Research Statement

Robotic Tactile Perception for Exploration and Manipulation in the Physical World

Wenzhen Yuan

In 2017, the AI program AlphaGo beat the best human player in the complicated board game Go, which made the public re-think the old sci-fi scene: Could robots act as humans to accomplish tactful and delicate tasks, like doing the house-keeping and washing the clothes? Unfortunately, it turns out that it is more difficult for a robot to pick up a dish from a messy dish pile than play Go. The physical world has much more variables and uncertainties, and is hard for robots to deal with. A major barrier lies in the robot perception system: how to understand the physical world, and how to interact with it accordingly. Recently, the fast development of computer vision has enabled robots to well understand visual information, but the perception from other sensory modalities, especially tactile sensing, which is crucial for physical-based perception, has been largely underdeveloped.

My research goal is to build an intelligent robotic tactile perception system for the sake of exploration and manipulation. Exploration means understanding the environments and objects; manipulation means interacting with the physical world, where sensory feedback is very important. With the tactile perception, a robot can know whether a chair is rigid or is covered by comfortable and soft cushions, it can know whether an avocado is ripe enough to eat by estimating its hardness, and it can know whether the ground surface is slippery that does not suit walking. When interacting with the environment, the tactile perception will enable a robot to secure a grasp by detecting potential slip, it will prevent a robot from crushing a rotten tomato, and it will enable a robot to pick an M&M from a bag of snack mix. Without tactile perception, a robot can hardly accomplish the those tasks.

Nevertheless, the current tactile sensors can only help with limited tasks. The major challenges come from both hardware and software: How to develop sensor devices that can obtain adequate tactile information, and how to interpret the raw signal into the relevant information to understanding the world. Traditional tactile sensors measure force or the pressure distribution over a small area. While they help robots to perceive contact location and magnitude, they are very insufficient to help robots to well understand the physical world or interact with it.

For building an intelligent robotic tactile perception system that helps robot with exploration and manipulation tasks, I plan to address the challenges from three aspects: **hardware**, **algorithm**, and **integration with other components**. For the hardware, I have been working on improving the tactile sensor design to obtain more information about the contact; for the algorithm, I plan to take advantage of the fast-developing AI and machine learning technologies, and have already applied the state-of-the-art neural network methods on the tactile signals to learn the embedded information from the high-dimensional input. In addition, I believe a good perception framework is a combination of different parts. I will integrate the tactile sensing with robot motion, and other sensing modalities such as vision and audio. The integration will enable robots to learn more information. For example, by combining tactile sensing and motion, we can conduct active touch, which means the robot can conduct some actions for the special purpose of sensing, like lifting an object to estimate its weight and pressing on an object to estimate its hardness.

Current Research

My current research is on the high-resolution tactile sensing for robotic perception. I have been working on both the hardware part on sensor developing, and the algorithm part of interpreting the tactile information for different manipulation and exploration tasks.

Hardware—the GelSight sensor. In our lab, we have developed an optical-based tactile sensor, GelSight (Figure 1 (a)-(f)), which can sense the high-resolution geometry of the contact surface, as well as force and shear. The sensor uses a soft

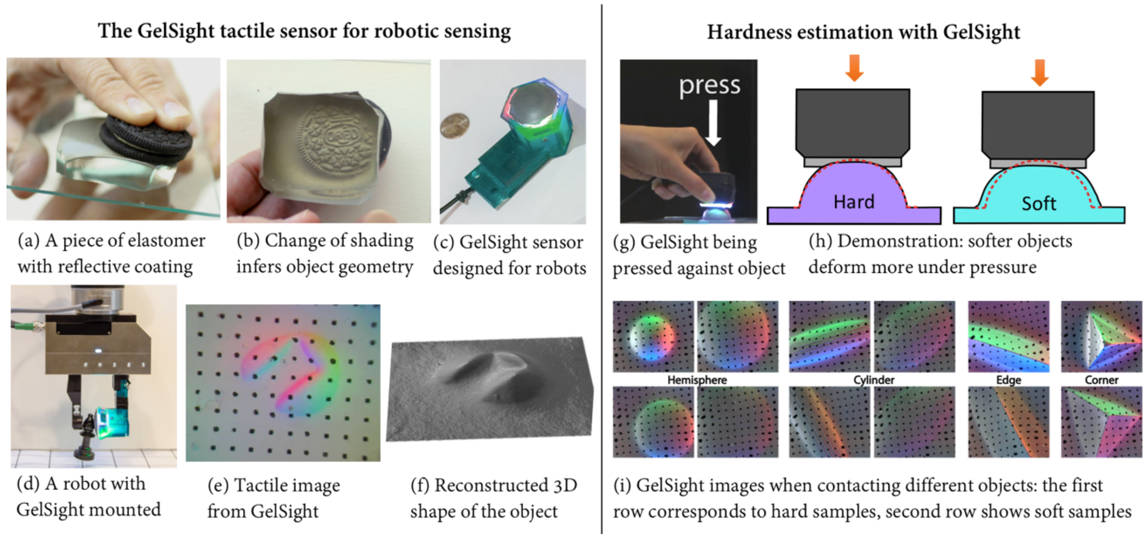

Figure 1 (a, b)[1]: principle of GelSight for shape measurement. (c-f)[2]: the fingertip GelSight. (g, h)[4]: when GelSight is manually pressed on an object, the object may deform, but softer objects deform more, making a flattened surface shape. GelSight measures the changing shape and force, and thus predicts the hardness. (i)[5]: examples of the dataset for hardness estimation

elastomer as the contact medium, and embeds a camera to capture the deformation of the elastomer. Using machine vision methods on the camera images, we are able to re-construct the high-resolution geometry of the elastomer, which is a replica of the object surface. The geometry information, especially the dynamic change of the geometry on the soft contact surface, reveals plentiful perceptual information that is unattainable by traditional tactile sensors. We also print black markers on the sensor surface to track the planar deformation of the sensor, which is directly related to the contact force and shear. The high-resolution tactile information was not available with other tactile sensors before, and it offers us opportunities to better understand the physical world with this rich information.

Algorithm—Model-based methods and deep learning. The output signal of GelSight is in the image format, which high-dimensional and highly nonlinear. To translate the raw signal into useful information, I firstly used model-based methods, such as directly calculate the object geometry, or shear on the surface. However, in some more advanced tasks, we need to extract subtle information for the highly-variant data. I applied the state-of-the-art convolutional neural networks (CNN) developed for computer vision, which proved highly effective in extracting semantic information from the high-dimensional images. The networks also successfully extract complicated information from the high-resolution tactile data.

Manipulation—Slip detection.[3][2] The slip detection is an important part of robot grasping, in that it practically tells the robots whether the grasp succeeds. People have developed different models and methods to predict slip in the past, but the robust methods that can detect slip in general cases remain to be explored. I discovered that, for objects with sharp shapes or obvious textures, slip can be inferred from the relative displacement of the detected object shape and the sensor surface; for the objects with smooth surface, slip can be inferred from the stretch degree of the sensor surface. The GelSight sensor measures the motion of both the object and the sensor surface, and the marker motion distribution reveals the stretch of the sensor surface. Experiments showed that by using the model-based method to measure the slip-related cues, a robot can predict or quickly detect slip for objects with different shapes or materials. We also conducted grasping experiments in which when slip is detected, we asked the robot to release and re-grasp the object with a greater force. Results showed that with the slip detection, the robot could successfully grasp various objects in 89% of the cases.

Exploration—Perceiving objects' physical properties. The physical properties of the objects, including hardness, roughness, and slipperiness, is an important part for human and robot to perceive the environment. The physical properties help robots to better understand the physical world, and evaluate the objects for different purposes. Those properties can only be learnt through physical contact with tactile sensing. However, those properties are hard to obtain in traditional methods, especially for the general objects. I have been trying to use GelSight to explore the properties of general objects under natural contact conditions. The projects include:

- **Hardness estimation**.[4][5] During the physical contact, an object may deform under normal force, while softer objects deform more under smaller forces, making a more flattened surface. The GelSight sensor measures both shape and normal force. By analyzing the temporal change of the shape and force indicators from the GelSight images, we can estimate the object hardness. For objects with unknown shapes, we trained a neural network (CNN+LSTM) to learn both the spatial and temporal features during the contact. Experiments showed that the method can estimate the hardness of the unseen objects under same or different contact conditions.
- **Fabric property perception by connecting vision and touch**.[6] The mechanical properties of objects largely decide our perception of them, while those properties can be hard to explicitly name or measure. In this project, we asked the question, is it possible to find a representation of objects that is related to multiple useful properties, while the names and values of the properties are unknown? We took fabric as an example, and designed an architecture to learn the property representations from both vision and touch (Figure 2 (a)(b)), based on the fact that the observations of the same piece of fabric may differ according to the sensory modality or viewing angles, but they are generated from the same property set. We used a joint CNN architecture to learn the invariant representation of the fabrics, and use the representation vectors to compare the similarity degrees between the source fabrics. The work was selected as the long oral presentation at the 2017 Conference on Computer Vision and Pattern Recognition (CVPR) (acceptance rate of 2.65%).
- **Active robotic tactile perception on clothes.** [7] The goal of this project is to build a robot system that can autonomously explore the properties of natural clothes. The system is shown in Figure 2 (c): an external Kinect sensor guides the robot to move to the proper positions for tactile exploration, and then the robot squeezes the clothes with a GelSight finger. We applied CNN to learn multiple clothing properties from the tactile data. The tactile output was used to improve the robotic exploration as well: the robot will learn good exploration locations from the Kinect images, and when the tactile data is not good, a repeat exploration procedure is required. With the trained network, the robot can explore the clothes online and learn the properties from touch.

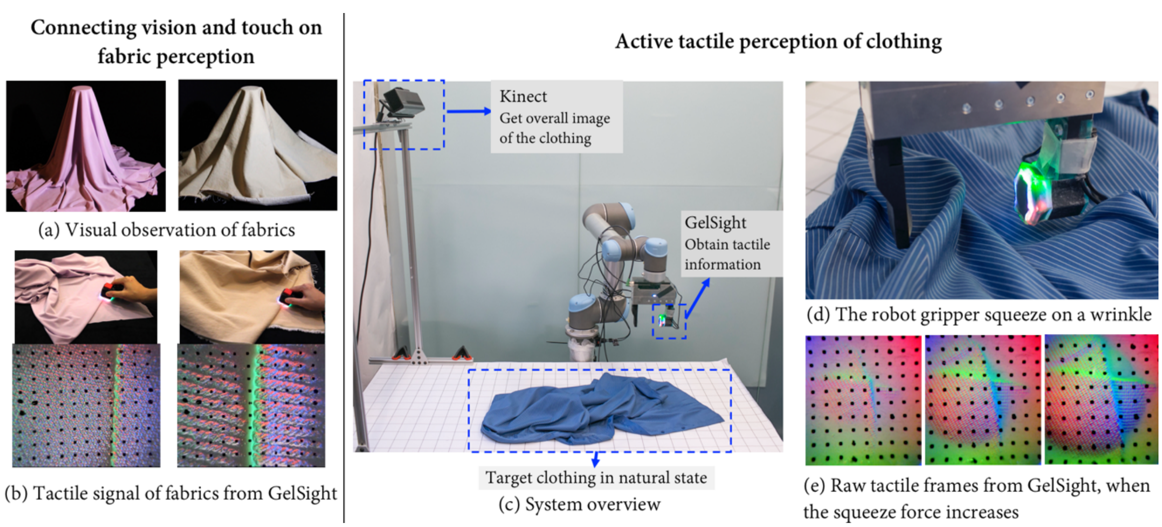

Figure 2 (a, b)[6]: examples of the visual and tactile data of fabrics. In vision, the fabrics' properties can be inferred from the shapes of draping and folding. For tactile data, we press GelSight on the flat folding of the fabrics, and the tactile images show the texture and the folding's shape, which are closely related to the physical properties. (c-d)[7]: the robot system that actively explore the properties of the clothing. The robot moves under the guidance of the Kinect sensor, and grips the wrinkles on the clothing, while GelSight records a successive sequence of tactile images. We use CNN to learn the properties from the GelSight images.

Future Research Plan

In the future, I plan to extend my research on building the intelligent robotic tactile perception system. I will research the broader use of the tactile system on exploration and manipulation tasks. Simultaneously, I am also interested in the broad applications of the tactile sensing technology. I look forward to closely collaborating with researchers working on robotic manipulation, machine learning, robotic vision, cognitive science, field robot application, and other related areas. Here are the details of the plan:

System—Hardware improvement and better machine learning algorithms. I will work on improving the hardware design and fabrication of tactile sensors to achieve more useful tactile information. I hope to redesign the high-resolution tactile sensor so that it can fit in smaller and dexterous robot hands with larger and more curved sensing areas. Besides, I also want to expand the sensor's capability in high-frequency sensing. For the algorithm, I expect to explore new machine learning models, especially neural network models, that are more suitable for tactile data, and can be trained with small dataset, instead of the vision CNNs which require large training set. At the same time, I will research the learning models or network architectures that connect the tactile data and the robot motion, and sensory data from other modalities.

Exploration—Active sensing with multi-modal input for thorough understanding of the environment. Active tactile sensing means actively performing actions to obtain tactile data. The generated actions are related to the specific perception purposes, and the combination of the motions will offer more information for sensing. I plan to develop a method to generate varied exploration actions, and use sensory feedback to guide further exploration on the objects. Other sensory modalities, like vision or audio, will be included in the framework as well. With the framework, the robot will be able to get more thorough understanding of the target objects. The perception goal could be measuring multiple physical properties of the objects, like 3D shapes or shape descriptors (such as curves, corners, edges), slipperiness, and weight; or it could be new object descriptors that are more semantic, such as whether a clothing material is warm, or whether a sofa is comfortable to sit on.

Manipulation—Closed-loop manipulation / industrial applications. I am interested in research further applying tactile sensing for extensive manipulation tasks. This includes building new manipulation models that contain tactile sensing for motion planning. For example, when grasping objects, the perceived properties of the objects, such as hardness, shape, texture and slipperiness, could suggest different grasping strategies; during the grasping, the tactile feedback could lead to adjustment of the grasping position or force, in order to secure the grasp. When trying to turn an unfamiliar knob, the tactile feedback could suggest whether the knob is turnable, or which is the proper direction to turn. At the same time, there are many industrial uses of tactile technologies. I communicated with engineers from different industries, and found that there is a great need for tactile sensing in different occasions, such as grasping and in-hand localization on the assembly lines, connector plugging for product testing, and surface friction detection for moving robots' safety in warehouses. I look forward to further collaborating with industrial partners and apply tactile systems in real-world applications.

Broader applications. In addition to the research for robotic tactile systems, I look forward to collaborating with other researchers and exploring applications of tactile technologies in other areas. Potential collaboration projects include tactile simulation for VR/AR, tactile sensing for prostheses, medical tests, surgical robots, human-machine interaction.

For grant supports, with the fast development and increasing demand for robotic technologies, many funding agencies have increasing support in the field. For example, NSF has multiple funding programs to support robotic research, such as NRI on collaborative robots, S&AS on robot platform including sensing, and EFRI on soft robots. ONR, DARPA, and NASA have been supporting robotic research for a long time. I believe they will be glad to support the fundamental and cutting-edge research on robotic tactile sensing. I will also actively connect with potential industrial sponsors. In the past years, Toyota Research Institute has been a main sponsor of our lab, and I plan to keep the connection and seek further funding supports or collaborations. I also have also talked to people from Mitsubishi, Huawei, and Midea, and they showed strong interest in using the advanced tactile sensors in their applications, and I believe they could be willing to support the research on tactile sensing. Other industries like Amazon, Google, Facebook and Foxconn also have a growing interest in robotics, either in intelligent robots or industrial robots. In both cases, tactile sensing will be important. I will connect with them and look for potential funding support.

With the fast development of robotic technology, tactile sensing will play an increasingly important role in making the robots better understand and interact with the physical world. I am excited to play a part in promoting the robotic tactile perception, and contribute to the establishment of intelligent robots. I look forward to the day that the intelligent robots could skillfully accomplish more physical related work and free humans from tedious low-level labor.

References

[1]. Micah K. Johnson, and Edward H. Adelson. "Retrographic Sensing for the Measurement of Surface Texture and Shape." In 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1070-1077. 2009.

[2]. Siyuan Dong, **Wenzhen Yuan**, and Edward Adelson. "Improved GelSight Tactile Sensor for Measuring Geometry and Slip." In Intelligent Robots and Systems (IROS), 2017 IEEE International Conference on, pp. 208-215. 2017.

[3]. **Wenzhen Yuan**, Rui Li, Mandayam A. Srinivasan, and Edward H. Adelson. "Measurement of Shear and Slip with a GelSight Tactile Sensor." In Robotics and Automation (ICRA), 2015 IEEE International Conference on, pp. 304-311. 2015.

[4]. **Wenzhen Yuan**, Mandayam A. Srinivasan, and Edward H. Adelson. "Estimating Object Hardness with a GelSight Touch Sensor." In Intelligent Robots and Systems (IROS), 2016 IEEE International Conference on, pp. 208-215. 2016.

[5]. **Wenzhen Yuan**, Chenzhuo Zhu, Andrew Owens, Mandayam A. Srinivasan, and Edward H. Adelson. "Shape-independent Hardness Estimation Using Deep Learning and a GelSight Tactile Sensor." In Robotics and Automation (ICRA), 2017 IEEE International Conference on, pp. 951-958. 2017.

[6]. **Wenzhen Yuan** *, Shaoxiong Wang*, Siyuan Dong, and Edward Adelson. "Connecting Look and Feel: Associating the visual and tactile properties of physical materials." In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5580-5588. 2017. (*equal contribution)

[7]. **Wenzhen Yuan**, Yuchen Mo, Shaoxiong Wang, and Edward H. Adelson. "Active Clothing Material Perception using Tactile Sensing and Deep Learning." Robotics and Automation (ICRA), 2018 IEEE International Conference on. 2018. (submitted)