Pre-touch Sensing for Sequential Manipulation

Boling Yang[1] Patrick Lancaster[2] Joshua R. Smith[3]

***Abstract*— The primary focus of this work is to examine how robots can achieve more robust sequential manipulation through the use of pre-touch sensors. The utility of close-range proximity sensing is evaluated through a robotic system that uses a new optical time-of-flight pre-touch sensor to complete a highly precise and sequential task - solving the Rubik's cube. The techniques used in this task are then extended to a more general framework in which ICP is used to match pre-touch data to a reference model, demonstrating that even simple pre-touch scans can be used to recover the pose of common objects that require sequential manipulation.**

I. INTRODUCTION

As robots continue to transition from operating in controlled, carefully designed environments towards human-centric, unstructured ones, they will have to make more sophisticated use of sensing to cope with the inherent uncertainty in our real world. In particular, robust robot manipulation is difficult to achieve because of the uncertainty involved in manipulating an object [1]. For tasks that require sequential manipulations, the robot's belief about the pose of the object at any point in time can be corrupted by a poor characterization of the initial pose, or through the accumulation of error caused by controller noise, previous imperfect manipulations, and perceptual errors. Depending on the precision required, such errors can cause the robot to fail at the task of interest.

In this work, we show that accumulating manipulation errors can be controlled by proximity sensors mounted to the robot's end-effectors, what we call 'pre-touch' sensors. First, we demonstrate that pre-touch sensing allows a robot to much more effectively complete a particular sequential manipulation task, i.e solving a Rubik's cube. In order to demonstrate that pre-touch sensing can be applied to other objects that may require sequential manipulation, we show that employing even a simple pre-touch scanning strategy allows the robot to infer the pose of a number of common objects.

We chose to first tackle the task of solving the Rubik's cube because the challenges that general purpose robots face when attempting to achieve sequential manipulation are well represented by this task. Solving a Rubik's cube can require up to 20 rotations [2]. For each rotation, the robot may have to alter its grasp on the cube by transferring it from one hand to the other and/or shifting the grasp point(s). Because each face of a Rubik's cube contains sub-cubes of a dimension of only 1.9 cm, each manipulation may only be able to tolerate error on the order of millimeters. Even if the robot is able to complete a rotation at a given point in time, an imprecise grasp could cause future rotations to fail by invalidating the robot's belief of where the cube is with respect to the robot's gripper.

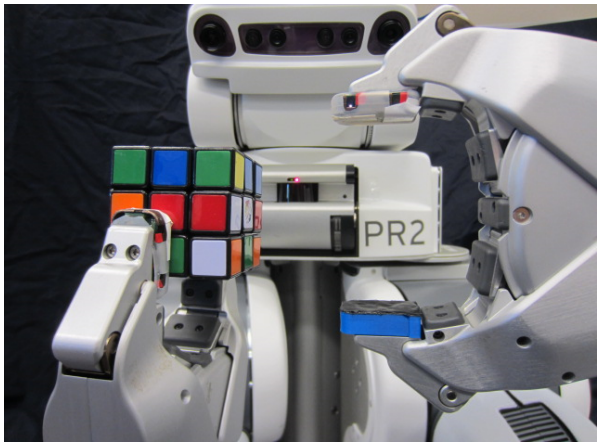

Fig. 1: The robot is able to precisely manipulate the Rubik's cube using the equipped pre-touch sensors.

Ultimately, the authors aim to show that pre-touch sensing enables robots to accomplish sequential manipulations more effectively. We hypothesize that pre-touch scanning enables a robot to gain the geometric information necessary to estimate the pose of an object of interest, and thereby perform actions that are common in sequential manipulation, such as re-grasping. The specific contributions of this work are:

1) Introduction of a new type of pre-touch sensing based on optical time-of-flight measurements, and full integration of this sensor with the PR2 system.
2) The application and evaluation of pre-touch sensing for robot manipulation in the context of solving the Rubik's cube. To the best of the authors' knowledge, pre-touch has never previously been applied to solving the Rubik's cube. Furthermore, the authors argue that solving the Rubik's cube is the most complex, sequential task to which pre-touch sensing has ever been applied.
3) The use of an Iterative Closest Point (ICP) algorithm to align a 1D pre-touch scan to a pre-existing reference point cloud in order to estimate object pose.

This work was supported in part by the National Science Foundation under grant IIS-1427419

[1] Boling Yang is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA bolingy@uw.edu

[2] Patrick Lancaster is with the Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, USA planc509@cs.washington.edu

[3] Joshua R. Smith is with the Department of Computer Science and Engineering and the Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA jrs@cs.washington.edu

4) A comparison of the performance of optical time-of-flight pre-touch (mounted in one finger) with our prior electric field pre-touch sensor (mounted in another finger), and the finding that the two sensing modalities are complementary.

The remaining sections of this paper detail the development of a new optical pre-touch sensor, and demonstrate the utility of pre-touch sensing in sequential manipulation. Section 2 discusses related work. Section 3 describes the utilized pre-touch sensors, and Section 4 details the scanning strategies used to estimate object pose. Section 5 outlines and presents experiments conducted to analyze the effectiveness of these methods. Section 6 summarizes this work and proposes future work.

II. RELATED WORK

Many approaches to robot manipulation have used cameras to localize target objects. For example, Maitin-Shepard [3] developed a vision-based algorithm to reliably detect corners of a piece of cloth, which they used to autonomously fold towels. In another example, Chang [4] singulates objects from a pile. The developed perception module uses image data to determine whether a grouping of one or more items has been singulated and how that group should be manipulated. Cameras have also been integrated into closed-feedback loops in order to achieve visual servoing. Vahrenkamp [5] demonstrate how a humanoid robot can utilize visual servoing to grasp multiple types of objects. For a more complete study of visual servoing, see [6].

While cameras can be employed by robots to sense objects from relatively far away, tactile sensors aid manipulation upon making contact with an object. Unlike cameras that are usually mounted to the head of the robot, tactile sensors are attached to the robot's end-effector. As a result, tactile sensors are typically more maneuverable, potentially allowing them to sense regions of an object that could not be imaged by a camera because of their relatively static nature. Li [7] created a tactile sensor that utilized GelSight technology to produce height maps with a resolution of 240x320 pixels. With these height maps, they were able to localize small parts within the robot's grasp and perform fine-grained manipulations, such as insertion of a USB connector. Petrovskaya and Khatib [8] use tactile measurements to estimate object pose. Particle-based methods are used to represent the belief distribution over possible poses, allowing their robot to locate and grasp a box and door handle.

Pre-touch sensors typically operate at a range intermediate to that of tactile sensors and vision-based sensors, endowing them with some of the same benefits that are achieved by sensors at both ends of the range spectrum. Similar to tactile sensors, pre-touch sensors are mounted to the robot's end-effector, making them more robust against occlusion than cameras. Also, by sensing at a closer range, they have the potential to measure more precisely. Similar to camera or depth sensors, pre-touch sensors do not have to make contact with a surface or object before a measurement can be taken. Contact is a disadvantage for tactile sensors because it can cause an object to be unintentionally displaced.

Optical pre-touch sensors have been used extensively in robot grasping because of their ability to accurately measure a wide range of materials. Hsaio [9] combine optical pre-touch sensors with a probabilistic sensor model to build a reactive controller that could grasp a number of common objects. In [10], Maldonado use measurements from an optical sensor based on computer mouse technology to reconstruct the shape of objects and grasp them. Data from the optical sensor is also used for classification of object surface and slip detection. Guo [11] utilizes a break-beam optical sensor to measure objects within the robot's gripper that are difficult for other pre-touch sensors, such as specular objects that present challenges for reflective optical sensors. All three of these works measure the amount of light reflected off of an object to detect it [10], [11] or infer distance [9], whereas the sensor we developed uses time-of-flight technology to measure distance. Compared to [9], our sensor has a longer range with comparable accuracy, and does not need to be calibrated for the surface reflectivity of the object.

Electric field sensing has been widely explored in the context of pre-touch sensing. Electric field sensing is typically achieved by transmitting an AC signal from one electrode to another. Objects near the sensor will alter the displacement current between the two electrodes, inducing a deviation from the baseline measurement (in which no objects are near the sensor) [12]. Electric field sensors are most adept at measuring conductive objects, but are unable to detect plastics, foams, or other objects with a dielectric constant similar to that of air. In [13], electric field sensing is used to localize an object and preshape the robot's fingers in order to achieve a stable grasp. Mayton et al. [14] extend electric field sensing to co-manipulation of objects between humans and robots. Finally, Mühlbacher-Karre [15] integrate electric field sensing into a robotic bar-tending system in order to determine the fill level of beverages.

More recently, acoustic sensing has been developed for pre-touch sensing. In [16], Jiang and Smith create a "seashell effect" sensor to localize objects that are typically difficult to sense with RGB-D cameras and other optical sensors, and then apply those measurements to grasping. The sensor consists primarily of a miniature metal pipe that has a microphone attached to one end. As objects approach the open end of the pipe, the effective resonant frequency of the pipe changes, which the microphone measures.

The above works focus on estimating the pose of the surface local to the pre-touch sensors for the purpose of single-shot grasping, whereas we are interested in full pose estimation of an object. Knowing the full pose becomes increasingly important when considering more complex manipulations, such as those that involve opening, twisting, or re-grasping an object. To the best of the authors' knowledge, our work is the first to estimate the full pose of an object by applying ICP (or any other similar algorithm) to a single, sparse pre-touch scan.

One of the disadvantages of the previously mentioned pre-

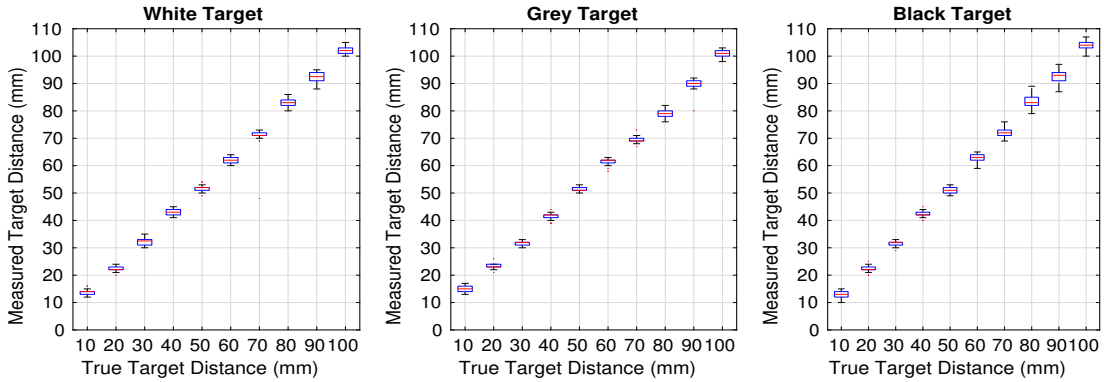

Fig. 2: Boxplots of sensor measurements over the specified range for white, grey, and black target objects. Each box consists of 30 measurements.

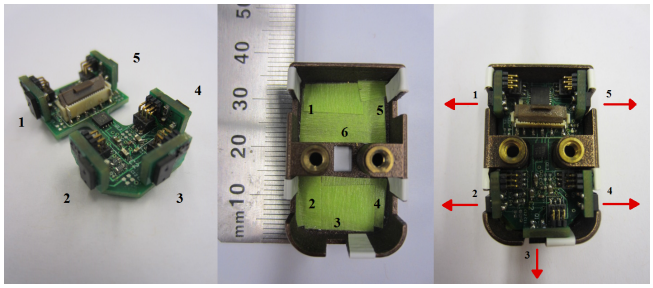

Fig. 3: Left: The printed circuit boards that compose the sensor. Middle: The 3D printed sensor casing. The hole in the middle of the case is for a sensing module soldered to the bottom of the main PCB. Right: The assembled sensor with arrows denoting the directions of five out of six sensor modules' infrared beams.

touch sensors is that they only make sparse measurements, requiring the robot to move its end-effector in order to densely measure a region of interest prior to manipulation. Thus, a number of works have developed a compromise between head-mounted cameras and pre-touch sensing by mounting cameras on the robot's wrist [17] [18]. However, although less actuation may be required to measure a region, the relative size of the cameras limit the robot's mobility. In addition, it may not be necessary to densely measure a region in order to achieve the given task, as demonstrated by the experiments in Section 5.

Solving the Rubik's cube has been previously suggested as a benchmark for robot manipulation [19], [20]. Beyond this suggestion, [20] and [21] develop robot systems for solving the Rubik's cube. In [20], the robot's grippers are somewhat specialized; their trough shape is used to grasp the cube at the corners and thereby reduce the uncertainty in the cube's state. The system in [21] was based on a PR2 robot without task specific modification of the grippers. They demonstrated that their system can successfully solve a basic Rubik's cube puzzle that required six rotations, which matches the performance of our baseline system that does not use pre-touch sensing for manipulation. However, there is no further information to indicate that this system is capable of robustly solving Rubik's cubes that require a larger number of rotations.

III. SENSOR HARDWARE

This work explores the use of two different pre-touch sensors in aiding sequential manipulation. One is an electric field sensor that was adapted from the work of [14]. The other is a new optical time-of-flight sensor, which the following paragraphs of this section will describe. Both sensors' casings were designed to match the form factor of the PR2's fingertip. On each of the robot's parallel jaw grippers, an optical sensor was attached to the left fingertip, while an electric field sensor was mounted to the right fingertip. Only the optical sensor was used when attempting to solve the Rubik's cube because the electric field sensor is not able to sense the plastic Rubik's cube effectively.

The optical pre-touch sensor measures the distance to a surface using the VL6180x optical time-of-flight sensor [22] module created by ST MicroElectronics. It is capable of measuring the distance to an object with millimeter accuracy at a range of 1cm to 10cm. Furthermore, some objects (particularly those with light colors) can be detected out to a range of 25.5cm. Fig. 2 demonstrates the sensor's performance over various colored targets.

The sensor supports up to six VL6180x sensing modules, where one can be placed at the tip of the finger, two on each side, and one in the pad of the finger. The location of each sensing module is labeled in Fig. 3, and will be referenced as such when referring to a specific sensor module throughout the rest of this work. The sensor's design consists of a main board (29x16.5mm) and a secondary board (8.25x5.75mm). The main board hosts an ATMega168PA microcontroller that communicates with each of the VL6180x sensing modules over I2C. The sensor module on the pad of the fingertip sensor is directly soldered to the main board. Every other module is connected to the main board through 1mm pitch headers that interface between the main board and secondary boards.

The robot is able to collect data from each of the six sensing modules at a rate of 30hz. Measurements are passed from the sensor's microcontroller to the robot through an SPI

interface built into the gripper and then published to a ROS topic. Two metal screw inserts are pressed into the sensor's casing so that it can be fastened to the robot's fingertip.

The components required to build one of these optical sensors cost less than $100. The circuit schematics, PCB files, sensor casing CAD files, and firmware are publicly available at https://bitbucket.org/planc509/optical-distance.

IV. METHODS

Throughout this work, the robot uses simple scanning strategies to estimate the pose of various objects. The simplicity of these scans suggests that they are applicable to a wide range of objects. The following subsections will detail the scanning strategies used, and Section 5 will present their effectiveness. In the remainder of this text, unless explicitly said otherwise, any coordinate references are with respect to the coordinate frame of whichever gripper is currently holding the object. That is, the y-axis extends along the direction that the gripper opens and closes, the x-axis extends out along the direction that the robot's fingertips point, and the z-axis is orthogonal to both the x and y-axes using a right-handed coordinate system.

A. Optical Pre-touch Scanning for Rubik's Cube

In this work, pre-touch sensing is used to estimate the pose of the cube not only to marginalize positional error in the next grasp, but also to correct any error that does occur from one manipulation to the next. As the robot solves the Rubik's cube, it will need to transfer it from one hand to the other, as well as change how it is grasping it. Before each re-grasp, the robot uses pre-touch scanning to refine its estimate of the pose of the cube with respect to the coordinate frame of the gripper that is currently holding it. The robot assumes that the cube is oriented such that its upper and lower faces are approximately parallel to the ground. This assumption is not always true, but works well in practice. Furthermore, the robot already has a good approximation of the center of the cube's position in the y direction and its rotation around the y-axis because the cube is held between the robot's fingers. However, due to errors in previous re-grasps, the cube could have shifted unexpectedly along the x and/or z directions. There are many pre-touch scanning strategies that could be used to estimate the position along these two directions, but our method used the following strategy in order to minimize the amount of actuation required:

1) The gripper that is not holding the cube is opened if it is not already open.
2) The optical pre-touch sensor on that gripper's fingertip is then oriented such that the beam of sensing module 3 is normal to one of the faces of the cube that is normal to the xz-plane, such as in Fig. 1. This gripper is positioned such that the beam is not yet broken by the cube.
3) The gripper begins to close, causing the sensor to move in the y-direction. As the gripper closes, any significant change in the sensed distance indicates the position of the edge of the cube, allowing the robot to infer the cube's position along one of the uncertain axes.
4) Once the gripper has finished closing, the robot uses the sensor's distance measurements at the current position to estimate the position of the cube along the remaining uncertain axis.

This pre-touch scanning strategy was integrated into a baseline system that uses a computer vision module to recognize the colors of the cube faces, an iterative deepening A* search [2] [21] to determine the necessary cube rotations, and a finite-state machine based motion planner to execute the trajectories necessary to solve the cube.

B. Pre-touch Scanning for Common Objects

Pre-touch scanning can also be applied to objects with more complex geometry. We aim to demonstrate that a simple 1D scan of an arbitrary object can contain enough distinctive features to estimate its pose when matched to a reference model. This estimate could be useful when initially picking up an object, or before performing a re-grasp.

Again, there are many different possible scanning strategies. One could try random trajectories until the robot is confident that it has correctly inferred the pose, or trajectories based on heuristics or learning models could be executed. We will explore the generation of such trajectories in future work. For this work, a single trajectory was chosen by the experimenter for each object that was likely to capture distinctive features. Each of the chosen trajectories consisted of the scanning gripper moving in a straight line with a fixed orientation. While executing each trajectory, the robot sampled the object at discrete points along it, where the interval between the sampling points was determined by the size of the object such that approximately 50 samples were collected. The process of obtaining a sample was slightly different depending on which pre-touch sensor was being utilized.

When using the electric field pre-touch sensor to scan the object, the robot oriented the front of its fingertip orthogonal to the trajectory and towards the object. The object affects the sensor's measurements by shunting displacement current away from the electrode located at the front of the fingertip sensor. At each sample point, the robot moves its gripper towards or away from the object, causing the amount of current shunted away, and therefore the change from the baseline measurement (i.e. the measurement when the object is far way from the sensor), to change. The robot obtains a distance measurement by servoing its gripper in this fashion until the change from the baseline measurement is sufficiently close to a pre-determined threshold. This threshold, which was recorded before beginning the scan and is different for each object, indicates when the fingertip is 1.5cm away from the object. This technique makes the assumption that the volume of the object local to the fingertip is uniform throughout the trajectory. Although this assumption is usually violated, in practice a satisfactory point cloud can often still be obtained with this method, as will be shown in Section 5.

When the robot uses the optical pre-touch sensor to scan, it again orients the front of its fingertip orthogonal to the trajectory and towards the object. At each sample point along the trajectory, the robot uses the sensor module at the tip of its finger to measure the distance to the object. Unlike the electric field sensor, the only actuation required is movement along the trajectory because the distance to the object at each sample point is directly reported by the optical sensor.

V. EXPERIMENTS

The following two experiments were undertaken to measure the ability of pre-touch sensing to aid robots in sequential manipulation. In both experiments, 1D pre-touch scans are used to estimate object pose by comparing the collected data to a reference model. By re-estimating the object's pose, the robot can negate previous manipulation errors and more accurately perform subsequent manipulations.

A. Rubik's Cube Manipulation Evaluation

In order to determine the effectiveness of pre-touch scanning in the context of Rubik's cube solving, a system (as briefly described at the end of Section 4a) for manipulating the cube was created in which pre-touch sensing could be enabled or disabled. When pre-touch is disabled, the system serves as a baseline for what is achievable without pre-touch sensing. Instead of scanning the cube after each re-grasp, the baseline system just assumes that the robot re-grasped the cube in the exact desired location.

1) Setup: We generated 10 random cube configurations that required between 20 to 23 rotations for the system to solve, and had both the baseline and pre-touch enabled versions of the system attempt them. In addition to reporting the success/failure rate, the robot's estimate of the cube's position throughout each trial is examined for both methods. Prior to each re-grasp, the robot's estimate of the cube position was recorded. All pose estimates were transformed into the frame of the gripper currently grasping the cube. In order to get ground truth measurements of the cube's position, an AR tag was attached to each face of the cube. We then used a Kinect mounted to the robot's head and cameras external to the robot to detect and estimate the position of the cube when appropriate. An AR tag was also added to each of the robot's grippers at a fixed distance away from that gripper's coordinate frame. This allowed us to compute the pre-touch enabled pose estimate (and corresponding ground-truth) of the cube without using the robot's coordinate transforms. Although our robot was re-calibrated prior to beginning this work, there was still significant error in the coordinate transforms (as there would be for any calibration of a high degree of freedom robot). Despite the use of AR tag detection as a 'ground truth' estimate of the pose of the cube, we are not implying that this method is better than pre-touch sensing for pose estimation. This method will have its own errors depending on how well the tag is detected, and has the disadvantage of requiring one or more tags to be placed on any object whose pose is to be estimated.

TABLE I: End-to-end Rubik's Cube Solving

Method \ Result	Success	Fail	Avg. Rotations Completed
Baseline	0	10	9.6
Pre-touch	8	2	20.1

2) End-to-end Results: The experiment demonstrated that the robot's ability to solve the Rubik's cube was significantly enhanced by pre-touch sensing, as shown in Table I. Using the pre-touch enabled method, the robot successfully solved 8 out of 10 cube configurations. As for the two failure trials, the robot finished 14 and 19 rotations out of 21 and 20 total required rotations before failing to complete a rotation. On the other hand, the baseline method did not successfully solve any of the puzzles. The maximum number of successful rotations for any of the 10 trials was 17, the minimum was 3, and the average was 9.6. All of the unsuccessful rotations occurred when the robot failed to re-grasp the cube; either as it tried to transfer the cube from one gripper to the other, or as it attempted to grasp the cube in order to rotate a face. These results demonstrate that although re-grasping motions are very sensitive to positional error, pre-touch sensing allows the robot to effectively compensate for them.

3) Intermediate Pose Estimation: Comparing the robot's positional estimate of the Rubik's cube for both methods to the ground truth values also yields interesting results. The robot re-estimated the pose of the cube prior to each re-grasp. For each pose estimate, the error was calculated as the root-mean-square deviation (RMSD) between the x and z axes of the estimate and the recorded ground-truth. The RMSD throughout each trial for both methods are summarized as box plots in Fig. 4.

The observed errors demonstrate that pre-touch sensing allowed the robot to significantly reduce the amount of error in the robot's estimate of the pose of the Rubik's cube. The robot has a finger width that is approximately equal to the size of a sub-cube of the Rubik's cube. Given that the desired grasp point is exactly in-between two sub-cubes, the margin of acceptable error is approximately half the width of a sub-cube. More error than this could cause the robot to unintentionally constrain one of the faces, causing a future manipulation to fail. Alternatively, error above this threshold could cause the robot to fail to even grasp the cube. The left plot of Fig. 4 demonstrates that for most trials, a large portion of the baseline method's pose estimates had an RMSD significantly larger than half of the size of a sub-cube. In contrast, the right plot shows that the RMSD of pre-touch aided estimates very rarely went above this threshold. In fact, for the pre-touch enabled method, most pose estimates had an RMSD of less than 0.8cm. Thus, through the use of pre-touch scanning, the robot has been able to limit the error in its estimate of the cube's pose and thereby solve it more robustly.

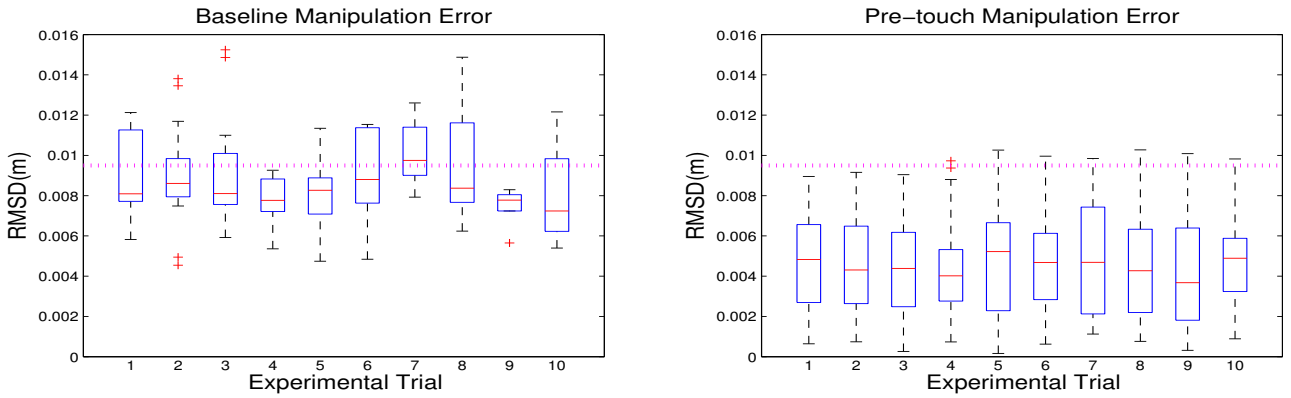

Fig. 4: Box plots plots of positional error for the baseline (left), and corrected pre-touch (right) methods. Each box corresponds to one of the 10 trials and consists of all cube pose RMSD errors observed during that trial. The RMSD error is recorded prior to each re-grasp. The horizontal line across each plot denotes half of the dimension of a sub-cube, demonstrating that the increased dexterity provided by pre-touch sensing is significant for this task.

B. Extension of Pre-touch Scanning to Common Objects

While pre-touch sensing was found to be very applicable to manipulation of the Rubik's cube, the following experiment examines how simple pre-touch scanning can be used in more general manipulation tasks. Here, we estimate the pose of an object by using ICP to match a single, simple pre-touch scan to a corresponding reference model. A future manipulation system could then utilize the transform between the pre-touch scan and reference model to compute the object's pose. Such an estimation may be useful for manipulation when the object is initially grasped, and/or as the object is re-grasped throughout the manipulation task.

1) Setup: We evaluate the use of pre-touch scanning on seven common objects that could require sequential manipulation: a metal bowl, banana, lemon, coffee can, hammer, bell pepper, and a glass soda bottle. Each object is separately scanned by an electric field pre-touch sensor and an optical pre-touch sensor as described in Section 4b. Each sensor has an operating regime in which it is most effective. In particular, the optical pre-touch sensor can sense non-transparent objects well, while the electric field sensor is limited to sensing objects with a dielectric constant significantly different from that of air. Furthermore, the optical sensor has a long, narrow sensing region while the electric field has a short, wide sensing region. It is conceivable that a future system could combine measurements from both sensors to obtain more accurate estimates over a wider set of objects than either sensor alone. For now, we explore each sensor's individual ability to provide geometrically discriminative features for pose estimation through pre-touch scanning.

2) ICP Matching Results: After obtaining a pre-touch scan of an object, we estimate its pose by matching the scan to a point cloud reference model. We used Kinect Fusion [23] to create each reference model, and the PCL library's ICP algorithm [24] to do the matching. The results of each scan are shown in Fig. 5 and the fitness scores (where lower scores indicate better matches) are given in Table II. However, fitness score is not always a clear indicator of successful matching. Ultimately, we qualitatively specify the result of the matching by examining if the pre-touch scan matches to the correct region of the object.

TABLE II: ICP Fitness Score

Sensor \ Object	1	2	3	4	5	6	7
E-field(10^{-6})	6	89	7	17	9	15	24
Optical(10^{-6})	23	24	6	24	286	33	210

The electric field sensor performed particularly well on two objects - the bowl and the hammer - most likely because they are made of metal. In fact, for these objects, the electric field sensor outperformed the optical pre-touch sensor, which failed to even find a match for the hammer. Although the optical sensor's raw point cloud of the hammer captures its general shape, we believe that a number of poor samples at the claw end of the hammer caused the matching to fail. These bad readings could have been caused by the reflectivity of the hammer head, the more complex geometry at the claw end, or a combination of the two. Although the coffee can is also made of metal, the electric field sensor did not perform as well, potentially due to the internal coffee powder not being uniformly distributed throughout the can.

Visually, the raw point clouds of the fruits created by the optical pre-touch sensor capture the profile of the corresponding fruit. However, the scan of the banana did not correctly match its reference model, most likely due to a failure to capture distinctive features in the pre-touch scan. The electric field sensor failed to get a distinguishable outline of the bell pepper, potentially due to the bell pepper being mostly hollow inside. Its raw scans of the other fruits capture the general shape of both the lemon and the banana, albeit with less fidelity than the optical sensor.

When scanning the soda bottle, the electric field sensor captured the shape of the body and the gradual transition

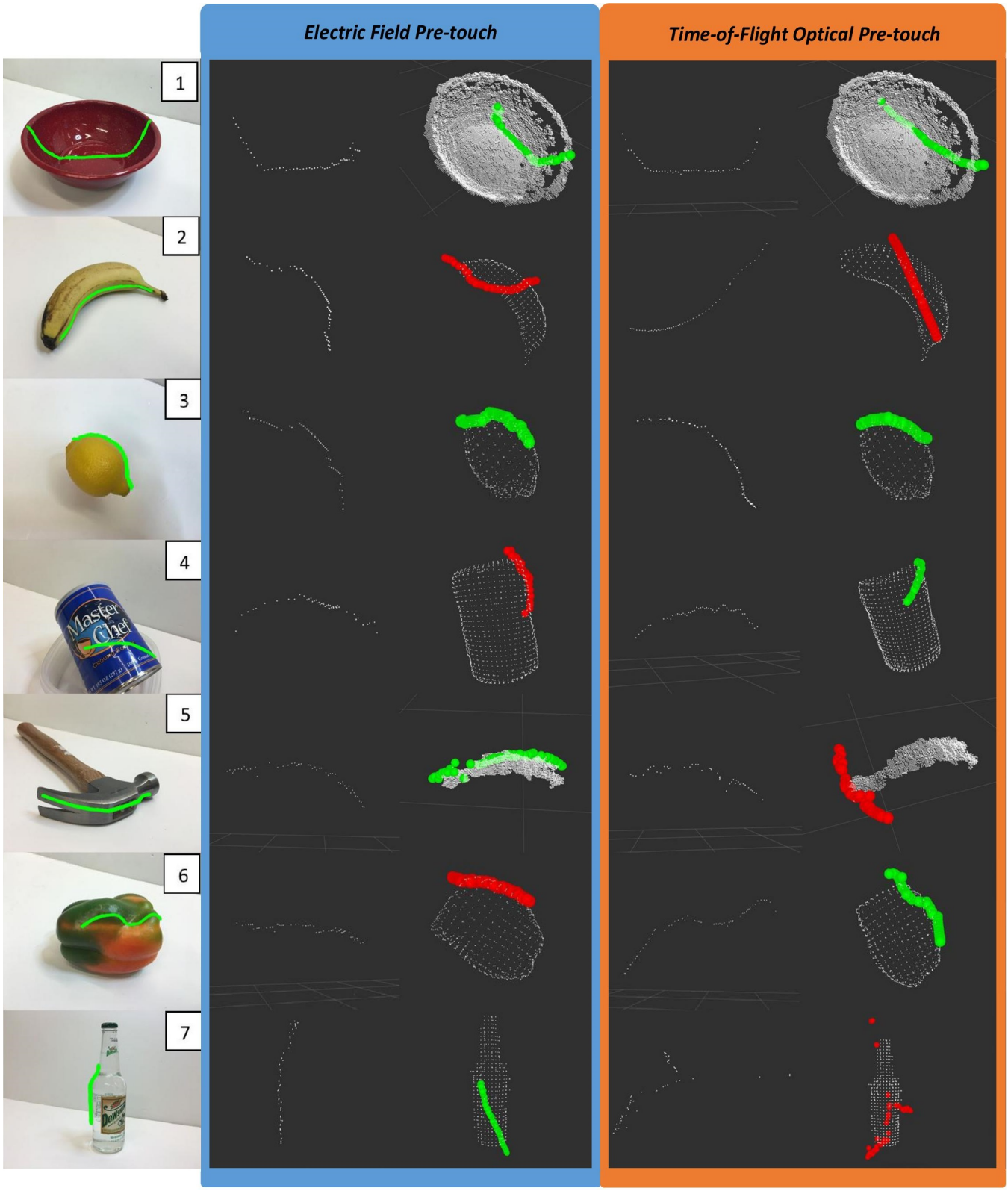

Fig. 5: The results of applying pre-touch scanning and ICP to seven common objects. The first column shows each of the seven objects and the region that was scanned in green. The second column shows the raw point cloud obtained by the electric field sensor for each object, and the third column shows how the ICP algorithm matched the raw point cloud to the reference model. The raw point cloud is colored green if the ICP algorithm found a correct partial or full match, and red if it failed. The fourth and fifth columns display the analogous results for the optical pre-touch scans.

from the body to the neck well. The matching gives a very rough estimate of the pose of the bottle. Note that the reference model was obtained by wrapping the bottle in opaque tape because transparent objects are difficult for the Kinect to sense. Accordingly, the optical pre-touch sensor did very poorly when scanning the uncovered bottle.

This experiment has demonstrated that for many objects, a single, simple scan is sufficient to get a good estimate of the pose of the object. Furthermore, only the electric field scan of the bell pepper and the optical scan of the bottle completely failed to retrieve any geometric information. This is encouraging because when one sensor completely failed, the other was able to capture at least some geometric information. Furthermore, it is possible that a method specifically designed to match these types of 1D scans to reference models would produce even better results than off-the-shelf ICP. It is therefore feasible that a system employing multiple scans with both of the sensors could have a robust ability to recover the pose of a wide range of objects.

VI. CONCLUSION

This work presented methods for using pre-touch scanning to help robots perform sequential manipulation. We first developed a new optical time-of-flight pre-touch sensor that is composed of inexpensive components. We then showed that this pre-touch sensor allows the robot to precisely re-estimate the pose of the Rubik's cube, endowing the robot with the dexterity necessary to robustly solve the cube. Finally, through the use of ICP, we extended pre-touch scanning to pose estimation of seven common objects and showed that even a single, simple scan can capture enough geometric information to perform the estimation well.

Now that we have demonstrated the utility of pre-touch scanning in achieving sequential manipulation, there are several directions in which we would like continue our work. One area is to explore how the robot can determine where to scan and how to generate a trajectory that has a high probability of capturing useful geometric information. In conjunction with this direction, we are also interested in how robots can use electric field sensors to create electric field images. Furthermore, we will aim to develop metrics for the quality of an executed scan. Such metrics will be important if the robot is trying to manipulate objects that have the potential to not be well detected by one or more of its sensors. Finally, we would like to examine how multiple modalities of pre-touch scanning can most effectively be fused together for the purpose of robot manipulation.

REFERENCES

[1] C. C. Kemp, A. Edsinger, and E. Torres-Jara, "Challenges for robot manipulation in human environments," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, p. 20, 2007.

[2] T. Rokicki, H. Kociemba, M. Davidson, and J. Dethridge, "The diameter of the rubik's cube group is twenty," *SIAM Review*, vol. 56, no. 4, pp. 645–670, 2014.

[3] J. Maitin-Shepard, M. Cusumano-Towner, J. Lei, and P. Abbeel, "Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 2308–2315.

[4] L.-Y. Chang, J. R. Smith, and D. Fox, "Interactive singulation of objects from a pile," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3875–3882.

[5] N. Vahrenkamp, S. Wieland, P. Azad, D. Gonzalez, T. Asfour, and R. Dillmann, "Visual servoing for humanoid grasping and manipulation tasks," in *Humanoid Robots, 2008. Humanoids 2008. 8th IEEE-RAS International Conference on*. IEEE, 2008, pp. 406–412.

[6] D. Kragic, H. I. Christensen *et al.*, "Survey on visual servoing for manipulation," *Computational Vision and Active Perception Laboratory, Fiskartorpsv*, vol. 15, 2002.

[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," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 3988–3993.

[8] A. Petrovskaya and O. Khatib, "Global localization of objects via touch," *Robotics, IEEE Transactions on*, vol. 27, no. 3, pp. 569–585, 2011.

[9] K. Hsiao, P. Nangeroni, M. Huber, A. Saxena, and A. Y. Ng, "Reactive grasping using optical proximity sensors," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 2098–2105.

[10] A. Maldonado, H. Alvarez, and M. Beetz, "Improving robot manipulation through fingertip perception," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. IEEE, 2012, pp. 2947–2954.

[11] D. Guo, P. Lancaster, L.-T. Jiang, F. Sun, and J. R. Smith, "Transmissive optical pretouch sensing for robotic grasping," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 5891–5897.

[12] J. R. Smith, E. Garcia, R. Wistort, and G. Krishnamoorthy, "Electric field imaging pretouch for robotic graspers," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE, 2007, pp. 676–683.

[13] R. Wistort and J. R. Smith, "Electric field servoing for robotic manipulation," in *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, 2008, pp. 494–499.

[14] B. Mayton, L. LeGrand, and J. R. Smith, "An electric field pretouch system for grasping and co-manipulation," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE, 2010, pp. 831–838.

[15] S. Muhlbacher-Karrer, A. Gaschler, and H. Zangl, "Responsive finger-scapacitive sensing during object manipulation," in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, 2015, pp. 4394–4401.

[16] L.-T. Jiang and J. R. Smith, "Seashell effect pretouch sensing for robotic grasping," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 2851–2858.

[17] A. Leeper, K. Hsiao, E. Chu, and J. K. Salisbury, "Using near-field stereo vision for robotic grasping in cluttered environments," in *Experimental Robotics*. Springer, 2014, pp. 253–267.

[18] G. Kahn, P. Sujan, S. Patil, S. Bopardikar, J. Ryde, K. Goldberg, and P. Abbeel, "Active exploration using trajectory optimization for robotic grasping in the presence of occlusions," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 4783–4790.

[19] B. Calli, A. Walsman, A. Singh, S. Srinivasa, P. Abbeel, and A. M. Dollar, "Benchmarking in manipulation research: Using the yale-cmu-berkeley object and model set," *Robotics & Automation Magazine, IEEE*, vol. 22, no. 3, pp. 36–52, 2015.

[20] C. Zieliski, T. Winiarski, W. Szynkiewicz, M. Staniak, W. Czajewski, and T. Kornuta, "Mrroc++ based controller of a dual arm robot system manipulating a rubiks cube," Citeseer, Tech. Rep., 2007.

[21] L. Riano. (2011) pr2 rubiks solver. [Online]. Available: https://github.com/uu-isrc-robotics/pr2_rubiks_solver

[22] *Proximity and ambient light sensing (ALS) module*, ST Microelectronics, 8 2014, rev. 6.

[23] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*. IEEE, 2011, pp. 127–136.

[24] D. Holz, A. E. Ichim, F. Tombari, R. B. Rusu, and S. Behnke, "Registration with the point cloud library: A modular framework for aligning in 3-d," *IEEE Robotics & Automation Magazine*, vol. 22, no. 4, pp. 110–124, 2015.