Dynamic Primitives in Human Manipulation of Non-Rigid Objects

Hui Guang, Salah Bazzi, Dagmar Sternad, and Neville Hogan

Abstract—This study examined strategies humans chose to manipulate an object with complex (nonlinear, underactuated) dynamics, such as liquid sloshing in a cup of coffee. The problem was simplified to the well-known cart-and-pendulum system moving on a horizontal line. This model was implemented in a virtual environment and human subjects manipulated the object via a robotic manipulandum. The task was to maneuver the system from rest to arrive at a target position such that no residual oscillations of the pendulum bob remained. Our goal was to test whether humans simplified control by employing dynamic primitives, specifically submovements. Experimental velocity profiles of the human movements were compared to those predicted by three different control models. Two models used continuous optimization-based control, the third control model was based on Input Shaping. Input Shaping is a method for controlling flexible objects by convolving a motion profile with impulses of appropriate amplitude and timing. To evaluate whether humans used Input Shaping, we decomposed the velocity profiles recorded from humans into submovements, as proxies for the convolved impulses. Comparing the motion profiles from the 3 models with the experimentally measured human profiles showed superior performance of the Input Shaping model. These initial results are consistent with our hypothesis that combining dynamic primitives, submovements, is a competent description of human performance and may provide a simpler alternative to computationally complex optimization-based methods of robot control.

I. INTRODUCTION

Interactions with non-rigid objects, such as folding laundry or carrying a cup of coffee, are ubiquitous in everyday life. However, modern-day robots are still far from mastering these skills. This disparity in performance raises the question of how humans achieve their remarkable dexterity. Better understanding of human motor control might inform the control of robots, exoskeletons and prostheses.

Insights gained from human motor control have already helped inspire new ways to address problems in the control of robot motion. For example, Central Pattern Generators for locomotion in vertebrate biological systems have inspired rhythm generators for artificial locomotion [1], algorithms for grasp planning in robots have been inspired by how humans control their hand [2]. Recent approaches in humanoid robotics were informed by human postural control [3], and understanding human object manipulation from a controls perspective [4] are promising. Further, in physical human-robot interaction a recent study showed that following a robot trajectory required less effort when the robot trajectory had a velocity-curvature relation similar to that found in natural human movement [5].

Manipulating objects with internal degrees of freedom is a complex control task due to nonlinear and underactuated dynamics. However, while simple arm-reaching movements have been investigated intensively [6] [7] [8] [9], the question of how humans transport and manipulate non-rigid objects is far less understood [10] [11] [12].

Several studies provided evidence that humans use internal models that map hand movements and interaction forces to object movements. Both Mehta and Schaal [13] and Mah and Mussa-Ivaldi [14] used the task of controlling an inverted pendulum to demonstrate that humans employed internal forward and inverse models of motion-force relations. When moving a mass-spring system, Dingwell et al. [15] presented results that were consistent with the hypothesis that humans developed an internal inverse model of the arm-plus-object dynamics. However, the non-rigid object used in that experiment was a simple linear mass-spring system. For tasks that require physical interaction with objects that exhibit more complex nonlinear, underactuated and potentially even chaotic dynamics, a control strategy that relies solely on an internal model appears challenging due to its complexity.

A series of studies by Sternad and colleagues examined just such a complex object, a cart-and-pendulum, inspired by the task of carrying a cup of coffee, although simplified to 2D. Results showed that subjects made the dynamically complex system simpler to predict and exploited its resonance structure [16] [17].

Based on earlier results presented by Flash and Hogan [6] that showed that unconstrained reaching movements can be described by a trajectory that maximized smoothness (minimized mean-squared jerk), several studies investigated whether human manipulation of non-rigid objects admitted similar smoothness-based optimization criteria [18] [19] [20] [21] [22]. Note that in each of these studies, the non-rigid object of interest was a simple mass-spring system.

Dingwell et al. proposed minimum-crackle of the object (MCO) as the governing optimization criterion in this task [18]. However, this model has several limitations: In the limit as the spring stiffness went to infinity, i.e., the manipulated object became rigid, one would expect to recover the minimum-jerk model, but this was not the case. Furthermore, Svinin et al. pointed out that the MCO model was only applicable to the specific task of manipulating a single-mass, single-spring system [20]. Svinin et al. proposed another model that circumvented these problems and showed that experimental results could be predicted by a dynamically-constrained minimum-jerk trajectory for the hand (DCMJH).

A third model proposed in the literature minimized the acceleration of the center of mass (MACM) [21]. In ad-
dition to accounting for all previous experimental results on the mass-spring object, MACM was able to account for other non-typical behaviors. Moreover, in contrast to DCMJH, MACM did not require re-optimization when the task changed from unconstrained reaching to object manipulation. Both the controlled variable, the center of mass of the system (hand-object), and the optimization criterion did not change between tasks and, hence, the analytical expressions representing the center-of-mass trajectory remained unchanged.

All three models suffer from several shortcomings. They all assumed that the dynamics of the manipulated object were those of a linear mass-spring system. This linear system does not represent the considerably more complex dynamics of the broad range of objects that humans interact with on a daily basis. Moreover, these models were merely descriptive and did not provide insight about how they would be implemented or generated by a controller. Online solution of such optimization problems combined with the inverse dynamics computation to generate the planned/desired trajectories would be a daunting task for human brains.

Therefore, we hypothesized that humans simplify the control of physical interaction with dynamically complex objects by using dynamic primitives. There is mounting evidence that the human sensorimotor control system indeed relies on a composition of primitives [23]–[30], but this has not been investigated in the context of complex object manipulation and controlling physical interaction. Dynamic primitives include submovements for discrete actions, oscillations for rhythmic actions, and mechanical impedance for physical interaction with objects [31] [32]. For robots, control by composition of dynamic primitives may provide a simpler alternative to computationally intensive approaches based on optimization [33] [34] [35].

To test this hypothesis, we designed an experiment in which participants transported a simplified model of a cup of coffee in a virtual environment. They were instructed to maneuver the cup from a starting position to arrive at a target position with zero oscillations of the ball. This ‘no-residual-oscillations’ task afforded a simple solution known as Input Shaping, which has been studied in the control literature [36] [37] [38]. Developed to control the motion of highly oscillatory systems, Input Shaping relies on convolving a motion profile with a sequence of impulses that cancels out transient oscillations induced by movement. The simulated movement profile is strikingly similar to the human profiles covered in the movement. However, since the hand and object were dynamically coupled, the required boundary conditions on the hand movement could be satisfied by imposing corresponding boundary conditions on the third and fourth derivatives of the object trajectory. This resulted in 10 independent boundary conditions

\[ \begin{align*}
  x_H(0) &= \dot{x}_H(0) = \ddot{x}_H(0) = 0, \\
  x_H(T) &= D, \dot{x}_H(T) = \ddot{x}_H(T) = 0, \\
  x_O(0) &= \dot{x}_O(0) = \ddot{x}_O(0) = 0, \\
  x_O(T) &= D, \dot{x}_O(T) = \ddot{x}_O(T) = 0 = 0.
\end{align*} \]

The lowest-order polynomial satisfying these boundary conditions was 9th-order. The optimization criterion that resulted in a ninth-order polynomial was to minimize the mean-squared-crackle, \( C \), of the object trajectory

\[ C = \frac{1}{2} \int_0^T \left( \frac{d^5(x_O)}{dt^5} \right)^2 dt. \]  

Enforcing the boundary conditions in (3), the optimal object trajectory was

\[ \begin{align*}
  x_O(t) &= 70D \left( \frac{t}{T} \right)^9 - 315D \left( \frac{t}{T} \right)^8 + 540D \left( \frac{t}{T} \right)^7 \\
  &\quad - 420D \left( \frac{t}{T} \right)^6 + 126D \left( \frac{t}{T} \right)^5.
\end{align*} \]

II. THREE MODELS FOR NON-RIGID OBJECT CONTROL

A. Minimum-crackle of the object (MCO)

The MCO model was proposed by Dingwell et al. [18], who investigated the task of transporting a mass with an attached spring from one point to another in the horizontal plane. The dynamics of the object were governed by

\[ \ddot{x}_O = \frac{K_O}{M_O} (x_H - x_O), \]  

where \( x_H \) and \( x_O \) denoted the position of the hand and the object, respectively; \( K_O \) was the stiffness of the object, and \( M_O \) was its mass. To move both hand and object from rest to rest required satisfying 12 boundary conditions, 6 for the hand and 6 for the object

\[ \begin{align*}
  x_H(0) &= \dot{x}_H(0) = \ddot{x}_H(0) = 0, \\
  x_H(T) &= D, \dot{x}_H(T) = \ddot{x}_H(T) = 0, \\
  x_O(0) &= \dot{x}_O(0) = \ddot{x}_O(0) = 0, \\
  x_O(T) &= D, \dot{x}_O(T) = \ddot{x}_O(T) = 0 = 0.
\end{align*} \]

B. Dynamically-constrained minimum-jerk of the hand (DCMJH)

The DCMJH model [20] proposed that, similar to unconstrained reaching movements, humans try to minimize the jerk of the hand trajectory, \( J \), when transporting an object from rest to rest

\[ J = \frac{1}{2} \int_0^T \left( \frac{d^3(x_H)}{dt^3} \right)^2 dt. \]

The key point is that \( x_H \) needs to be expressed in terms of \( x_O \). From (1),

\[ x_H = \frac{M_O}{K_O} \ddot{x}_O + x_O, \]  

Taking the third derivative of (7) for hand-jerk, the problem became minimization of object crackle, as in the MCO model, combined with object jerk. Due to its length and
complexity, the analytical expression for the optimal object trajectory is not included here, but the interested reader is referred to [20].

C. Input Shaping (IS)

Input Shaping (IS) is a strategy for maneuvering objects with highly-oscillatory internal dynamics (e.g. a payload suspended from a crane) in a rest-to-rest motion, such that there are no residual oscillations at the target position [36] [37] [38]. Input Shaping is achieved by convolving a motion profile with a sequence of impulses that excite transient oscillations in the system in a manner that results in their destructive interference at the target position. An example with 2 impulses is shown in Fig.1(a).

Consider a crane controlled by a human. The operator could simply move the trolley to the desired position using a desired velocity profile e.g. a minimum-jerk profile. However, without any other input, this would cause the payload to oscillate around the new trolley position, as displayed in Fig.1(b). On the other hand, it is possible for the trolley to reach the desired position with zero residual oscillations of the payload, as shown in Fig.1(c). This can be achieved by convolving the minimum-jerk velocity profile with a sequence of impulses that cancel residual oscillation (an input shaper), as shown in Fig.1(a). The resulting ‘shaped’ command is then used as input to the system to generate the behavior illustrated in Fig.1(c).

For the cart-and-pendulum system moving from rest to rest over a distance \( D \), an input shaper consisting of only 2 impulses suffices to ensure zero residual oscillation [39]. For this input shaper, the two impulses \( I_1 \) and \( I_2 \) and their respective timings \( t_1 \) and \( t_2 \) were

\[
I_1 = \frac{\exp \left( \frac{-\zeta \pi}{\sqrt{1-\zeta^2}} \right)}{1 + \exp \left( \frac{-\zeta \pi}{\sqrt{1-\zeta^2}} \right)}, \quad I_2 = 1 - I_1, \quad t_1 = 0, \quad t_2 = \frac{\pi}{\omega_d},
\]

where \( \zeta \) was the system’s damping ratio and \( \omega_d \) was its damped frequency.

III. SIMULATIONS AND CONTROL MODEL PREDICTIONS

A. Dynamical model of the “cup of coffee”

From a computational perspective, simulating a realistic 3-dimensional cup with sloshing coffee governed by nonlinear equations from fluid mechanics is taxing. This computational complexity was especially challenging as the study employed a virtual setup in which real-time simulations of the system dynamics were required. The cup of coffee was therefore simplified to a semicircular 2-dimensional arc, the cup, with a ball rolling inside, representing the coffee; the motion of the 2D cup was limited to a horizontal line. Restricting the ball to only slide along the cup without rolling or friction, the system was reduced to the well-known cart-and-pendulum system with no damping. Fig.2(a) illustrates the real and simplified task, while Fig.2(b) displays the mechanical model of the simplified system.

The equations of motion of the simplified mechanical model were

\[
(m + M)\ddot{x}_C = m l \left( \dot{\phi}^2 \sin(\phi) - \dot{\phi} \cos(\phi) \right) + u,
\]

\[
l\ddot{\phi} = -g \sin(\phi) - \ddot{x}_C \cos(\phi),
\]

where \( x_C \) denoted the position of the cart, \( \phi \) denoted the pendulum angle, counter-clockwise defined as positive, \( m \) the mass of the pendulum bob, \( M \) the mass of the cart, \( l \) the length of the massless pendulum rod, and \( g \) the gravitational acceleration. The force exerted by the human subject was denoted by \( u \). The parameters used to simulate this system in the virtual environment were: \( M = 1.9 \ kg, \ m = 1.1 \ kg, \ l = 0.5 \ m, \ and \ g = 9.8 \ m/s^2 \).
B. Linearization of the cart-and-pendulum model

To simulate the three models, the equations of motion (9) and (10) were linearized around the pendulum’s downward equilibrium position, \( \phi = 0 \). The reasons for this were two-fold: First, the MCO and DCMJH models were developed for linear mass-spring models; a fair comparison should be based on a similarly linearized model. Second, the standard application of Input Shaping is to a linearized model. While designing an input shaper for a nonlinear system is possible, it is not straightforward and was beyond the scope of this study.

The linearized equations were

\[
(m + M)\ddot{x}_C = -ml\ddot{\phi} + u, \quad (11)
\]

\[
l\ddot{\phi} = -g\phi - \ddot{x}_C. \quad (12)
\]

The position of the internal degree of freedom, the pendulum position \( x_P \), which was required to simulate the MCO and DCMJH models, can be inferred from Fig.2(b)

\[
x_P = x_C + l\sin(\phi). \quad (13)
\]

Using the linearized version of (13)

\[
x_P = x_C + l\phi, \quad (14)
\]

\[
\ddot{x}_P = \ddot{x}_C + l\ddot{\phi}, \quad (15)
\]

and substituting into (12), the following relation was obtained

\[
\ddot{x}_P = \frac{g}{l} (x_C - x_P), \quad (16)
\]

which closely resembled (1). For simulating the MCO and DCMJH control models, the position of the hand \( x_H \) was assumed to coincide with the cart position \( x_C \). This linearization was only used to compute predictions of the three models; the virtual system with which subjects interacted simulated the full nonlinear equations (9-10).

C. Predictions of the three control models

Simulations of all 3 control models generated velocity profiles of the linearized system that succeeded in the ‘no-residual-sloshing’ task. The desired distance \( D \) was equal to the distance between the starting and target positions in the experiment, \( 0.45 \ m \); all other model parameters were identical to those used in the virtual environment (see next section). Since the two optimization-based models (MCO and DCMJH) were variants of maximizing smoothness, the input shaper model used a maximally smooth (mean-squared jerk) profile. That was convolved with the impulse sequence.

Fig.3 presents the simulation results of all three models for 12 motion durations between 1.2 s to 2.6 s (color-coded and normalized to time). Even though the 3 control models produced very similar motion profiles, there were also important differences. In particular, the Input Shaping model predicted that cup velocity would never be negative. Nevertheless, the similarity shows that the simple Input Shaping strategy compares well to the computationally expensive optimization-based methods.

IV. The ‘NO-RESIDUAL-SLOSHING’ Task

The task of transporting a cup filled with coffee requires physical interaction with a dynamically complex non-rigid object. Motivated by this real-life example, in this experimental task subjects maneuvered an underactuated object from a starting point to a target point, with the goal of arriving at the target position with no residual oscillation.

A. The virtual experiment

The dynamical model presented in (9) and (10) was simulated in a virtual environment interfaced with a robotic manipulandum, as depicted in Fig.2(c). The projection screen displayed the cup as a 2D arc (corresponding to the cart) and ball (corresponding to the pendulum bob), and a start and target box separated by 0.45 m, as shown in Fig.2(d). Subjects were instructed to move the cup from the start box on the left and arrive at the target box on the right with no residual oscillations of the ball. They were also instructed not to lose the ball along the way. Subjects moved at their desired pace; only if the trial duration was longer than 3.5 s subjects were told to move faster on the next trial. At the end of each trial, the maximum angle that the ball reached inside the target box was displayed on the screen as performance feedback to the subjects. The smaller the angle of these terminal oscillations, the better their performance.

The experiment consisted of 4 blocks with 50 trials each. Before the first block, subjects familiarized themselves with the virtual task for 5 minutes. At the beginning of each trial, the cup was centered in the left box and the ball rested at its equilibrium position.
B. Apparatus and data acquisition

The participants stood in front of a large backprojection screen (2.4 × 2.4 m) at 2 m distance. Physical interaction with the virtual object occurred via a 3-degree-of-freedom force-controlled robotic manipulandum (HapticMaster, Motekforce, NL [40]). By applying a force to the handle of the robotic arm, participants controlled the horizontal x-position of the virtual cup. The robotic arm was restricted to move in the horizontal direction along the subject’s frontal plane to ensure a linear horizontal motion of the cup, consistent with the model. The robotic arm provided haptic feedback, allowing participants to sense the system’s inertia and the force of the ball on the cup. More details on the manipulandum’s end-effector position resolution, haptic resolution, and force sensitivity, are provided in [40].

The force applied by the participants to the manipulandum (u) and the kinematics of the cup and the ball (x_c, \dot{x_c}, \ddot{x_c}, \phi, \dot{\phi}, \ddot{\phi}) were recorded at 120 Hz. A total of 10 subjects participated in the experiment. The protocol was approved by the Institutional Review Board of Northeastern University and participants signed a consent form prior to the experiment.

C. Submovement analysis

All three models assumed perfect execution of the motor task: zero residual oscillations. For example, the Input Shaping model required two perfectly timed and precisely identical impulses (Fig.3(c)), observed as two symmetrical peaks in the velocity profile. To evaluate whether humans used Input Shaping, we decomposed the velocity profiles recorded from humans into submovements, regarded as proxies for the convolved impulses. Unlike the exact model, the analysis allowed two different submovements, each approximated by a support-bounded lognormal function. These differences accounted for the inevitable variability of human performance. Therefore, we decomposed the measured velocity-profiles into a sequence of support-bounded lognormal functions, commonly referred to as LGNB [41]. This choice was also motivated in previous work showing that LGNB submovements provided the best fit to rapid movements [42] [43]. Using the scattershot algorithm presented in [43], the sequence of LGNB functions that produced the best fit with the experimental velocity profile for a given trial was extracted. An example of a typical decomposition is shown in Fig.4.

V. Experimental results

Subjects chose a range of movement times to perform the task. The histogram in Fig.5 pools the movement times of all trials, separating the successful from the unsuccessful ones. A trial was considered successful if the maximum angle in the residual oscillations was ≤ 3 degrees. These results show that the majority of the trials lay within the range of 2.25 s to 3.25 s, and that the proportion of successful trials increased for slower movements.

A. Comparison of the three control models

To determine which of the three models, MCO, DCMJH, or IS, best described the experimental data, the velocity profiles of all successful experimental trials were compared to the model-predicted profiles. These comparisons were made for 8 different movement times between 1.7 s and 2.5 s. Trials with movement times in the intervals of [1.7, 1.8], [1.8, 1.9], ..., [2.4, 2.5] seconds were averaged and standard deviations were calculated. For each interval, the velocity profiles for the three models at this duration were plotted together with the mean velocity profile of all successful trials. The results of these comparisons for 4 exemplary intervals are displayed in the top row of Fig.6. Although the three control models predicted quantitatively different velocity profiles, these differences were comparable considering the variability of the experimental data. The rightmost panel in the top row summarizes the fitting error for all 3 models for the 9 different motion durations.

An Input Shaping strategy based on a minimum-jerk assumes an exact and perfect execution of motor actions. A more realistic approach to allow for the variability of human performance assumed a pair of non-identical submovements, specifically LGNBs. We therefore used the extracted LGNB profiles to approximate Input Shaping. The velocity profiles generated with these LGNB submovements are displayed in the bottom row of Fig.6. As summarized in the rightmost panel in the bottom of Fig.6, the two LGNB submovements

3787
show a better fit with the experimental data than the two optimization models.

B. Sensitivity analysis

Imperfectly-timed or wrong-amplitude submovements would result in non-zero residual oscillations, quantified by their maximum angle. To further assess the appropriateness of this model, we tested the sensitivity of Input Shaping to errors in amplitude and timing of the impulse sequence was analyzed, Fig.7(a). For each trial, errors of impulse amplitudes and timing were inferred from the extracted submovements. The resulting maximum angle of residual oscillation was determined as summarized in Fig.7(a) and compared with the experimental performance in Fig.7(b). As illustrated, the predicted residual oscillations were close to the experimentally observed oscillations. This sensitivity analysis supports that humans used a strategy similar to Input Shaping.

VI. DISCUSSION AND CONCLUSIONS

This paper addressed the problem of manipulating a non-rigid object, such as a cup of coffee, with internal degrees of freedom. Motivated by previous work on Input Shaping, the hypothesis was that humans simplify control using appropriately scaled and timed submovements. To test this hypothesis the task of maneuvering a cart-and-pendulum system was implemented in a virtual environment. Experimental velocity profiles of successful trials were compared to those predicted by three different control models, two based on proposed optimization criteria, the third based on Input Shaping.

The difference between the three model predictions was largely obscured by the variability of the experimental data. When variability was taken into account by decomposing experimental velocity profiles into a sequence of overlapping LGNB submovements, that sequence best described subjects’ behavior.

A. Limitations

The most prominent limitation of this initial study was that subjects avoided fast movements. For faster movements, the proportion of trials that failed became larger. This initial study has not yet analyzed the reasons and errors in the respective strategies leading to failure. Further, this initial experiment comprised a limited number of trials and was not designed to study learning. It is possible that with extensive practice subjects may converge to a behavior that more closely resembled one of the optimization-based models. That is a topic for future research.

B. Levels of explanation

An important distinction between the models considered here is the level at which they describe motor behavior. At least three levels with progressively increasing detail (observational, compositional, physiological) have been proposed [31]. The two optimization-based models describe actions at the level of observable behavior, but remain silent about how that behavior might be implemented at lower levels. The model based on submovements describes the compositional level where primitive actions are composed to produce observable behavior. The fact that Input Shaping and its experimental approximation successfully reproduced human performance suggests that it may provide deeper insight about human control of complex dynamic objects. This insight suggests that primitives-based controllers are simpler solutions for robot control compared to computationally complex optimization-based methods.