Human reach-to-grasp compensation with object pose uncertainty

Qiushi Fu, Arash Ushani, Leif Jentoft, Robert D. Howe, Marco Santello

***Abstract*—This paper examined how human alter reach-to-grasp behavior to compensate environmentally-induced object orientation uncertainty. We used a novel motion tracking framework to capture hand-object interactions, as well as a custom-made cylindrical object to detect contacts. Subjects were instructed to reach, grasp, and lift the object with or without vision. The orientation of the object was randomly changed on each trial. We hypothesized subjects would use reach-to-grasp strategy that minimize post-contact adjustments. However, our results indicate that (1) subjectsare more likely to use the hand as a sensing apparatus prior to contact, and (2) the reach-to-grasp kinematics may be optimized for efficient sensing of object orientation. Our findings could provide potential solution to efficient tactile sensing for robotic hand in unstructured environment.**

I. Introduction

It has been demonstrated that optimal control theory can be used to explain many behavioral data on the ability ofhuman sensorimotor systemto compensate for variability and uncertainty [1]. Most of the research on optimal sensorimotor controlhas used pointing and reaching tasks. However, few studies have investigated motor compensation during reach-to-grasp movement. These tasks are more complex because the purpose of these movementsis tomake contact with objects,which requires anticipation the task constraints, and control the hand to position the fingers on specific positions for subsequent manipulation[2]. It could be challenging to compensate environment-induced uncertaintyin object poseduring reach-to-grasp behavior, because of the high dimensional nature of hand-object interactions[3]. Less is known about how human adjusts reach-to-grasp kinematicsin response to uncertainties. Recently, it was found that, with 2D position uncertainty, subjects adjust the approaching direction of the fingers to maintain grasp efficiency. The adjusted finger kinematics was thought to provide better chance of stable two-digit grasp at first contact, thus minimizing post-contact movements[4].This strategy has been later implemented in robotic grasping [5].

The uncertainty in daily grasping tasks usually results from lack of accurate visual feedback. Vision of the hand and object could be removed due to darkness or simply paying attention to other visual target(s). However, two-digit precision grasps are usually controlled with visual feedback. Therefore, it is important to study how human compensates uncertainty in a more naturalistic scenario with whole hand grasping. In this paper, we investigate a case in which subjects were instructed to reach and grasp a cylindrical object with all five digits under object orientation uncertainty. The kinematics and contact events of the hand-object system were measured using a novel algorithm and contact sensors. We hypothesized that subjects would adjust their reach-to-grasp movement to minimize post-contact adjustments similar to what was found in [4]. However, our results indicate that post-contact adjustments were not minimized. Instead, we found that subjects might have used the hand to sense the actual object orientation until first contact occurred, and then repositioned their fingers for subsequent manipulation.

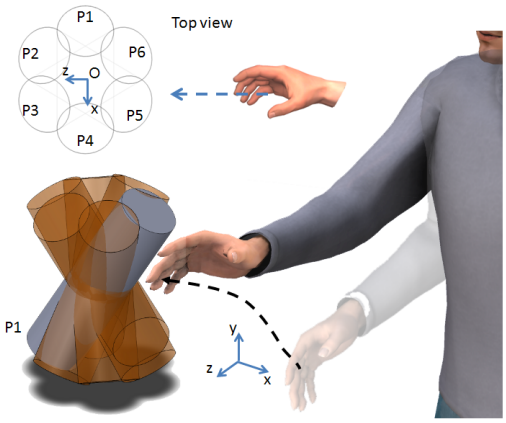

Figure 1. Experimental setup. The grey cylinder is the actual object orientation and the brown cylinders are other possible orientations. The circles in the top view are the base positions for six possible orientation.

II. Method

A. Experimental setup

Eight healthy right-handed subjects participated in this study. All participants were naïve to the purpose of the study and gave their informed consent according to the declaration of Helsinki. The protocols were approved by the Office of Research Integrity and Assurance, Arizona State University.

All subjects were instructed to pull the designated object off the base along object's central axis using all five digits of their right hands. The custom-made object was a PVC cylinder (48 mm diameter; 160 mm length) with a protruded peg which sits in a hole on a base. The base can be set to different positions and orientations thereby setting the

Manuscript received February 4, 2013. This work was supported in part by a National Science Foundation Grant IIS 0904504.

Q. Fu and M. Santello are with School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ 85287 USA. (e-mail: qiushifu@asu.edu; marco.santello@asu.edu).

A.Ushani is with Liquid Robotics, Sunnyvale, CA 94089 USA. (e-mail: arash.ushani@liquidr.com).

L.Jentoftand R.D. Howe are with School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138. (e-mail: ljentoft@seas.harvard.edu; howe@seas.harvard. edu).

position and orientation of the cylinder. The cylinder is covered by a grid of 8 × 8 touch sensitive taxels 20 × 20 mm each. Each taxel consists of a copper sticker with conductive adhesive. Voltage signals from each taxel are routed to an analog-to-digital converter on a PIC microcontroller mounted inside the cylinder. The microcontroller serially measures the voltage of each taxel continuously. To determine if a given taxel is being touched, the amplitude of that pixel's voltage is compared to a configurable threshold which is tuned to be subject dependent. When in electrical contact with a human body, the input to the high-impedance amplifier picks up the prevalent 60 Hz noise from nearby AC lines, making even a light touch (< 0.1 N) on a taxel easily detectable. The map of contacted taxels is sent to a host computer via USB. Using this technique, we determined the timing and spatial distribution of digit contacts on the cylinder.

The kinematics of the hand was measured using a motion tracking system (Impulse, PhaseSpace Inc.) with a custom hand-tracking algorithm based on extended Kalman filter. This tracking framework operates at 120 Hz to estimate the 29 DoF spatial positions of all the linkages of the hand. We also tracked the 6 DoF pose of the cylindrical object and estimated the digit making contact with the object at any given time using custom collision detection algorithm [3]. The motion- tracking framework was used as supplement to provide more details about hand-object interaction when the contact was detected (i.e., which finger is in contact).

B. Protocol

With the adjustable base, the object can be positioned to six different configurations. These configurations share two commonalities: 1) the center of the cylinder is always located at a point 30 cm in front of subject's right shoulder, and 2) the angle between the major axis of the cylinder and gravity is always 30 degrees. The six possible configurations are then generated by rotating the object about a vertical axis (direction of gravity) passing through the center of the object by 60 degrees incrementally. For instance, Position one (P1) is a cylinder tilted to the left from subject's perspective (Fig. 1). Subjects were instructed to reach, grasp the cylinder with all five digits (without palm contact) and pull it out of the base along object's central axis. Each subject performed two blocks of 30 trials. In the first block (NV), subject were first told to "close their eyes", then initiate the reach after hearing a verbal "go" signal from the experimenter. The time between the instructions "close your eyes" and "go" was 5 seconds. Subjects were told that the center of the object would always remain at the same position but that the orientation of the object would be set to one of six possible orientations. In the second block (FV), subjects had full vision all the time. The object orientation was randomized for each block of trials but each orientation was presented 5 trials within each block. The subjects were instructed to move at a natural, self-selected speed.

C. Data analysis

For all analyses, the origin of the global coordinate frame was set to coincide with the center of the object, the y-axis was defined as pointing in the direction opposite to gravity, the x-axis was parallel to the subjects' coronal plane, and the z-axis was parallel to the subjects' sagittal plane (Fig. 1). To simplify the analysis, we defined several task-relevant variables to describe in a lower dimensional space of the large number of degrees freedom tracked by the motion capture system. Since the subjects were required to use all digits for each grasp and the target object was a cylinder, we focused our analyses on the following variables (Fig. 2):

1) *Grasp aperture and aperture axis*. The grasp aperture is defined as the distance between the tip of the thumb and the position of the 'virtual finger' (averaged position of four finger tips). The aperture axis is a unit vector that originates from the 'virtual finger' and ends on the tip of thumb. This vector denotes the direction of the opening of the hand.

2) *Grasp axis*. The grasp axis is defined as a unit vector thatoriginates from the tip of little finger and ends on the tip of the index finger. This vector indicates the orientation of the finger side.

3) *Grasp center*. The center of the grasp is defined as the mid-point of the grasp aperture. This point indicates the position of the grasping hand.

These three variables summarize key features of the grasp configurations while treating the four fingers as one 'virtual finger' with variable orientation.

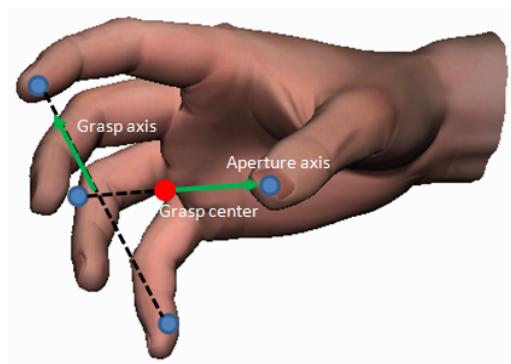

Figure 2.Definition of grasp kinematic variables: Grasp axis, Grasp center, and Aperture axis.

III. Experimental Results

A. Time course of reaching kinematics

Since no visual input is available for NV condition during reach and pull off, subjects had to completely rely on 1) memory of the object center location, 2) proprioceptive sensory information to transport the hand towards the remembered location of the object, and 3) tactile input to detect the actual object orientation. To compensate for the uncertainty in object orientation, subjects made slower reaching movements and used larger grasp apertures compared to the FV condition. In addition, subjects in the NV condition tended to deviate from the straight path earlier and further than in FV condition. These behaviors are shown in Fig. 3 where the reach and grasp kinematics from one subject grasping the object at orientation P1 is shown. These

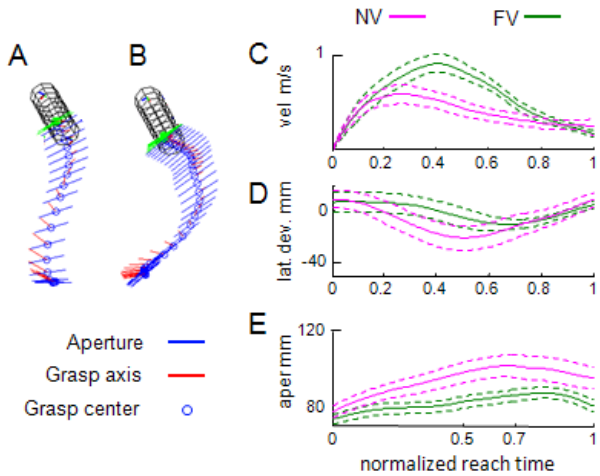

Figure 3.Temporal comparison of main kinematic variables. Panel A and B show the top view of kinematics of the reach-to-grasp movement in FV and NV conditions, respectively, from representative trials. The circle, the blue line and the red line are grasp center, grasp aperture and grasp axis. Panel C, D and E show the time course of the velocity of grasp center, the grasp aperture, and the lateral deviation of the grasp center, respectively. The solid red and blue curves are the mean values averaged across trials for NV and FV conditions, respectively. The dotted curves are the corresponding standard deviation.

observations were confirmed by data averaged across all subjects.To compare the reaching kinematics between two vision conditions, we first normalized the time course of the related kinematic variables between reach onset to initial contact. The reach onset was determined by the zero-crossings of the velocity profile and the initial contact was determined by the touch sensors on the object. The time of peak velocity for FV condition occurred at~40% of the reach whereas the peak velocity for NV condition occurred earlier (~25% of reach, Fig. 3C).Subjects reached peak grasp aperture earlier during reach in NV condition (~60% of reach) compare to FV condition (~ 80% of reach) andthe aperture used in NV condition was larger (Fig. 3D).The time of peak velocity and aperture in FV condition are consistent with previous studies [6]. The time courses of lateral deviation of grasp center along the global x-axis suggest that subjects consistently started to deviate to the right (negative x) earlier in NV condition (~40% of reach) than FV condition (~65% of reach) and the deviation is larger (Fig. 3E).

In addition, in the NV condition the object orientation could be sensed only after subjects made contact with the object. Therefore, subjects had to recover from the initial contact posture to a final posture that was suitable to pull-off the object. As expected, the contact to pull-off time was more than two times longer in NV condition than in FV condition (1048±158ms vs. 347±49ms).

B. *Grasp postures at contact and object lift*

As revealed by different reaching kinematics, subject clearly and consistently used a specific strategy in NV condition. We now ask whether subjects behaved differently at the end of pull-off, i.e.:did subjects use different grasp configuration for NV and FV? To address this question, we applied discriminant analysis. Specifically, we used the grasp axis at object pull-off in FV condition (FV@lift) as training data set to find the discriminant functions for six different orientations. These discriminant functions served as the 'reference grasp posture' subjects tend to use with full vision. The grasp axis at object pull-off in NV condition (NV@lift) was then classified using the discriminant function. A high rate of correct classification would indicate that subjects used grasp configurations similar to those in FV condition for specific orientation. In addition, the established discriminant function constructed with FV@lift data was also used to test the grasp configuration that subjects used at initial contact. The confusion matrices generated by discriminant analysis are shown in Fig.4. For the NV condition at object pull off (NV@lift, Fig. 4C), 92.5% of 240 trials were correctly classified, thus indicatingthat subjects used similar final grasp configurations in both NV and FV conditions. However, the two conditions differed significantly at initial contact. The grasps at initial contact for FV already took the form of final grasp (FV@touch, 91.25% correct classification, Fig. 4A) since subjects had acquired the object orientation before they started reaching, thus allowing accurate planning. In contrast, the grasps at initial contact for NV could not be distinguished very well (NV@touch, 23.75% correct classification, Fig. 4B).Similar classification rates were foundby additional discriminant analyses using all joint angles instead of the grasp axis (not shown due to space limit). These resultssuggest that the grasp configuration at the time of initial contact was not consistent with the final grasp subjects used to pull off the object. Instead, subjectstended to adjust their hand posture significantly *after* initial contact for the subsequent object pull-off.

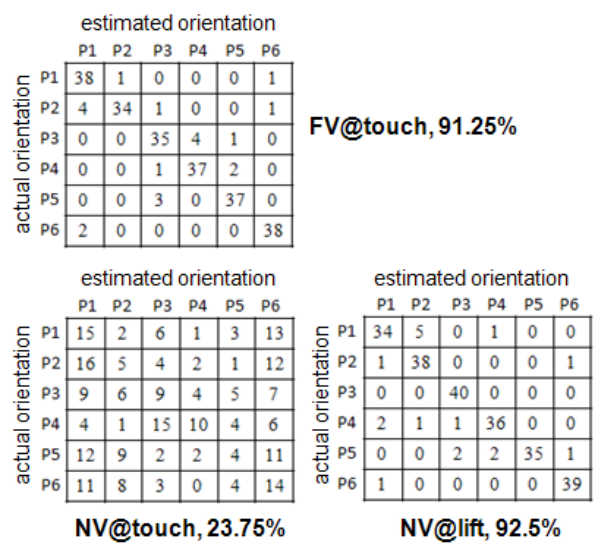

Figure 4.Confusion matrices using discriminant analysis show hand readjustment between touch posture and lift posture. The discriminant function was constructed with hand posture in Full-Vision at object pull-off

C. *Digit contact*

We examined how the hand was used during and after the initial contact by computing which digit made first contact with the object as well as how the post-contact adjustments

were made to converge to an object orientation-dependent final grasp. The first contact digit wasdetermined by first calculating the distance between each phalange and the major axis of the cylinder. The contact digit would have the smallest distance at the time of contact as detected by the sensor array on the cylinder. The finger re-positioning after contact was assessed by calculating the distance that each fingertip traveled from time of contact to time of pull-off. If a digit remained at the position where it landed initially, it should move less during re-position. In contrast, if all digits moved similar distance, it would suggest that subjects repositioned their grasp completely regardless of which digit made the first contact.

We found that index finger was most likely to make the first contact in NV condition (49.1%) whereas the chances for other digits to make first contact were similar (~10%). However, there was no significant difference across digits in post-contact adjustments (one-way ANOVA, $p>0.05$). These resultsdo not support our initial hypothesis and indicate that subject did not need to keep the first contacting finger in touch with the object while adjusting the grasp configuration.Furthermore, these results suggest that the hand might have been used as a sensing apparatus to initially sense the actual object orientation by posing in a way that is efficient for quick sensing. After sensing the orientation, subjects adjusted their hand configuration to the final pull-off posture similar to the posture in FV condition, without necessarily keeping invariant contact with the object.

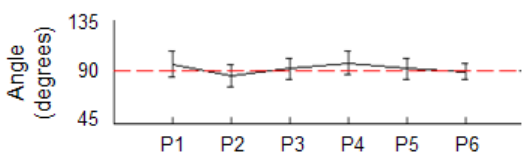

Figure 5.Approaching angle for all six orientations

D. *Optimal approaching angle without vision*

Since we found that the hand might have beenused only as a sensing tool at initial contact, the last analysis was designed to assess the strategy subjects used for reaching in relation to the initial contact. We hypothesized that subjects approached the object in an optimized direction to maximize the possibility of contact given the presence of sensorimotor noise without visual guidance. Because the aperture is the maximum width of opening of the hand by definition, it is desirable to make contact within the range of aperture. This means that the optimal strategy would be approaching the object in a direction perpendicular to the grasp aperture such that the movement error during blind reaching and the uncertainty of object orientation can be accommodated. We define the approaching direction as the unit vector pointing from the position of grasp center measured at 80% of the reach to the position of grasp center at the end of reach(initial contact) for each trial. During this last 20% of reaching movement (0.21 ± 0.03 sec), the grasp center moved 29.6 ± 4.5 mm. We chose 80% because the aperture starts to decrease around this time point (Fig. 3).

The result supports our hypothesis by showing that the angles between grasp aperture and defined approach direction were close to 90 degrees consistently across six orientations (t-test, $p>0.05$; Fig. 5).

IV. Conclusion and Future Work

By using motion tracking of the hand and a contact sensing object, we examined how human compensate orientation uncertainty during whole hand reach-to-grasp movement without online visual guidance. It was found that, unlike previous studies using precision grasp, subjects did not try to minimize post contact adjustment but rather maximized the probability of initial contact within grasp aperture. This might be due to the fact that the task requires post-contact manipulation (object pull-off) with all five digits. It was shown that the task constraint can lead to preferred hand configuration for optimal subsequent manipulation [7]. Instead of finding the best strategy to allow the fingers to land on the object under large uncertainty for minimal post-contact motion, repositioning the fingers may lead to a more optimal behavior over the whole course of reach-grasp-lift movement. Therefore, our results suggest that the reaching movement was adjusted to compensate sensorimotor noise for more efficient sensing of actual object orientation. Since there were only six possible orientations, and the cutaneous receptor can provide rich information about local contact feature with very high speed [8], initial contact maybe sufficient enough to inform the sensorimotor system to select the corresponding final hand posture. Our ongoing work aims at modeling the reach-to-grasp strategies described in this paper. Such strategies might be very useful for robotic hands to select optimal reaching angle for efficient tactile sensing of object contours under uncertainty [9].

References

[1] D. W. Franklin and D. M. Wolpert, “Computational mechanisms of sensorimotor control.,” *Neuron*, vol. 72, no. 3, pp. 425–42, Nov. 2011.

[2] U. Castiello, “The neuroscience of grasping.,” *Nature Reviews Neuroscience*, vol. 6, no. 9, pp. 726–36, Sep. 2005.

[3] Q. Fu and M. Santello, “Towards a complete description of grasping kinematics: a framework for quantifying human grasping and manipulation,” in *IEEE Engineering in Medicine and Biology Society*, 2011, pp. 8247–50.

[4] V. N. Christopoulos and P. R. Schrater, “Grasping objects with environmentally induced position uncertainty.,” *PLoS Computational biology*, vol. 5, no. 10, p. e1000538, Oct. 2009.

[5] F. Stulp, E. Theodorou, J. Buchli, and S. Schaal, “Learning to grasp under uncertainty,” in *IEEE International Conference on Robotics and Automation*, 2011, pp. 5703–08.

[6] M. Jeannerod, “The timing of natural prehension movements,” *Journal of motor behavior*, vol. 16, no. 3, pp. 235–54, 1984.

[7] Q. Fu, W. Zhang, and M. Santello, “Anticipatory planning and control of grasp positions and forces for dexterous two-digit manipulation.,” *Journal of Neuroscience*, vol. 30, no. 27, pp. 9117–26, Jul. 2010.

[8] R. S. Johansson and J. R. Flanagan, “Coding and use of tactile signals from the fingertips in object manipulation tasks.,” *Nature Reviews Neuroscience*, vol. 10, no. 5, pp. 345–59, 2009.

[9] L. Jentoft and R. Howe, “Determining object geometry with compliance and simple sensors,” in *IEEE/RSJ IROS*, 2011, pp. 3468–73.