Predicting Slippage and Learning Manipulation Affordances through Gaussian Process Regression

Francisco E. Viña B., Yasemin Bekiroglu, Christian Smith, Yiannis Karayiannidis, Danica Kragic

Abstract — Object grasping is commonly followed by some form of object manipulation — either when using the grasped object as a tool or actively changing its position in the hand through in-hand manipulation to afford further interaction. In this process, slippage may occur due to inappropriate contact forces, various types of noise and/or due to the unexpected interaction or collision with the environment.

In this paper, we study the problem of identifying continuous bounds on the forces and torques that can be applied on a grasped object before slippage occurs. We model the problem as kinesthetic rather than cutaneous learning given that the measurements originate from a wrist mounted force-torque sensor. Given the continuous output, this regression problem is solved using a Gaussian Process approach.

We demonstrate a dual armed humanoid robot that can autonomously learn force and torque bounds and use these to execute actions on objects such as sliding and pushing. We show that the model can be used not only for the detection of maximum allowable forces and torques but also for potentially identifying what types of tasks, denoted as manipulation affordances, a specific grasp configuration allows. The latter can then be used to either avoid specific motions or as a simple step of achieving in-hand manipulation of objects through interaction with the environment.

I. INTRODUCTION

Interaction with and manipulation of objects are essential abilities of robots operating in realistic environments. As humans, robots may need to grasp objects for simple tasks such as moving them from one position to another. More complex tasks, such as using objects as tools, requires a more advanced ability of manipulating an object in the hand so to achieve a suitable grasp configuration. In this process of achieving and loosing contacts with the object in the hand, events such as slippage commonly occur. The knowledge of contacts and slippage provides important information about the status of the task one is executing.

For both humans and robots, sense of touch is paramount for safe and flexible interaction with objects and the environment. As reviewed in [1], components of tactile perception in humans depend on the sensory inputs within muscles, tendons and joints (kinesthetic) and stimulus mediated by receptors in the skin (cutaneous). Most of the research in robotic tactile sensing addressed the problem of finger-object interactions and grasp stability assessment. If the contact locations as well as the friction coefficients of the contacting surfaces are known, the problem can be formulated in terms of the Grasp Wrench Space (GWS) [2], [3]. However, it is difficult to construct the GWS in practice since it requires the exact values of those parameters.

Besides planning stable grasps, the robot should also acquire knowledge of the maximum forces and torques that can be applied on the object before slippage occurs. Various methods have been proposed for detecting slippage [1], [4]–[6]. Apart from addressing the problem at the signal processing level in terms of cutaneous tactile sensing, general machine learning methods have proven adequate for analysis in cases where noise and imperfect models are inherent to the problem, [7], [8].

Our work follows the direction of using kinesthetic sensing for slip detection in combination with machine learning techniques. Autonomous learning and a physical model of the friction forces are used to estimate the maximum static friction forces and torques on objects the robot is interacting with. We approach the problem through Gaussian Process regression, resulting in a model that can predict forces and torques that a grasp can tolerate before the held object starts slipping. As such, the model can also be used to identify the affordances of a specific grasp such as, for example, what type of in-hand rotation can be applied to an object while still keeping the object in the hand.

The learned bounds can be used as constraints at the control level to avoid certain motions and thus prevent slippage of the grasped object while executing the task. In addition, the approach also identifies in which directions the object might translate or rotate in the hand and thus be exploited in tool use and in-hand manipulation to actively change the pose of the object in the hand – either through specific motion or interaction with the environment. This is also commonly done by humans, for example prior to putting a key in a keyhole we may change its orientation between the fingers by pushing the key toward a surface.

Fig. 1 : A dual arm robot setup for estimating maximal allowable forces and torques for a grasp.
Thus, differently from commonly addressed grasp affordances [9], we facilitate the system to identify manipulation affordances. Our method uses force-torque and proprioceptive feedback different from commonly used tactile or skin sensors which in practice can be fragile and easily damaged. However, when possible, the cutaneous and kinesthetic methods can be integrated resulting in a more biologically inspired approach [1]. Our approach also takes advantage of the dual arm capabilities of humanoid robots since the training actions can be executed autonomously through dual arm manipulation procedures. Fig. 1 shows our dual-arm robot as an example of a platform that can be used to implement the method we propose in this paper.

The paper is organized as follows: Section II presents the related work, Section III our learning framework, including the friction model and the use of Gaussian Process regression while in Section IV we proceed to describe how our system learns manipulation affordances from doing regression on the static friction. Finally, we provide our experimental results in Section V as well as the conclusions, discussion on the results and future directions in Section VI.

II. RELATED WORK

Early works studying the physics of robotic grasping and contact between rigid bodies are reviewed in [3]. The review addressed the basic closure properties of grasps, force and form closure, which describe the equilibrium conditions of an object grasped by a robotic hand by assuming frictional and frictionless point contacts respectively. Given that friction forces play a central role in robotic grasping, some of the works reported in the literature have focused on studying their properties [5], [10]. These studies cover not only the translational Coulomb friction, but also the rotational friction. Moreover, by combining different sensor modalities (tactile and force-torque) it is shown in [5] that it is possible to detect and control both translational and rotational slippage.

Besides modeling the physics of grasping and the friction forces, quantifying the quality of grasps in terms of the capability to counteract external disturbances has been one of the main research questions in the grasping community. In order to plan stable grasps with robotic hands, many grasp planners have been proposed in the literature which optimize these quality measures [2], [11], [12]. These planners are constructed in terms of approximations of wrench spaces or heuristic algorithms that consider a subset of a wrench space.

The main drawback of these methods is that these require precise 3D models of the object as well as prior knowledge of the friction coefficient and the location of the contact points of the robot’s hand. To cope with this problem, [13] proposes a set of manipulation actions to estimate properties such as weight, stiffness and friction in order to set appropriate grasping forces.

In order to overcome the uncertainties and problems with modeling errors in grasping, learning approaches have also been proposed. Example works of [7], [8], [14] consider learning of grasp stability and grasp affordances. Our previous work on grasp stability assessment performs learning mainly through tactile (cutaneous), proprioceptive and visual feedback in order to predict the stability of the grasp prior to lifting and manipulating the object [8], [14]. In [7] the proposed system learns grasp affordances which are defined as hand-object relative poses that lead to successful grasps on a particular object. These affordance densities are learned through exploration and visual features. The main strength of these learning approaches originates from the fact that these do not require prior knowledge of physical contact parameters as the system is trained using supervised learning without explicitly modeling the physics of grasping.

Our work makes use of the physics models of friction described in the seminal work of [5], [10]. However, instead of employing geometrical, analytical or signal processing based approaches [2], [4], [5], [11], [12] we follow a kinesthetic learning approach for predicting slippage. In this sense, our work follows more closely approaches in which the robot first interacts with objects and assesses their contact and friction properties prior to executing tasks [13]. Our method also follows the motivation behind learning based approaches in order to deal with the issue of modeling errors and uncertainties in grasping [7], [8], [14].

Within the broader scope of haptic sensing, which consists of both cutaneous and kinesthetic sensing as shown in Fig. 2, our approach falls under the subcategory of kinesthetic sensing and perception while most of the related work discussed so far including our own work on grasp stability assessment cover mostly the domain of cutaneous/tactile sensing [4], [6], [8], [14].

III. PHYSICS AND LEARNING MODEL

The main objective of our system is learning the maximum static friction forces and torques for various grasp configurations through force-torque sensing. In this section we present
the modeling aspects of our framework, beginning with a description of the friction model used and the selection of input features for training. We finalize the section with a brief overview of Gaussian Process regression and explain how we apply it within our work.

A. Friction Model

According to the Coulomb friction model, when an external force is applied parallel to the surface of contact between two bodies, there is a reaction friction force \( f_j \) which relates to the normal force \( f_n \) according to the following inequality

\[
f_j \leq \mu_s f_n
\]

(1)

where \( \mu_s \) is the static coefficient of friction. This equation holds until the external force exceeds the maximum static friction force. The object then starts slipping when Eq. (1) becomes an equality. From this point, a dynamic friction force with a lower friction coefficient starts acting on the object as depicted in Fig. 3. The peak of this curve corresponds to the maximum static friction force \( f_{\text{slip}} \) given by

\[
f_{\text{slip}} = \mu_s f_n
\]

(2)

The static torsional friction typically displays a nonlinear behavior given by

\[
\tau_{\text{slip}} = \beta_s f_{n}^{4/3}
\]

(3)

where \( \beta_s \) depends on geometric and elasticity factors of the contact [5]. However, slippage still occurs at the point in which the friction torque reaches its maximum value, which we denote as \( \tau_{\text{slip}} \).

In order to achieve a more general physical model for prediction, we take into consideration the effect of both rotational and translational friction forces as discussed in [5], [16]. When an object is subject to both rotational and translational shears, the translational and rotational friction components become correlated as shown in Fig. 4. The curve \( f_{t} = h(\tau_n) \), where \( f_t \) is the component of the force tangent to the contacting surfaces and \( \tau_n \) the component of the torque in the normal direction, represents the boundary at which the object starts slipping due to the loads exerted on the object. If the tangential force \( f_t \) applied on the object is above the curve for a given applied torque \( \tau_n \), then the object will slip and the grasp is thus unstable.

A number of mathematical approximations have been formulated in the literature to describe this slippage boundary. We will use the linear approximation described in [5] that defines a conservative bound on the magnitude of the forces and torques that cause slippage on an object. This linear bound is denoted by \( f_{t}(\tau_n) = h_{\text{lin}}(\tau_n) \) in Fig. 4 and can be expressed using the following equation:

\[
\frac{f_{t}}{\mu_s} + \frac{\tau_n}{\beta_n} = f_n
\]

(4)

B. Learning Framework

Our goal is to learn the mapping between a set of input features \( (X) \) and the resulting maximum friction forces and torques \( (Y) \), which is a regression problem due to the continuous outputs. While there are several types of regression techniques that could be used within our framework, we have chosen Gaussian Process (GP) regression which can capture the nonlinearity in the data and provide estimates for uncertainty in the predictions.

1) Gaussian Processes: Given a dataset \( D = \{x_i, y_i\}_{i=1}^{n} \) with \( n \) observations where \( x_i \in \mathbb{R}^N \) and \( y_i \in \mathbb{R} \) is a scalar output, regression analysis aims at learning a model for the relationship \( y = f(x) + \varepsilon \) which is composed of a latent function of the input and a noise component \( \varepsilon \). As a result of this learning, given a new input \( x' \), the aim is to obtain the predictive distribution for \( y' \).

A GP [17] defines a distribution over functions and is parametrized by a mean and a covariance function as

\[
GP \sim (m(x), k(x, x'))
\]

(5)

The mean function is assumed to be zero. The covariance function expresses how similar two outputs, \( f(x_i) \) and \( f(x_j) \), are for some input pairs \( x_i, x_j \).
are given the inputs $x_i$ and $x_j$. Our covariance function is based on the squared exponential, which is given by

$$k(x_i, x_j) = \sigma_f^2 \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right) + \sigma_n^2 \delta(x_i, x_j).$$

(6)

The hyperparameters of the covariance function, $(\sigma_f, \sigma_n, l)$, are optimized based on $D$, where $\sigma_f$ denotes the signal variance, $\sigma_n$ is for the noise variance and $l$ is the length-scale which determines how relevant an input is, i.e., if $l$ has a large value the covariance will be independent of that input.

We are interested in the conditional probability $p(y^*|D, x^*)$ as we want to find how likely is a certain prediction for $y^*$, given the data and the new input. Based on a trained GP model, the estimate for $y^*$ is given by the mean value at the test point with the confidence being the variance. The interested reader can refer to the literature [17] for additional details on Gaussian Processes.

2) Feature Selection: As an input to the regressor, we need a set of informative features $X$, that can reliably represent the behavior of the maximum static friction forces and torques. In our case, we have selected the $x$ component of the hand $H$ pose with respect to the object $O$ as shown in Fig. 5

$$X = \begin{bmatrix} O & x_H \end{bmatrix}$$

(7)

We have selected this feature for illustration purposes, yet more features can easily be incorporated into the system, such as for example the joint angles of the fingers and their grasping force which can modify the friction forces present in a grasp. If more features are incorporated into the system, a preprocessing stage with dimensionality reduction would be necessary [18].

![Fig. 5](image.png)

Fig. 5: Grasp preshape used for training on the maximum static friction forces and torques, with the corresponding reference frames of the hand and the object used for training.

The outputs $Y$ of the regression system are the maximum static friction force and torque

$$Y = \begin{bmatrix} f_{slip} \\ \tau_{slip} \end{bmatrix}$$

(8)

which can be measured through force-torque sensors by interacting with the object. We isolate the components of $Y$ and train two GPs, one for the translational friction $f_{slip}$ and one for the rotational friction $\tau_{slip}$. In our case, we learn friction forces $f_{slip}$ in the $y_H - z_H$ plane and friction torques $\tau_{slip}$ around the $x_H$ axis of the tip of the hand reference frame as shown in Fig. 5, given that these are the directions in which the object can move within the hand. Forces and torques around the remaining axes are trivial to learn since they will be constrained by the operational safety limits of the hand, given the geometry of the grasp.

IV. TOWARDS LEARNING MANIPULATION AFFORDANCES

Once the robot has interacted with an object and learned the maximum friction forces $Y = [f_{slip}, \tau_{slip}]^T$ for a range of grasp configurations, it can use this information to infer what type of motions the object can withstand given the current grasp. The details of the training data generation for learning are provided in the next section.

For a given wrench $w^*$ measured by the robot while executing a task, the robot can detect how close the object is to slipping according to the model discussed in Section III-A. In order for the object to remain fixed in the robot’s hand the measured force should lie below the torque dependent slippage boundary $h(\tau)$

$$f_t^* < h(\tau_n^*)$$

(9)

where $f_t^*$ and $\tau_n^*$ are the tangential force and normal torque components of the wrench measured by the robot.

In the training stage we isolate the translational and rotational components of the friction and thus we can approximate $h(\tau_n)$ linearly with $h_{lin}(\tau_n)$ by joining the end points $(f_t, \tau_n) = (f_{slip}, 0)$ and $(f_t, \tau_n) = (0, \tau_{slip})$. In the case of a linear approximation the following condition ensures a stable grasp in terms of zero relative motion between the object and the hand $H \times_O = 0$:

$$f_t^* < h_{lin}(\tau_n^*)$$

(10)

$$f_t^* < -\frac{f_{slip}}{\tau_{slip}} \tau_n^* + f_{slip}$$

Thus, our approach makes it possible to identify stable grasps through identification of forces and torques that can be applied on an object before slippage occurs. In a broader sense, the methodology also identifies directions of motion constraints – that is, in which directions the object is more likely to translate or rotate.

In the case of the grasp studied in this work, see Fig. 5, the model would inform that the object can translate in the $y_H - z_H$ plane and rotate around the $x_H$ axis. Moreover, if a large torque is detected around the $x_H$ axis with relatively low forces in the $y_H - z_H$ plane then we can expect the object to rotate around the fingertips rather than translate once the force-torque measurements reach the slippage boundary of Eq. (4).

This knowledge is necessary for manipulation tasks where a predicted slippage of the object may be facilitated to complete a task. An example scenario is shown in Fig. 6, in which the robot exploits the rotational slippage to pour the contents of the cereal box into the bowl by letting the box rest against an edge of the bowl and allowing it to rotate slightly in the hand while the manipulator moves upwards.
V. EXPERIMENTAL EVALUATION

Our experimental setup consists of a dual arm robot as shown in Fig. 1. Each manipulator has 7 DOF and these are equipped with ATI Mini45 6-DOF force/torque sensors mounted at the wrists and they are sampled at a 650 Hz frequency. We start by describing the training data collection process.

A. Training Data Collection

For collecting training data autonomously with the robot we use three dual arm manipulation procedures: one sliding action for measuring the maximum static linear friction $f_{slip}$ and the other two are a rotational motion and pushing action for measuring the rotational friction $\tau_{slip}$.

Fig. 7 shows an illustration of the sliding action along with the forces and torques measured during the execution. In this case the robot holds the object firmly with the parallel gripper shown on the right while the hand on the left, which is the one we train for, slides up in the $y_H$ direction of the hand. The $y$-component of the force signal $f_y$ measured in the force-torque sensor of the arm is then similar to the one shown in Fig. 3, and $f_{slip}$ is obtained from the peak of the signal.

For obtaining training data for the maximum static friction torque $\tau_{slip}$, we used the pushing action shown in Fig. 8. This action is performed by grasping the object with the hand we train for, while the parallel gripper shown on the right pushes the object on a corner so that the object rotates around the $x_H$ axis of the tip of the robotic hand. We selected this action given that we expect collisions with the environment to be a source of rotational slippage when the robot performs tasks with the object.

For verification purposes we also trained a separate GP for $\tau_{slip}$ by applying a different type of training action as shown in Fig. 9. This training action consists of performing a rotational motion with the grasping hand while the object is kept on a fixed grasp with the parallel gripper shown on the right. Even though in this case we also train for $\tau_{slip}$ as with the pushing action, we can expect different outcomes from the learning given that each training action represents a different kind of interaction with the environment. The pushing action gives $\tau_{slip}$ for tasks in which the object is grasped by the robot’s hand and it collides with the environment while being grasped by the robot hand, whereas the rotational motion models a task in which the object is fixed with respect to the environment and the robot’s hand rotates around the object.

B. Experimental results

We collected 14 training examples for the friction force and 10 training examples for the torque by varying the relative pose between the robot hand and the manipulated object along one dimension as described in Section V-A. To learn the Gaussian Processes and obtain the hyperparameters we
used Rasmussen and Nickisch’s Gaussian Process Regression and Classification Toolbox [17]. The hyperparameters were calculated by maximizing a Gaussian likelihood function.

Fig. 10 shows the resulting learned Gaussian Process for $f_{\text{slip}}$. This plot shows the mean function of the learned GP (solid blue line) which follows the training points, along with the two standard deviation confidence bounds (dashed red lines) enveloping it. Given this result, we take the lower confidence bound as stability boundary for $f_{\text{slip}}$ given that the Gaussian Process predicts that 95% of the points of the process will lie above this boundary.

For testing and validating the learned GP, we manually pushed the object while it was being grasped by the robot in different configurations compared to the ones used for training. Fig. 10 confirms that the sliding action performed on the object is valid for training $f_{\text{slip}}$ as most of the test points lie above the lower confidence bound of the Gaussian Process.

Fig. 11 shows the learned Gaussian Process for $\tau_{\text{slip}}$ trained by using the pushing action shown in Fig. 8.

Fig. 11 shows the learned Gaussian Process for $\tau_{\text{slip}}$ when using the pushing action. Once again, we manually pushed the object while it was grasped by the robot in order to collect the test points shown in the figure. These test points show that the pushing action and the learned Gaussian Process succeeded in capturing the behavior of $\tau_{\text{slip}}$ with respect to the object to hand relative pose.

Fig. 12 shows the result of learning $\tau_{\text{slip}}$ by using the rotational motion, while we collected test points by manually pushing the object as in the previous case. The clear offset between the learned GP and the test points shows that the training and testing actions are not anymore physically consistent. In the case of the rotational training motion, the interaction between the active robot hand and the object involves both forces and torques, while pushing actions, performed either by the robot hand or manually by ourselves for testing, exert only forces on the object. This result can thus be used to inform the system that the action is not proceeding according to the model and provide the basis for replanning. This is something we plan to address in the subsequent work.
Fig. 12: Learned GP of $\tau_{\text{slip}}$ with two-standard deviation confidence bounds trained with the rotational motion shown in Fig. 9.

VI. CONCLUSIONS AND FUTURE WORK

In this work we have presented a learning framework for prediction of slippage of grasps through kinesthetic perception which provides a basis for learning manipulation affordances. Our method uses Gaussian Process regression and the training is performed by isolating the translational and rotational components of the friction. The novelty of the approach lies on using a machine learning approach together with a physical model of the friction to determine continuous bounds on the forces and torques that a grasped object can withstand before slipping for a set of different object-hand relative poses. The experimental results show that our system is able to generate reliable predictions which agree with tests performed by manually pushing the object in the hand of the robot for previously unencountered grasp configurations.

Future directions of work include expanding our sensor modalities from kinesthetic perception to cover a wider spectrum of haptic perception (see Fig. 2) by use of tactile sensing. We also aim to incorporate into our system the estimation of the axis of rotation of the object in the hand of the robot as it can improve the results shown here. We have assumed a constant axis of rotation around the fingertips of the hand that might not correspond precisely with the actual axis around which the object rotates when it is manipulated. In order to cope with this issue, we aim to use adaptive control techniques previously used for estimating the kinematic constraints of hinged doors [19] and treat the object as a virtual hinge. We are also interested in coupling this work with probabilistic grasp assessment techniques and object categorization as demonstrated in our previous work in [20], [21].

ACKNOWLEDGMENT

This work has been supported by the European Union FP7 project RoboHow.Cog (FP7-ICT-288533), the Swedish Research Council (VR) and Swedish Foundation for Strategic Research (SSF). The authors gratefully acknowledge the support.

REFERENCES