Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 EECS Building
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 06-003

Internal models for object manipulation: Determining optimal contact locations

Paul R. Schrater and Erik J. Schlicht

February 09, 2006

Internal models for object manipulation: Determining optimal contact locations

Paul R. Schrater [1,2] & Erik J. Schlicht[1]

[1]*Department of Psychology*
[2]*Department of Computer Science & Eng.*
University of Minnesota

Although there are an infinite number of ways that humans can lift an object, they tend to reach in a predictable manner. This suggests that people are aware of, and optimizing, some sort of loss function. This paper outlines a natural loss function that may be used to predict people's actions in everyday reaching tasks. The loss function is based on the physics of object manipulation and the assumption that people are planning for the intended motion of the object. Using this framework, we are able to make predictions about how people should reach if they are minimizing their expected risk. To test the model, we required people to reach to objects at varying orientations. Our experimental results indicate that people are reaching in a manner that minimizes their expected risk for the task. These findings suggest that people planning for the intended motion of the object and that our brain is aware of the physics involved with object manipulation.

Introduction

Natural tasks are usually specified in terms of high-level goals like lifting an object. However, to execute a movement, high-level goals must be converted into a detailed control strategy at the level of forces and torques (or muscle contractions). This problem is difficult because the conversion is underconstrained – there are an infinite number of ways to reach to and lift an object. As a result, understanding how the brain accomplishes this conversion is a major goal in motor control research.

Optimal control theory provides an elegant solution to this dilemma, explaining why people reach in a very predictable and stereotypical manner, even with ambiguity (e.g. Harris & Wolpert, 1998). The basic idea is that a specific movement strategy is selected by minimizing an expected loss function defined on movement parameters and task goals. Despite a considerable amount of effort over the past thirty years to reverse engineer a loss function for reaching (e.g. minimum jerk–Flash & Hogan, 1985; minimum end-point error–Harris & Wolpert, 1998; minimum control–Todorov & Jordan, 2003; etc.), there are still many gaps in our knowledge. For example, one of the key planning problems for lifting an object is where to place finger and hand contacts. Without this information, trajectory plans cannot be formulated. Nevertheless, previous proposals do not address this problem. Instead they focus on understanding trajectory planning after the contact conditions (location, velocity and acceleration) have been experimentally imposed. This approach has yielded a series of loss functions that can successfully predict trajectory data (e.g. minimum jerk–Flash & Hogan, 1985; minimum end-point error–Harris & Wolpert, 1998; minimum cost-to-go –Todorov & Jordan, 2003; etc.). However, specifying good contact conditions is important because they can vary drastically across tasks. For example, in a task where the subject has to lift an object, finger locations determine the relative controllability of the object. Moreover, the velocity and acceleration at contact could be planned to allow the hand's momentum to lift the object.

The goal of this work is to propose a theory for how the brain solves the contact selection problem. The theory maintains that the brain selects contact conditions through an internal model of the physics of object manipulation – how and where forces must be applied to effectively move objects. It will be demonstrated that this knowledge, coupled with a strategy of moving objects with a minimum control strategy, can be used to derive a natural loss function for optimal contact point selection. The empirical validity of these ideas are investigated through human reaching experiments that require subjects to lift or touch objects arranged at different orientations.

The core idea of this proposal is that the strategy for lifting an object begins by planning for the intended motion of the object. It is further assumed that object trajectory planning follows the same minimum control principles as hand and finger trajectory planning. Intuitively, when an object is manipulated, it is functionally an extension of the hand and is (presumably) controlled by the same system that controls an empty hand. Specifically, this theory proposes that the object's motion plan is controlled by encoding the desired object via-points and goal-state (locations, velocities, accelerations) in a loss function on object motion. Furthermore, it is assumed that an optimal control strategy is used to find the set of forces required to generate an optimal object trajectory. However, the required object control forces are a function of finger placement and external forces (like gravity). Using equations of rigid body motion, contact dynamics and kinematics, it will be demonstrated that the object motion loss function can be rewritten in terms of a loss on contact conditions. Optimizing this loss function results in contact conditions that can achieve the desired object motion with minimum force and torque. Finally, these optimal contact conditions induce a loss function on approach trajectories.

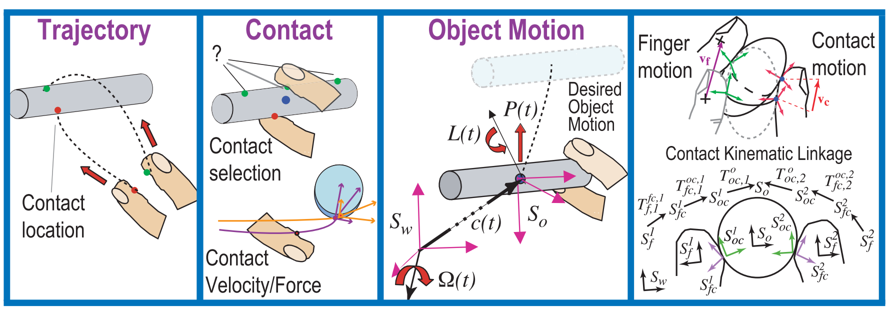

Figure 1: Demonstrates the three components of a reach task. Red dots signify the thumb's position. Green dots represent the index finger's position, and blue dots show the object's center of mass. **(a)** *Approach trajectory.* Optimal control trajectory to contact conditions. **(b)** *Selecting contact conditions.* Select contact conditions that can produce the desired object motion with minimum force and torques. **(c)** *Planning for object's motion.* Optimal control of object motion. Notice that this separation has a hierarchial arrangement (from right to left). **(d)** Illustrates relationships between object motion and finger motions. Given frictional contact without slippage, the object and fingers are connected by a kinematic chain. From this view, contact selection amounts to choosing a particular object-hand linkage.

This proposal for converting a goal for object movement into a motor plan can be thought of as forming a planning hierarchy (Figure 1). At the first level, the task goal is converted into object motion goals. At the next level the intended object motion is used to determine good contact conditions for control of the object. Finally, the contact conditions form the goals for approach trajectory generation. For a given desired object motion, only a subset of finger contact configurations will minimize the risk. Thus we can use the proposed framework to predict what contact conditions people should use to achieve a desired object motion.

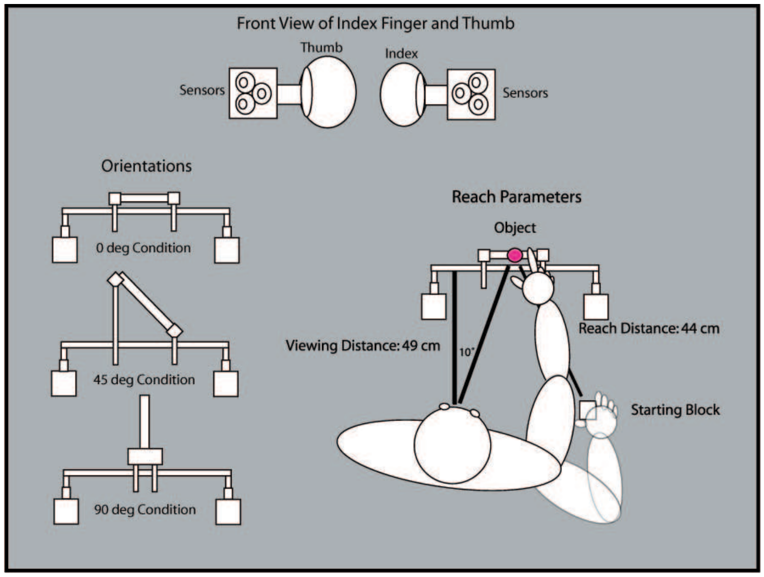

Figure 2: Reach set-up used to test the *natural* loss function. See methods for full description.

To test if people were reaching in a manner predicted by the natural loss function, subjects were required to touch and lift cylindrical objects at varying orientations (Figure 2: horizontal, 45 degrees, and vertical, with respect to gravity). When the fingers support an object, gravity induces torques on the object that varies with the object's orientation and finger locations. Notice the effects of gravity must be taken into consideration when lifting the object, but not when touching the cylinder. Therefore, we predict that subjects in the lift condition will place their fingers in locations that achieve the desired object motion while cancelling gravity-induced torques. Whereas, in the touch condition, subjects will successfully contact the experimentally imposed location.

The heart of our proposal is that the brain understands object manipulation and can bring object motion within its control framework. Hence, we will begin by explaining how object motion can be rewritten in terms of the motions of finger contacts and forces. We will then show how to rewrite the cost function for the object movement task in terms of contact conditions, which allows the selection of optimal contact conditions. Finally we compare predictions of this approach with contact data from experiments.

1 Controlling object motion with fingers

A fundamental idea in robotic manipulation is that a stably grasped object forms a closed kinematic linkage with the hand through frictional contact points (Montana, 1992; Bicchi & Kumar, 2000). What this means is that (with force closure) the finger contacts act as virtual links that attach the object to the hand. Roughly speaking, human fingertip contact (without sliding) is instantaneously equivalent to a (virtual) spherical joint with two degrees of freedom. Given a description of the object-finger contacts, the motion of the object is determined by the motions of the fingers, as long as contact is maintained.

The motion of a rigid object can be described using generalized coordinates and generalized momentum. Generalized coordinates $X(t) = [\vec{c}(t),\ \Omega(t)]^T$ combine a vector to the object's center of mass $\vec{c}(t)$ with a vector that parameterizes the three degrees of freedom in the rotation $\Omega(t)$ between the reference and object's coordinate frames. Specifically, the direction of $\Omega(t)$ is the rotation axis, and the length is the angle. Then the rotation is given by the matrix exponential $\mathbf{R}(t) = \exp(\Omega^{\times}(t))$, where $\Omega^{\times}(t)$ is the skew symmetric matrix that would accomplish a cross product by $\Omega(t)$ via a matrix multiply. Generalized momentum is similar, combining the linear momentum vector $\vec{P}(t) = m_o \frac{d\vec{c}}{dt}$ with the angular momentum $L(t) = I(\Omega(t))\omega(t)$ into one vector, where m_o is the object's mass, and $I(\Omega(t)) = \mathbf{R}(t)\, I_{body} \mathbf{R}^T(t)$ is the inertia tensor in the world frame and I_{body} is the fixed inertia tensor in the object's frame.

Using generalized coordinates and momenta, the equations for the dynamics of an object's motion are easy to write down. We consider the special case of object manipulation using two finger contact in stable force closure (without slippage). The dynamics of the object's motion as a function of the generalized finger contact forces $\vec{u}_i$ for two fingers is given by the Newton-Euler equation:

$$M(\Omega(t))\ddot{X}_o(t) = \vec{u}_1(\vec{r}_1) + \vec{u}_2(\vec{r}_2) + \vec{G}(\vec{r}_1, \vec{r}_2) \tag{1}$$

where $X_o(t) = [\vec{c}(t)\theta(t)]^T$ is a Cartesian vector representing the position and orientation of the object, $M(\Omega(t))$ is generalized mass matrix specified by the object mass and inertia tensors, $\vec{r}_i$ are the contact locations in object coordinates, and $\vec{G}$ the force and torque from gravity. These forces and torques need to be supplied by the fingertips while preserving finger-object contact.

Preserving object contact constrains both finger forces and finger motions. To maintain static frictional contact, finger forces $\vec{u}_i$ must have tangent plane components less than the static coefficient of friction μ times the surface normal component. Finger motions required to maintain contact can be written in terms of the object's motion, and the location and local geometry of contact points. Let S_o,$S_{oc,i}$, $S_{fc,i}$, and $S_{f,i}$ denote coordinate frames for the object (at the center of mass), the Gauss frame defined by the object's surface normal and tangent plane at the i^{th} contact point, finger i's Gauss frame at it's contact point, and finger i's reference frames, respectively. Then we can define a chain of transforms $T^{fc,i}_{f,i}$, $T^{oc,i}_{fc,i}$, $T^{o}_{oc,i}$ between the coordinate frames above, where subscripts denote the origin frame and superscripts the destination. Moreover, the velocity of the object can be written in terms of finger velocities using Jacobians derived from the transforms above and the kinematic contact constraints (see Montana (1995 for details):

$$\dot{X}_o(t) = \mathbf{J}_{f,i}\left(X_{f,i}(t), \vec{r}_i\right)\dot{X}_{f,i}(t) \tag{2}$$

Given non-sliding contact and force closure, these Jacobians are sufficient to determine the finger motions given an initial contact point. Thus the object motion can be rewritten in terms of desired finger motions and contact forces, which are straightforward to convert into joint torques given inverse dynamics models for finger motion. Abstractly, there are different system models for the fingers before: $\dot{X}_f = h_{free}(t, X_f, \vec{u})$, during: $\dot{X}_f = h_{contact}(t, X_f, \vec{u}, \vec{v}_c, \vec{r}_c)$ and after contact $\dot{X}_f = h_{f+obj}(t, X_f, \vec{u}, \vec{r}_c)$, which leads to different control laws $\vec{u} = \pi(t, X_f, \vec{v}_c, \vec{r}_c)$ applying in each period, where $\vec{u}$ is the vector of required finger forces. Note the dependence of the contact and object systems on contact location $\vec{r}_c$ and contact velocity $\vec{v}_c$.

2 Cost function for object motion planning

Our cost function combines two elements, task constraint costs penalizing deviations from desired object via states $Y^*_{o,j}(t_j)$ and finger control costs that are always present:

$$\begin{aligned} C_{obj} &= \sum_{j=1}^{Nvia} Q_j\left(X_o(t_j), Y^*_{o,j}\right) + \sum_{i=1}^{Nfing} \int_0^{t_{fin}} R_i(\vec{u}_i(t))dt \\ &= \sum_{i=1}^{Nfing}\sum_{j=1}^{Nvia} Q_j\left(T^o_f(t_j)X_{f,i}(t_j), Y^*_{o,j}\right) + \int_{t^+_{con}}^{t_{fin}} R_i(\pi^i_{f+obj}(X_{f,i}, \vec{r}_c))dt + \int_0^{t^-_{con}} R_i(\pi^i_{free}(t, X_{f,i}))dt \\ &\quad + \int_{t^-_{con}}^{t^+_{con}} R_i(\pi^i_{contact}(X_{f,i}, \vec{v}_c, \vec{r}_c))dt \end{aligned}$$

The optimal control strategy will minimize the expectation of this cost function over control laws $\pi(.)$. Because the control law after and during contact depends on contact conditions, the optimal overall control law will indirectly choose optimal contact conditions. In this formulation, optimal contact points could be selected online as a part of an expected cost-to-go minimization, or offline to serve as intermediate task goals.

To make testable predictions from this theory, we made the following simplifying assumptions: $Q_j\left(X_o(t_j), Y^*_{o,j}\right) = \alpha||X_o(t_j) - Y^*_{o,j}(t_j)||^2$, and $R_i(\vec{u}_i(t)) = \beta||\dot{u}_i(t)||^2/dt^2$, where α and β adjust the relative importance of the two costs. In addition, because the cost on controls is related to the finger jerk, we expect that the hand and object trajectories produced by an optimal control will be close to minimum jerk predictions on average. These simplifications were used to generate loss functions on contact location as described in the methods section.

3 Observed contact positions and velocities depend on the task

The theory makes some simple qualitative predictions. Contact locations that make it easy to move and stabilize the object should be selected. Moreover, the direction and magnitude of contact velocities should produce a smooth, low control transition between the approach and manipulation components of the reach. External forces that change the location where stabilizing forces should be applied can be used to test the contact location prediction. A simple way to accomplish this is to vary the orientation of an object with respect to gravity. As shown in figure 3, stably holding a cylinder requires cancelling torques due to gravity. Gravity induced torques depend on the distance between the cylinder's center of mass and the pivot point (L) and the orientation of the cylinder. The amount of torque that can be applied by the fingers is proportional to the distance between the fingers. Thus, planning for object motion would predict that subjects should adopt a staggered finger arrangement at contact when lifting the cylinder horizontally, parallel finger locations when lifting the object vertically, and intermediate finger offsets for orientations between the two extremes. A plausible alternative to planning for object motion is to choose a generic strategy which selects contact conditions that work in many situations. For lifting the cylinder in our tasks, such a strategy exists: reach for the center of mass with a parallel finger arrangement.

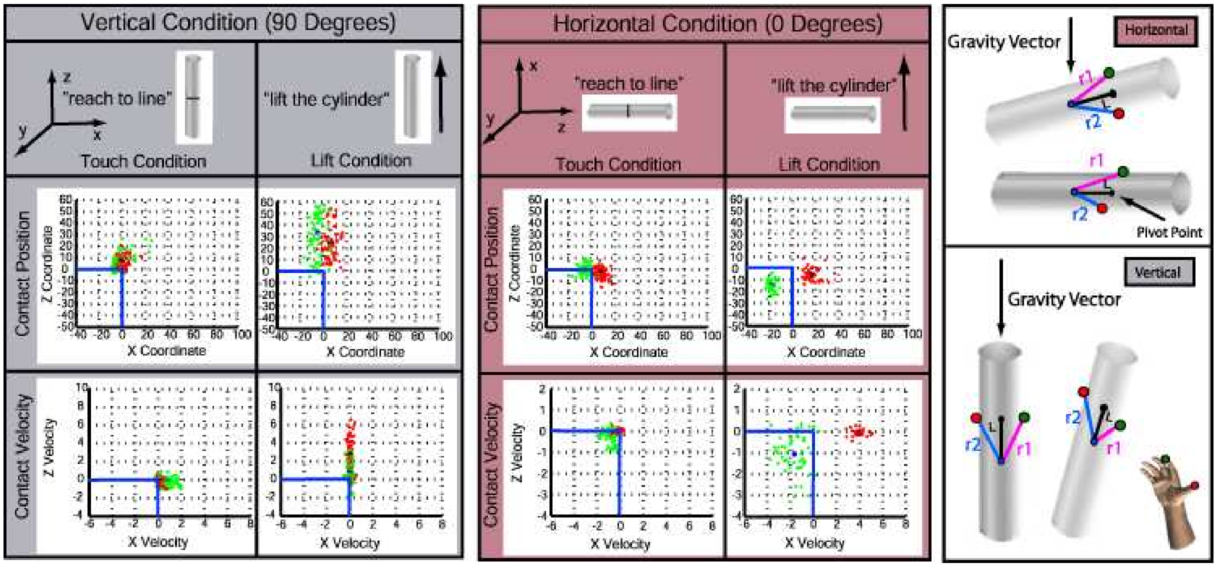

Figure 3: Empirical contact condition data for different tasks. End-point locations and velocities of the thumb (red dots) and index finger (green dots) are given. Blue lines on graph demonstrate the location of the center of mass $(0, 0, 0)$, and the zero velocity (mm/csec^2) location. **Left:** Contact location and velocity data for a vertically oriented cylinder. Notice in the touch condition people were reaching to the center of mass and terminating with zero velocity. However, in the lift condition, people were contacting the object above the center of mass and with positive z-velocity. **Middle:** Contact location and velocity data for a horizontally oriented cylinder. Notice in the touch condition people were reaching to the center of mass and terminating with zero velocity. However, in the lift condition, people were contacting the object with a staggered finger arrangement (thumb slightly to right of the center of mass, but finger more right) and with positive x-velocity. **Right:** Diagram illustrating the important concepts for the empirical predictions. r_1 & r_2 are the distance of the finger's and thumb's contact location from the center of mass, respectively. L is a point that bisects an imaginary line between the contact locations.

Assuming that people are planning for the object's motion also leads to contact velocity predictions. More specifically, it predicts that contact velocities will be selected that achieve the desired object motion for a given object orientation. To experimentally test these velocity predictions, we additionally required subjects to touch (i.e., no object motion) the same cylinders. It should be noted that there were two separate touch conditions where subjects were asked to touch imaginary lines and real lines at the object's center of mass for each orientation.

Figure 3 (middle) shows the results from a typical subject in the horizontal and vertical lift condition. It is apparent that the subject is following the object motion planning prediction: lifting with a staggered finger arrangement in the horizontal condition (where the thumb is slightly to the right of the center of mass, and the finger is further right of the center of mass), and lifting with a parallel configuration in the vertical condition. All subjects showed an identical trend, with the finger offset in the 45 deg condition falling between the 0 and 90 deg conditions. These results are not simply due to motor limitations–subjects were quite good at touching the imaginary line in all orientations. In fact, contact locations for the touch condition did not vary significantly across orientations.

In addition, contact velocities vary with object motion and orientation in ways qualitatively predictable from planning for object motion. Figure 3 shows that a typical subject in the touch condition contacted the object with zero velocity for both orientations in agreement with task demands. In the lift condition, it may not be desirable to contact the target with zero velocity. Instead, we can use the momentum of our hand to help carry the object to the desired location. Therefore, we predicted that subjects will contact the object with a positive thumb x-velocity for the horizontal condition, and a positive thumb and finger z-velocity for the vertical condition. All subjects showed finger velocities that match these predictions, and the horizontal and vertical data for one subject is shown in figure 3.

From this section it is clear that contact conditions change in the expected directions across tasks and orientations. The next section will outline evidence from the horizontal condition that suggests the empirical lift data are in agreement with the loss function detailed above.

4 Are empirical contact locations optimal?

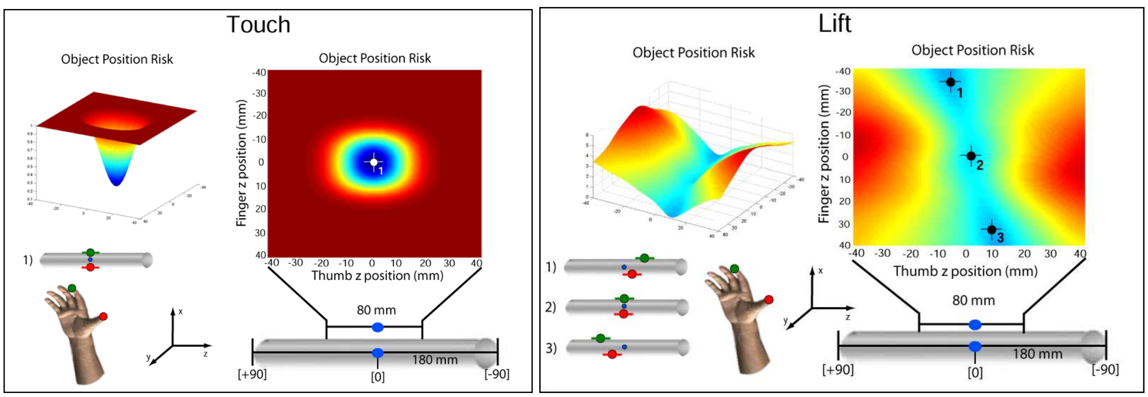

Figure 4: Risk landscapes for the horizontal touch and lift conditions. Index finger (green dots) z-positions are plotted against thumb (red dots) z-positions. The blue dot represents the object's center of mass. Only the middle 80 cm of the cylinder are shown. **Left:** *Touch risk landscape.* Heaviside function was used as loss function as it was assumed that the task required subjects to place any part of their finger and thumb (both assumed to be 1 cm) on the line. **Right:** *Lift risk landscape.* Loss function detailed above was used to develop landscape. Notice that there are a set of finger arrangements that produce accurate object motion.

Figure 4 illustrates the object position risk landscape for a horizontally oriented object as a function of the task. Blue areas show the low cost finger arrangements that overlap the contact line in the touch condition and produce good object motion in the lift condition. Figure 4 shows that there are essentially three distinct finger arrangements that produce good object motion. Conceptually, the top and bottom arrangements (Figure 4 (Right: 1,3)) are such that if the thumb falls to the right of the object's center of gravity, then the finger should be more right, and vice versa. Given our setup(i.e. right-handed reaching and start point right of center of mass), the (Figure 4 (Right: 1)) finger arrangements are more likely than those in (Figure 4 (Right: 3)). The middle arrangement (Figure 4 (Right: 2)) takes-on a more central approach where the fingers on parallel with each other. However, in order to achieve accurate object motion, it requires the fingers to be very near the center of mass. The area of low risk object motion for a central strategy is much smaller than those for a staggered strategy. Therefore, when making a right handed reach, the first finger arrangement may be the best of the possible low-risk arrangements. Finally, it should be noted that this risk function results solely from the physics of the task and motor noise estimates and has no free parameters.

Figure 5 shows the object position risk landscape with the four subject's data superimposed on the landscape. In the touch condition, it's apparent that subjects were doing as the task demanded - reach to the line located

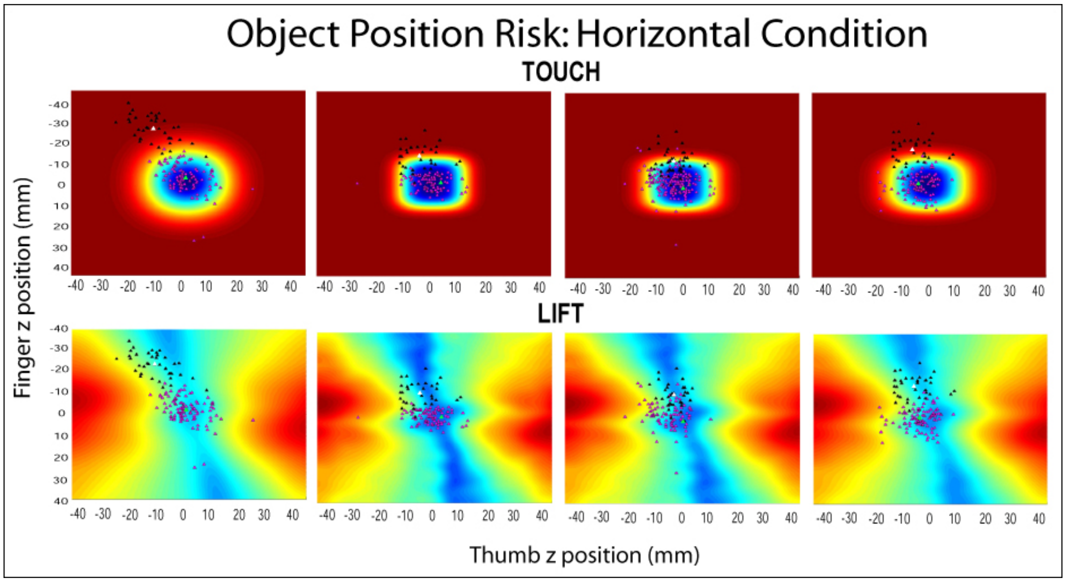

Figure 5: Risk landscapes plotted under data from four subjects. Purple triangles represent finger positions from an individual touch trial, and the green triangle represents the mean finger position across all of the touch trials. Black triangles are the finger locations from an individual lift trial, and the white triangle is the mean finger position across the lift trials. **Top:** Risk landscapes for touch data. It is apparent that the touch data falls within low risk areas for the touch landscape, but the lift data does not fall within the low risk areas for this landscape. **Bottom:** Risk landscape for the lift data. From the lift data, it seems that people are placing their fingers in positions that minimize the object position risk. This suggests that people are planning for the object's motion and placing their fingers in locations that produce accurate object motion.

at the center of gravity. There is little variability between subjects because there was only a small range of positions that satisfied the task (Figure 5 (Top)). Figure 5 (Top) also demonstrates that the lift data does not fall within areas of low risk for the touch condition. This is important because it is plausible that subjects are adhering to a generic control strategy that works in several different situations (e.g., reach to the center of mass with a parallel finger arrangement). However, this data suggests that people are not employing a generic control strategy.

In the lift condition (Figure 5 (Bottom)), finger arrangements varied slightly across subjects. However, they were not arbitrarily placing their fingers on the object as every subject contacted the object with a staggered finger arrangement located on an area with low object-motion risk. Notice that the areas of low object motion risk are extremely small - for the average finger position, there are $\leq 5mm$ of low cost thumb positions. Since the landscape puts a cost on the object's motion, it can be assumed that people were planning for the object's motion and placing their fingers in positions to achieve the desired motion, with minimum force and torque.

In agreement with our predictions, the results show that subjects systematically vary the contact location and velocities of their fingers as a function of the task and orientation. These results support the idea that contact points are selected to optimize a loss function on object motion. More generally, these results suggest that people understand the physics involved with lifting an object and that reach planning is hierarchical with object motion determining contact conditions and trajectory controls.

Methods

Experimental details Four head-fixed subjects were required to reach to a spatially fixed object located approximately 44 cm away. The viewing distance of the object was 49 cm. Subjects were instructed to reach as quickly (less than 1200 ms) and accurately as possible. Once the reach was completed, their hand was returned to the starting block for the next trial. Sensors were attached to the fingers and their movements were recorded via an Optotrak 3020 sampling at 100Hz. Each subject ran in one reach condition (i.e., touch or lift) at one object orientation (horizontal, 45 degrees or vertical) per day. For each condition the starting block was rotating to make the relative finger paths in space nearly identical for the three orientations. The session terminated after they completed 120 reaches per condition. Every subject ran in each of the six possible conditions. A calibration procedure was used to determine a probabilistic relationship between a fingertip model and the finger sensors and the position of the object relative to the Optotrak. Fingertips were modeled as elliptical generalized cylinders (15mm peak width and 10mm peak thickness) whose radius followed a profile roughly matched to PS's finger. Finger-object impact events were determined from the trajectory data the large magnitude jerk events with the highest probability of first collision (based on the calibration data). Contact positions were computed from the highest probability contact location at the end of the impact event, while contact velocities were referred to time immediately preceding impact.

Loss function computation To simplify risk function evaluation, we simulated the rigid body manipulation equations above over a dense grid of contact finger locations. Because the movements involved no intended rotation, an overall net force impulse could be precomputed that would drive the object along the desired minimum jerk trajectory to ($[0, 0, 50]$) in 500 ms. At each time step finger contact forces and finger movements of the above model finger were computed that would produce the value of the net force impulse while maintaining contact with least effort. Simulation parameters were: friction=1.1, hand mass of $.0474$ kg, object's mass 0.185 kg, and inertia was computed from the standard formula for a cylinder. Different measures of cost were then computed from the computed control signals and finger motions. Object motion cost is a sum of the deviation from desired object end state and the amount of relative motion between object and fingers introduced by rolling of the contacts. Object jerk and control cost measure the object jerk (linear and rotational) and the change in required controls, respectively. These cost measures differed very little qualitatively so only object cost is shown. Finally, the risk associated with a contact location was computed by convolving the evaluated cost function with the contact motor error distribution estimated from each subject's touch data.

References

Bicchi, A. and Kumar, V. (2000). Robotic Grasping and Contact: A Review. *In Proc. IEEE Int. Conf. on Robotics and Automation*, 1, pp. 348-353.

Flash, T., and Hogan, N. (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *The Journal of Neuroscience*, 5, pp.1688-1703.

Harris, C.M., and Wolpert, D.M. (1998). Signal dependent noise determines motor planning . *Nature*, 394, pp.780-784.

Montana, D. (1992). Contact Stability for Two-Fingered Grasps, *IEEE Transactions on Robotics and Automation*, 8, pp. 421-430.

Todorov, E., and Jordan, M. (2003). Optimality feedback control as a theory of motor coordination. *Nature Neuroscience*, 5, pp. 1226-1235.