 2 behaves similarly as a "one-way mirror", also known as a "transparent mirror". When one side of the membrane is held brighter than than the other, the membrane acts as an opaque mirror in the well lighted compartment and as a transparent window in the dark section. This phenomena is the same as the one used in police interrogation rooms, where the interrogation room is illuminated while the detective room is maintained dark.

As the illumination within the sensor increases, the half mirrored membrane behaves as an opaque mirror that renders a high resolution image of the contact geometry, effectively acting as a tactile sensor. As the light is decreased within the sensor, the light rays from outside the sensor can penetrate through the sensor and render a view of the external world, effectively acting as a visual sensor.

## III. Tactile Simulator

Quantifying the tactile sensation from the distortion of the gel surface has been traditionally performed using photometric stereo. Although this approach can be successful in obtaining quantitative measures of forces over the surface of the membrane, the approach is not without its limitations. It requires a fair amount of computation at each pixel and solving for the surface normal at each pixel location on the surface of the gel can be computationally expensive. It is also possible that the deformation data might be utilized in a more direct manner. For example, it might be possible to exploit a mapping directly from the distortion of the gel surface to a prior for the object being interacted with or to control in an intelligent manner a set of illuminants that might be used a a given time.

More generally, given that the STS is an optical sensor that outputs high resolution images of the contact surface and the world, this type of measurement is ideally suited for the application of convolutional neural network (CNN) based computer vision techniques. These algorithms require large amounts of data to make accurate predictions. Obtaining large datasets of interactions with the tactile surface both pre- and during-contact, although possible, is very time consuming. In order to assist in the generation of data, as well as acting as a testbed for the application of novel AI algorithms, we have developed a simulator that leverage ray casting and collision detection capability of existing physical simulators such as PyBullet and scene graph rendering capability of OpenGL for realistic physical modeling and simulation.

Accurate modeling of a tactile surface requires high-fidelity simulators capable of modeling deformable objects. However, it is computationally expensive to model deformable objects, and hence current robotic simulators have either limited or no support for soft body dynamics. Thus, data-driven [1, 3] and render-based [2] approaches have been proposed recently as alternatives. We take the latter approach and simulate the tactile sensor by applying a reflectance model on the its surface. This is essentially the inverse of the surface reconstruction problem, denoted by:

$$\mathbf{I}(x, y) = \mathbf{R}(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}), \quad (1)$$

where $\mathbf{I}(x, y)$ is the image intensity, $z = f(x, y)$ is the height map of the sensor surface, and $\mathbf{R}$ is the reflectance function modeling the environment lighting and surface reflectance [5].

The surface function $f$ can be readily obtained from the depth buffer of an OpenGL camera in PyBullet; we further clip it to the thickness of the elastomer (5mm, in our case). To compute the surface normal at each point, we find its adjacent points and calculate their principal axis using covariance analysis. Following [2], we then implement the refelectance function $\mathbf{R}$ using the Phong reflection model that breaks down the lightning into three main components of ambient, diffuse, and specular for each RGB channel:

$$\mathbf{I}(x,y) = k_a i_a + \sum_{m \in lights} \left( k_d(\hat{L}_m \cdot \hat{N}) i_{m,d} + k_s(\hat{R}_m \cdot \hat{V})^{\alpha} i_{m,s} \right), \tag{2}$$

where $\hat{L}_m$ is the direction vector from the surface point to the light source $m$, $\hat{N}$ is the surface normal, $\hat{R}_m$ is the reflection vector computed by $\hat{R}_m = 2(\hat{L}_m \cdot \hat{N})\hat{N} - \hat{L}_m$, and $\hat{V}$ is the direction vector pointing towards the camera. Through extensive search and taking into account the suggested parameters in [2], we set the specular reflection constant $k_s$ to 0.5, the diffuse reflection constant $k_d$ to 1.0, the ambient reflection constant $k_a$ to 0.8, the shininess constant $\alpha$ to 5, and RGB channel of specular and diffuse intensities ($i_s$ and $i_d$) of each corresponding light source to 1.0. Similar to [2], we apply a darkening mask based on the pixel penetration depth and tilt the light sources slightly towards the elastomer for creating more contrast between active and inactive regions.

In an expansion of [2], we enable the simulator to take into account the weight of objects for estimation of the surface penetration. We approximate the deformation of the tactile surface by modelling it with an array of springs, one at each pixel, and solve for static equilibrium given the known contact geometry and reactive forces from the simulator.

## IV. Summary and Future Work

This paper introduces a novel sensor that has the dual ability to act as a tactile sensor and as a camera. Using a half-silvered coating, we show that the tactile surface can be made to be reflective or translucent depending on the sensor's lighting conditions. The sensor's ability to provide rich multi-modal measurements makes it ideal for autonomous robotic manipulation, where the physics are driven primarily by the relative motions and forces at the frictional interfaces between the robot's end-effector and the object. We also present a simulated environment for the tactile sensor within the PyBullet software. This simulation package renders RGB images of the contact interface and provides an efficient environment for the training of visuo-tactile manipulation policies.

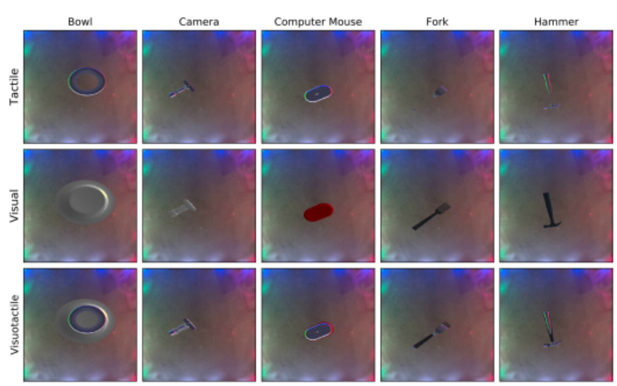



Fig. 3. Simulated tactile, visual, and visuotactile signatures of five common household objects within the simulated PyBullet environment. The simulated dataset contains five more classes (paperclip, scissors, spoon, watch, and wine bottle). The sensor computes surface deformations based on the contact forces obtained from the physics engine. Simulated images are overlaid on a background image captured from the real sensor.

## References


[1] Zihan Ding, Nathan F Lepora, and Edward Johns. Sim-to-real transfer for optical tactile sensing. *arXiv preprint arXiv:2004.00136*, 2020.

[2] Daniel Fernandes Gomes, Achu Wilson, and Shan Luo. Gelsight simulation for sim2real learning.

[3] Carmelo Sferrazza, Thomas Bi, and Raffaello D'Andrea. Learning the sense of touch in simulation: a sim-to-real strategy for vision-based tactile sensing. *arXiv preprint arXiv:2003.02640*, 2020.

[4] A. Yamaguchi and C. G. Atkeson. Combining finger vision and optical tactile sensing: Reducing and handling errors while cutting vegetables. In *2016 IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, pages 1045–1051. IEEE, 2016.

[5] Wenzhen Yuan, Siyuan Dong, and Edward H Adelson. Gelsight: High-resolution robot tactile sensors for estimating geometry and force. *Sensors*, 17(12):2762, 2017.