A Robot with Chopsticks: How do Interfaces and Expertise affect Demonstrations?

Liyiming Ke\textsuperscript{1}, Ajinkya Kamat\textsuperscript{1}, Jingqiang Wang\textsuperscript{2}, Tapomayukh Bhattacharjee\textsuperscript{1}, Christoforos Mavrogiannis\textsuperscript{1}, Siddhartha S. Srinivasa\textsuperscript{1}

Abstract—Humans have proved to be remarkably effective using chopsticks to perform a variety of manipulation tasks. Inspired by human performance, we focus on manipulation tasks requiring the use of a robot manipulator with a chopsticks-equipped end-effector. We present a novel teleoperation interface that maps the tracked motion of a human-controlled pair of chopsticks to the motion of the robot's chopsticks. Our key insight is that leveraging human adaptability in learning how to control the robot through our teleoperation interface could enable the collection of high-quality datasets for robot learning of chopsticks-based manipulation. As a first step towards this goal, we studied 25 subjects in which we investigate the factors governing human performance in chopsticks-based manipulation of everyday-life objects across three methods including our teleoperation interface, motion-capture tracked chopsticks, and normal chopsticks. Findings include: a) humans can teleoperate the robot to solve very challenging manipulation tasks such as grasping a slippery glass ball with a pair of slippery metal chopsticks, without the use of haptic feedback; b) teleoperation in some cases is even preferred over using normal chopsticks, opening up the landscape for collecting on-hardware demonstrations that the robot can directly learn from and c) subjective ratings found the teleoperation interface to be the least comfortable and most difficult to use though it achieved equivalent success rate to other methods.

I. INTRODUCTION

Roboticists build both complex manipulation tools and end-effectors (such as vacuum suction tools and anthropomorphic hands) and simple ones (such as parallel jaw grippers and forks \cite{1}–\cite{7}). Complex tools can simplify complicated tasks, while simple tools are easier to design but can require more diverse manipulation strategies. In this paper, we focus on chopsticks, a simple tool with versatile use cases that many humans are already familiar with. We analyze data quality when collecting demonstrations through the use of motion-capture feedback; teleoperation vs a teleoperation system. Our goal is to acquire the highest-quality dataset to train robots to perform chopstick-based manipulation tasks.

Every day, billions of people use chopsticks for food-related manipulation, i.e., to pick up sushi, eat fried rice, or twirl noodles. One wily and dexterous human even used chopsticks to pick-pocket cellphones \cite{8}. Researchers have adopted the general design of chopsticks for various applications, such as meal assistance \cite{9}, \cite{10}, surgery \cite{11}.

\textsuperscript{1}Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA 98105, USA \{kayke, arkamat, tapo, cmavro, siddh\}@cs.washington.edu

\textsuperscript{2}Department of Mechanical Engineering, University of Washington, Seattle, WA 98105, USA jqw123@uw.edu

Fig. 1: Subjects teleoperating a robot holding chopsticks using a motion-capture-marker instrumented chopstick to pick up different objects.
quality data. To better understand how using mocap-marker instrumented chopsticks and teleoperated chopsticks impact task performance and users’ manipulation strategies, we conducted a within-subjects user study. Our study explored both methods on a pick-and-place task and compared these methods to direct manipulation with normal chopsticks.

We make the following contributions:

- A custom-built, 6 DOF robot manipulator with a pair of chopsticks attached to its end-effector
- A teleoperation interface that maps the tracked motion of human-controlled chopsticks to joint commands for robot to complete chopsticks-based manipulation tasks.
- A study of human subjects using real chopsticks, motion-captured, chopsticks and teleoperated chopsticks to pick up objects of diverse grasping difficulty.
- An extensive analysis reporting the performance variance across a number of variables including user expertise, ease of learning, and grasping task difficulty.

Humans successfully teleoperated our robot to complete a very challenging manipulation task: picking up a slippery glass ball with equally slippery metal chopsticks, without haptic feedback. Our dataset of hundreds of diverse grasping trajectories from users will help us build a framework for learning from demonstrations [19].

II. SYSTEM DESIGN

We designed a pair of chopsticks for motion-capture (“MoChop”) and an interface for users to teleoperate a robot holding chopsticks (“Teleop”). Both methods used the same consumer chopsticks (“Chop”). See Fig. 2a.

A. Motion-Capture Chopsticks

We 3D printed light-weight, ball-shaped markers and wrapped them in reflective material. We mounted 5 such markers onto MoChop. To ensure users could still hold the chopsticks, we placed the markers near the tips and tails of chopsticks at different positions on each stick.

B. Tracking of Chopsticks Pose

To track the motion of MoChop held by users, we used the OptiTrack motion capture system [20] with 11 cameras (Fig. 2b). The system uses optical reflection to track the position of markers. Its tracking accuracy is about 0.04mm, and the tracking updates at 100Hz. From the tracked markers’ positions, we extracted MoChop’s pose.

C. Teleoperated Robot Platform

Users held MoChop (master device) to teleoperate the chopsticks on a robot’s end-effector (slave device). We custom built a 6-DOF robot manipulator, assembled from components provided by HEBI Robotics [21]. We actuated the manipulator using 6 rotational actuators attached at its joints. We attached an actuated pair of chopsticks to the manipulator’s end-effector. The two chopsticks operated on the same plane: one was fixed at the actuator’s body, and the other is attached at the actuator’s output shaft. We used HEBI’s X-Series actuators [21] since they provide built-in controllers for position, velocity, or torque control, running at a loop rate of 1KHz.

D. Controller

We designed a controller that translated tracked MoChop poses to joint commands for the robot. Depending on how the user held MoChop, the two chopsticks were or were not on the same plane. To make the pose of the human chopsticks consistent with the robot end-effector chopsticks, we projected the two chopsticks onto the same plane. This became the target pose that our controller tracked. Our controller guided the robot’s chopsticks to smoothly mimic MoChop’s pose. Upon receiving its pose, the controller first computed an Inverse Kinematics solution for the desired pose. However, since smooth chopstick trajectories for humans could require jumps in the robot’s joint space, we applied a convolution smoother on the obtained pose and got the joint position command. We used a position PID controller to follow the joint position command. To add smoothness to the robot’s movement, we also supplied a joint torque command based on gravity compensation and passed it through a torque PID controller. We added both PID controllers’ PWM outputs to compose the final command.

III. EXPERIMENTS

We conducted an in-lab user study with human subjects in which participants performed a series of pick-and-place

\footnote{The amount of torque commanded was based on the mass of the robot as specified on the manufacturer’s datasheet [21].}
manipulation tasks using chopsticks on a variety of objects. These experiments aimed to compare human performance on pick-and-place manipulation tasks using our teleoperation interface, mocap-marker-instrumented chopsticks, and normal chopsticks. We also wanted to explore how humans adapted to different interfaces during this process. Our object set for the manipulation tasks included objects (foam, chip, nut, pencil and ball) with varying levels of chopstick grasping difficulty. See Table I.

A. Participants

We recruited 25 human participants (13 male, 11 female, 1 non-binary of age $M = 27.28, SD = 7.89$) for our human-subject studies, all of whom were approved by the Institutional Review Board of the University of Washington. The participants all had experience using chopsticks ($M = 14.44, SD = 8.79$ years of experience). Fourteen out of 25 subjects reported having experience with teleoperation, but none was previously exposed to our system. To offset individual differences among users, we chose a within-subjects design, where all subjects performed the same set of grasping tasks under the same conditions.

B. Experiment Procedure

Before beginning the experiments, participant signed a consent form and reported demographic information. They were informed that the experiment was intended to evaluate the interactions with all three grasping methods. Prior to the recorded trial, participants went through a training procedure. First, they held Chop and MoChop and tried to open and close them. Second, they watched the researcher demonstrate how to use teleoperation. They then tried to initiate the teleoperation by matching the orientation of MoChop and the robot’s chopsticks. Subjects finished training by picking up a piece of broken foam a single time using teleoperation.

Subjects then proceeded to recorded trials. All used three methods for grasping, shown in Fig. 2a: chopsticks (“Chop”), motion-captured chopsticks (“MoChop”) and teleoperated chopsticks (“Teleop”). All methods used titanium chopsticks that differed only in color. Participants manipulated Chop and MoChop to directly pick up items; for Teleop, they held MoChop to teleoperate the robot to pick up items. All subjects tried to pick up five different items. For each combination of item and method, they had three trials. Each user therefore contributed 45 trials in total ($5 \times 3 \times 3 \times 3$ trials).

During each trial, participants were asked to pick up a specified object and hold it statically in the air for 1 second to show the grasp was firm. If the chopsticks moved the object without procuring it, or the object immediately slipped away from chopsticks after being picked up, the trial failed. However, the subject was allowed to re-try the task until successfully procuring the object or 20 seconds have passed since the trial began. The trial ended when either condition occurred and the experimenter would stop the trial. We chose 20 seconds as the maximum trial length because we wanted to (1) give subjects an opportunity to learn from interacting with the object, and (2) to control the total trial time so the subject would be less likely to feel frustrated.

Each subject worked with a randomized order of objects. For each object, the subject used a randomized order of methods for pick up. All methods for one object were completed before subject moved on to another object. Upon finishing all tasks for one object, they rated the difficulty and comfort of each method on a 5-point Likert scale.

Upon completing all trials, participant responded to an open-end post-task questionnaire. Samples of all questionnaires (pre-task, during-task, post-task) are available in [22].

C. Data Acquisition

We recorded all trials using two RGB cameras and collected written questionnaires from users. We tracked and recorded mocap chopsticks movements during both MoChop and Teleop trials. We recorded joint commands and robot states during teleoperation trials.

IV. Analysis

Among the factors we analyzed were:
- Did subjects perform differently with different objects?
- How did different methods of collection affect subjects’ performance?
- Did different methods result in different manipulation strategies?
- Did subjects themselves affect collected data?
- Was the teleoperation system interface user friendly?
A. Subjective rating of object difficulty matched objective performance

In Fig. 3, the variation in participants’ success rates shows that the selected set of tasks presents an variance of difficulty (ANOVA $F=38.49$, $p < 0.001$), ranging from very easy (foam) to very difficult (ball). The ranking derived from subjective ratings of difficulty roughly correlates with the corresponding ranking of performance, with the only exception being the order between nut and pencil. Paired-T tests on subjective ratings across objects were all significant ($p < 0.0001$). Paired-T test on success rates across objects were all significant ($p < 0.01$) except between nut and pencil ($p = 0.65$).

![Graph](image.png)

Fig. 3: Success rate (gray bar) and subjective rating (multicolor bars) of difficulty for each object. From left to right, objects were easy to difficult.

B. What factors affected the success of grasping?

We evaluated how participants’ performance varied by (1) grasping method, (2) chopstick expertise, (3) effect of learning by interacting with the environment, and (4) the ability to repair failed grasps by re-trying.

a) Different grasping methods affected the success rate for each object. Fig. 4 depicts the success rate per method and how performance varied per method for each grasping task. Overall performance was similar ($F=0.8256$, $p=0.44$) between the grasping methods, although teleoperation performed slightly better. However, for each individual object, we observed statistically significant differences across methods ($F > 3$, $p < 0.04$). Chop and MoChop were significantly better ($p < 0.005$) at picking up foam, while Teleop was significantly better at picking up a nut ($p < 0.05$). Foam was the simplest task and participants had 100% success rate when directly using Chop and MoChop. They were familiar with the environment, and more specifically, the interaction between chopsticks and objects; unfamiliarity with teleoperation might explain the drop in success rate for this object. A nut, on the other hand, is a flat and slippery item that required a firm grasp after being picked up. Failures occurred mostly because the nut slipped from chopsticks, possibly because users were tired of supplying a concentrated force on their fingers. Teleoperation could potentially have alleviated this problem by offering a firm grasp without participants supplying force, thus yielding a higher success rate.

b) Chop expertise doesn’t imply Teleop expertise.: We identified expert subjects by choosing individuals that achieved higher than median success rates when using Chop to pick up items ($SR_{\text{exp}} > 60\%$). Based on this standard, our sample comprised 11 expert and 14 non-expert subjects. Fig. 5 depicts the performance of expert and non-expert chopstick subjects. Both cohorts had statistically significantly different success rates when using Chop ($p < 0.001$) but not when using MoChop or Teleop ($p = 0.06, 0.89$ respectively). Non-experts had a statistically significant improvement using Teleop relative to using Chop ($p < 0.01$). We suspect that non-expert participants may have had less stable and precise control of chopsticks, which would be critical in directly picking up a nut, for example. However, teleoperation removed this requirement and therefore enabled nonexperts to perform better. This suggests that teleoperation can be a desirable interface for future data collection, especially for collecting data on certain challenging tasks for humans, such as prolonged grasping, as researchers can improve both hardware and controllers to alleviate the burden of controls from users. We observed that some subjects had to hold chopsticks differently when holding MoChop.

c) Participants improved performance within 3 trials.: How quickly did subjects learn to adapt to the use of chopsticks for the trials? We look at the change in success rate over the three trials as a proxy for estimating the effect of learning on participants’ performance. For each task (5 in total) and each method (3 in total), each subject attempted 3 trials. Fig. 6 depicts the success rate per trial number. As the trial count increased, performance increased, suggesting the existence of a learning effect ($F=5.47$, $p < 0.001$). The variance in performance across trials was significant for Chop and Teleop methods ($F = 3.79$, $p = 0.027$, and $F = 3.6$, $p = 0.032$ respectively), suggesting that subjects had an easier time adjusting to these two methods through trials. Additionally, subjects significantly improved performance over trials when picking up a nut and pencil. For the nut, they realized the need to supply a firm grasp after failed trials. For the pencil, they learned that its center of mass is actually towards its tail. Subjects usually mastered a more effective grasping point that was closer to CoM after experiencing failures.

d) Given chances to re-try after a failed grasp, participant success rates can increased to above 75% even for the most difficult task.: Many robotics grasping tasks evaluate success based on one-shot grasping, i.e., whether the robot successfully grasps the object in one try. However, humans learn from and adapt to failures. In our experiment, we let subjects re-try a task after a failed trial, even though the first failed attempt might have changed the object’s configuration. We evaluated re-try success rates based on whether subjects could pick up the same object within 20 seconds (see Fig. 7.) We see a significant boost in the success rate where ,
Fig. 4: Success rate for each method. Error bars indicate 95% confidence interval. No significant variance was observed across the three methods for all objects \( (p = 0.44) \). However, for each individual object, we observed significant variance across methods \( (p < 0.05 \text{ in ANOVA}) \).

within 20 seconds, subjects managed to pick up the most difficult item (ball) with a 79.1% success rate even though their initial success rate at first try was only 26.2%. Humans demonstrated an impressive robustness, which suggest that alternative metrics for evaluating manipulation performance might consider allowing successive tries.

C. How did participants adjust their manipulation strategies?

a) Subjects applied different strategies for different shapes of objects.: To adapt to the curved surface of chips, 21 of 25 subjects rotated the chopsticks procure the object. However, for a flat object like nut, all subjects held the chopsticks parallel to the nut from the top view, as shown in Fig. 8a.

b) For different methods, participants used different strategies for the same object.: Fig. 8b shows a subject using MoChop to pick up the nut from a different angle than was used via Teleop. The different strategies might also contribute to the performance difference in addition to the reason we discussed above. We observed similar phenomena for chips: many participants used Teleop to grip the chip on one pressure point and used MoChop and Chop to lock the chips between chopsticks. For the glass ball, they needed to first grip it and then lift it up while retaining grasp strength. However, the glass ball frequently slipped during lifting. When using Teleop, subjects seemed to pay more attention to gripping the ball and less to keeping the grip firm.

c) Participants changed strategies after failure.: As shown in Fig. 8c, the long pencil shape made it tricky to pick up. Most subjects tried to grip its center. Because some did not estimate the CoM accurately, the pencil rotated and dropped. After failure, many subjects kept adjusting the gripping point on the pencil and the contact point on the chopsticks. Interestingly, some had a different adaptation strategy when using Teleop: they adjusted how much they closed their master chopsticks (and how much force the robot output from slave chopsticks) until the robot could grasp the pencil firmly.

D. Subjective Ratings

On average participants rated teleoperation as the least comfortable and the most difficult to use \( (p < 0.0001) \), except when to pick up the ball. Participants explained that the negative ratings came from the sense of indirection added by the teleoperation interface, lack of haptic feedback and the misalignment between the robot arm and human arm. However, participants’ performance using teleoperation was empirically better than using other methods to pick up 3 out of 5 objects chosen (chips, nut and ball shown in Fig. 4). Possible reasons of such a contrast include (1) teleoperation alleviated the burden of firm and stable control from the subject and that (2) subjects found a strategy that worked better for teleoperation method. A third possibility could be perhaps believing that the teleoperation interface was more difficult and less comfortable made subjects spend more effort and concentration using this method. We found qualitative evidence supporting first two hypothese: (1) 15 of 25 subjects in the post-task questionaire reported that Teleop simplified grasping because it “(was) easier to maintain a constant grip,” “(required) less effort on my hand to grip the object between the chopsticks,” and ”(users) only need to care about the motion of chopsticks without, considering how much force to exert”, and (2) 19 of 25 subjects developed a different strategy for Teleop compared to the other two methods for grasping the same object. Some subjects commented on how they adapted to work with the Teleop system: “It takes a little time to learn and be familiar with the movement (of Teleop),” and “I started focusing on the robot joints and how to move my arm in a way that translated to the robot joint, the placement tasks became easier.”
E. Position as a proxy for force

One main benefit of our system is its ability to collect force commands without needing specialized force measurement devices. We achieved this by using position as a proxy for force. Subjects could close the master chopsticks to direct the robot to close its chopsticks. However, the robot’s chopsticks could hold a ball between its tips rather than closing the chopsticks as instructed. The error between commanded and the actual position incurred a proportional torque output from our PID controller. To pick up any object that requires gripping, this mechanism is necessary. On post-task questionnaires, subjects commented on how firm the Teleop grasp and that “(they) didn’t have to squeeze that hard to increase the tension.” Fig. 9 shows the recordings of displacements, which indicate the generation of forces. Note that the clear displacement gaps for the nut and ball correspond to subjects gradually closing master chopsticks to add force to firmly grip those objects.

V. DISCUSSION

We observed that the chosen set of objects yielded significant variance in task performance and human strategies. Despite being rated the most difficult and least comfortable interface by users, Teleop achieved a higher success rate on 3 of 5 objects. We hypothesize that this might be due to: (1) relaxing the burden of control from subjects and facilitating stable grasping, and (2) human adaptability and effectiveness in learning to use a novel teleoperation interface. Supporting evidence includes the fact that (1) the three objects that Teleop had a better performance on were Chips (brittle), Nut (hard, slippery) and Ball (small, hard, slippery) which demand precision and stability in control, and have been challenging for Chop and MoChop, (2) we observed gradual closing of chopsticks yielding higher forces in Teleop and (3) qualitatively, 20 out of 25 users reported that Teleop made picking up certain items easier as they said “The grip seemed firmer than the grip that I could produce”, “The support (from Teleop) was good enough for me to be a bit careless” and “(It) holds small objects robustly (in a way I cannot do easily with real chopsticks)”. Regarding (2), we observed that users adapt their strategies for teleoperation and achieved higher success rates as trials progressed.

Regarding the ineffectiveness of Teleop on foam, we hypothesize that even though users are familiar with chopsticks and can adapt to different grasping methods and tasks: (1) it is still a challenge for users to control an arm that is not configured like their own, (2) the more expert one were with normal chopstick use, the harder it might be to adapt to Teleop, and (3) the system’s movement and range is constrained by the robot’s configuration. For (1), we hear from users that “Movements of Teleop ... were opposite of what I’d naturally want to do”. For (2), we see that Teleop couldn’t boost expert users performance as much as it did to non-experts. For (3), we observed occasionally robot’s IK solution jumped when users rotated their MoChop, after which the users adapted to move MoChop in a more constrained way.

Teleop has proved to be more helpful in manipulation tasks involving hard, small and slippery pieces that demand precision and stability in control than soft, compliant, lightweight objects. Human subjects have demonstrated in our study their impressive capability to serve as the controller of an imperfect hardware that is not precise to control, can counter their intuition during usage and lacks the haptic feedback that they are familiar with to achieve challenging manipulation task. This brings out the potential of involving chopsticks in robotics manipulation and have the robot pick up challenging items without relying on a complicated end-effector. We intend to extend this work by learning from the demonstrations we collected and build a robot that is skillful at using chopsticks autonomously.
VI. ACKNOWLEDGEMENT

This work was (partially) funded by the National Institute of Health R01 (R01EB19353), National Science Foundation CPS (1544797), National Science Foundation NRI (1637748), the Office of Naval Research, the RCTA, Amazon, and Honda Research Institute USA.

REFERENCES