Guiding Grasping with Proprioception and Markov Models

Peter Deckers, Aaron M. Dollar, and Robert D. Howe

Abstract— **This paper describes the application of a partially observed Markov decision process (POMDP) to guide the control decisions made during the task of grasping objects with a simple compliant grasper in unstructured environments. The decision process relies only on the sensing of angular deflection of the compliant gripper joints – proprioceptive information available on most robot hands and grippers. This information is used to infer the state of contact between the gripper and the object and guide a set of actions to be undertaken in order to lead to a successful grasp. We believe that the performance of the gripper under a POMDP model built from this limited sensory information will serve as a valuable baseline for comparison with more complex sensing modalities, allowing for quantitative analysis of the tradeoffs between commonly available sensory suites.**

I. Introduction

The uncertainty associated with interacting with an unstructured environment presents a number of challenges. For grasping, the lack of a precise model of the object, environment, and contact state with the grasper makes the task of reliably acquiring the target object difficult. Indeed, the fidelity of the available sensory information can vary widely. We are interested in the performance of a gripper when only proprioceptive information such as joint angles is available. Since the vast majority of robot hands incorporate this type of sensing, we would like to evaluate the performance of a gripper in a scenario in which only this basic set of information is available. This can serve as a baseline for future studies to address the cost-benefit tradeoffs of adding further sensory systems such as contact and force that can be used for the control of a hand during the grasping task.

This raises the question of how to best use the very limited sensory information available from finger joints, especially in an unstructured environment where visual and a priori information is prone to error so there is great uncertainly about object properties. The framework of partially-observable Markov decision processes (POMDPs) provides a formal means of dealing with the uncertainty inherent in these tasks and enabling robust control of the robot manipulator [1]. Utilizing POMDPs is a way of estimating the probabilities of the different states between the grasper and object and allows intelligent decisions to be made regarding a sequence of actions to be undertaken in order to converge to the target configuration. The POMDP framework has recently been applied to grasping using contact sensors on the tips and sides of the fingers [2]. In this paper we examine the use of the POMDP framework to guide manipulation by sensing joint deflections of a compliant gripper.

We begin this paper with a general description of POMDPs, paying particular attention to the physical interpretation of the various mathematical components. We then describe the model of the gripper that we are interested in analyzing, and identify a set of states and actions that we believe embody the most important aspects of the task we are interested in analyzing, which is a generalization of grasping tasks with frequently used robot hand architectures. Finally, we derive the equations governing the observation model and suggest practical limits on the unknown state variables which will enable the derivation of the probability density function.

II. Model Construction

A. Partially observed Markov decision processes

A partially-observed Markov decision process (POMDP) is a model for formalizing the decision process under uncertainty in the classification of system state, in order to choose an appropriate action [1,2]. The POMDP model is constructed and used in the following manner:

- Define a set of *states* S based upon the specifics of the task and equipment.
- Define a set of *actions* A which will be undertaken based on the prediction of the state and which will lead towards some goal.
- For every state s in S and action a in A, define a *reward* R(s,a) which will determine which action to take based upon the state prediction.
- For every combination of state s_t and actions a_t, define a *transition matrix* Q(a_t) which represents the probabilities of transitioning from state s_t to s_{t+1}: $Q(a_t)_{i,j} = \Pr(s_{t+1}=j|s_t=i,a=a_t)$. This matrix is called the Markov transition matrix.
- Choose a set of *observations* O which consist of the

P. Deckers was supported by a Fellowship of the Belgian American Educational Foundation and a Fulbright Fellowship from the United States Department of State. He is with the Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138 USA (e-mail: peterdeckers@post.harvard.edu).

A. M. Dollar is with the Harvard/MIT Division of Health Sciences and Technology and the Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (phone: 617-253-2941; fax: 617-253-8542; e-mail: adollar@mit.edu).

R. D. Howe is with the Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138 USA (e-mail: howe@deas.harvard.edu).

available sensory information.

- Generate an observation model $P(o|s)$ which will be used to predict the state based on the specific observation of sensory information: For every observation, define the diagonal matrix $B(o)$ which is a diagonal matrix with diagonal elements (i,i) the probability of observing o , given the state is i: $B(o)_{i,i} = Pr(o|s=i)$.

If the state probabilities $p_{t+1}(i)=Pr(s_{t+1}=i|o_1 \ldots o_{t+1}, a_1, \ldots a_t)$ are represented as a row vector p_{t+1}, we can now represent the state-transition model as $p_{t+1}=1/N.p_t.Q(a_t).B(o_{t+1})$ where N is chosen such that $\sum_i p_{t+1}(i)=1$.

In this way we can now iteratively choose the appropriate action a_{t+1} based on our guess of the guess of the state probabilities at time t+1, p_{t+1}. The action a_{t+1} we choose at time t+1 maximizes $\sum_i p_{t+1}(i).R(s_{t+1}=i, a_{t+1})$.

In the context of our grasping problem, states are contact conditions (e.g. fingertip–to-object, finger side-to-object, etc.), actions are motions of the base of the hand or robot, and the observations are the sensed joint angles of the compliant gripper.

B. Grasper Model

Our goal in this study is to gain insight into how common compliant robot fingers can be used as "feelers" to determine contact state with little or no visual information, and use this information to guide the manipulator into a successful grasping configuration with the object. The basic approach is to use kinematic information provided by joint angle sensing and the known kinematics of the fingers to infer object location and geometry [3,4] and carry out a set out actions to lead to a target configuration in which a successful grasp can be achieved. The use compliant finger joints also enable the inference of some aspects of contact forces from this kinematic information.

To reduce the parameter space of the problem, we focus on a simple planar gripper with two fingers, each with two compliant revolute joints (Fig. 1). This gripper, proposed by Hirose [5], is perhaps the simplest configuration that is able to grasp a wide range of objects. This mechanism is the same as that used in the 100G hand [6] and is similar to the planar, power-grasp configurations of a number of popular robot hands.

In previous work, we examined the optimization of the preshape, joint stiffness, joint coupling, and actuation of this mechanism [7,8]. Additionally, we constructed a four-fingered gripper similar to this model and experimentally demonstrated that the compliance and adaptability designed into the mechanism (based upon the results of the optimization studies [7,8]) was able to reliably grasp a wide range of target objects in the presence of sensing uncertainties resulting in larger positioning errors [9]. Furthermore, the hardware was designed to be simple to use (feed-forward control, a single actuator for eight degrees of freedom) and robust to impacts and other large forces that are likely to occur in unstructured grasping tasks.

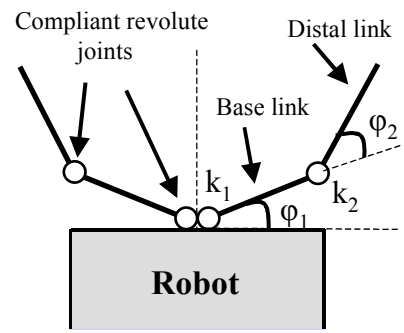


Fig. 1. Gripper model used in this study

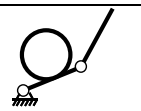
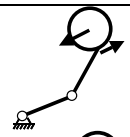
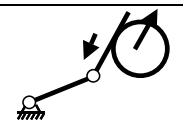
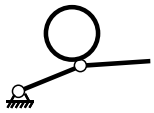
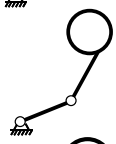
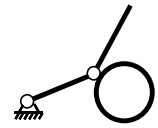
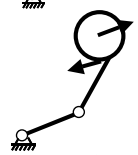
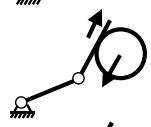
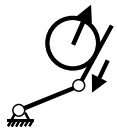
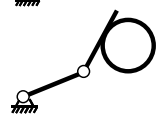
For the above gripper model, we refer to the two links closest to the robot as 'base links' each having length l_1. The other two links are referred to as 'distal links' and have length l_2. The base links are connected to the robot via compliant revolute joints, each having rest angle φ_1 and joint stiffness k_1. The base and distal links are also connected by compliant revolute joints each having rest angle and stiffness φ_2 and k_2, respectively. Because of the symmetry of the setup, only the finger on the right is further considered in the analysis.

C. States

The first task in finding a representative POMDP for the above grasper is defining the finite set of states S. Since we are using this model to make decisions with regards to movements of the gripper, we consider as separate states a quasi-static representation of the relative motion within contact states (Table I). Thus, for example, states 3, 4, and 5 all refer to contact between an object and the inside surface of the distal link, with each of these states distinguished by different relative motion between the link and object. Distinguishing these different motions is necessary as each produces different frictional forces and thus different deflections in the compliant joints. This is also the reason that there is only one state for contact with each surface of the proximal link (states 1, 2, 12, and 13); in these states any frictional forces would act on a line through the base joint and thus produce no joint torque. The use of a circular object in Table I is arbitrary – the states presented encompass all object geometries.

Note that typical use of the finger as a "feeler" would make many of the states listed unlikely to occur in practice (e.g. states 11-13), since the robot would likely approach the target object in the direction of the hand opening. However, for completeness, we present them all in Table 1. Also note that state 2 can only occur for objects having a sharp protrusion or corner at which contact with the distal joint of the grasper occurs. For an object with a finite contact point radius (e.g. the circular object shown in the diagram), this state is impossible as it would require infinite contact force on link 2. Along these lines, we believe these states apply to all objects in single point contact with the finger.

TABLE I
DISCRETE STATES FOR THE POMDP OF A 4 LINK GRASPER. THE CIRCLE REPRESENTS A GENERIC OBJECT

State	*Description*		*State*	*Description*		*State*	*Description*	
1	Link1 In		**6**	Tip 2 Slip Out		**11**	Link 2 Out Slip Down	
2	Tip 1 In		**7**	Tip 2 Stick		**12**	Tip 1 Out	
3	Link 2 In Slip Up		**8**	Tip 2 Slip In		**13**	Link 1 Out	
4	Link 2 In Roll		**9**	Link 2 Out Slip Up		**14**	No Contact	
5	Link 2 In Slip Down		**10**	Link 2 Out Roll				

On the tip of link 2, 3 different states are possible: slip out, slip in and stick. On both sides along link 2 the possible states are slip up, slip down and "roll". This last term is a general term meant to encompass rolling (both up and down) and sticking, which can only happen for corner type contacts and is impossible for a circular object except for contact on the tip of the finger. These behaviors are lumped together as a single state since further sub-classification is not possible based upon the available information. Also, if point contact occurs on link 1 and link 2 (which is common in the transition between those two individual states for a circular object), we classify the configuration as 'Link 1 Inside' (state 1) since contact on link 1 kinematically determines the grasper state. The 'no contact' state (state 14) could also be subdivided according to where the center of the object is with respect to the grasper.

D. *Actions, Observations and Reward*

One of the key elements that determine the utility of a POMDP is defining an effective set of actions. Although we don't go very deep into this subject here, we suggest a possible set of actions for this specific grasping setting:

- 'Move the base a fixed distance to the right'
- 'Move the base a fixed distance to the left';
- 'Move the base a fixed distance up'
- 'Move the base a fixed distance down'
- 'Move base in the direction of $\theta_1 + \theta_2$ down until $\theta_2 = \varphi_2$'
- 'Move base in the direction of $\theta_1 + \theta_2$ up until $\theta_2 = \varphi_2$ '

The final two actions are chosen to allow the gripper to "trace" the object along the inside or outside of link 2 to position the hand without exerting large forces on the object.

TABLE II
NOMENCLATURE

Parameter	Definition
s_t	State at time t
a_t	Action at time t
o_t	Observation at time t
p_t	State probability vector at time t
$Q(a_t)$	Markov transition matrix for the action a_t
$B(o_t)$	Observation matrix for observation o_t
l_1, l_2	Link lengths
φ_1, φ_2	Spring rest angles
θ_1, θ_2	Deflected angles
k_1, k_2	Joint stiffness values
F_n, F_t	Normal and tangential force
l_{cont}	Contact length
μ	Friction coefficient

As mentioned in the introduction, we are interested in using the proprioceptive joint angle values, θ_1 and θ_2, as our observations. We believe that this case is a baseline for sensory information since information about the kinematic configuration is available on the vast majority of robot hands and grippers. Additional sensory information may be useful, but we would eventually like to perform a quantitative

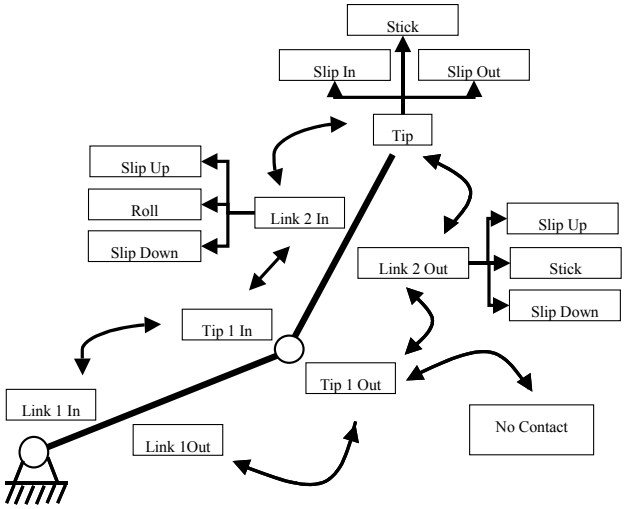


Fig. 2. Representation of possible transitions in the transition model.

analysis of the tradeoffs between sensory suites, and this case in which we restrict sensory information to only joint angles will serve as a baseline for this comparison.

A reward needs to be assigned for every state/action combination (for our proposed model, 14 states x 6 actions = 84 rewards). We will not present this matrix here, but these rewards will be chosen based upon intuition as to whether the given state/action combination will lead towards the goal of the object being in a graspable location with respect to the hand. For example, when the state is 'Link 2 Inside Roll' (state 4) the reward of the action 'Move base to the right' should be more positive than 'Move the base to the left'.

E. The Transition Model

The transition model represents the probability of transitioning from state s_t to s_{t+1}, given an action a_t. For every action in our action set and for every combination of (s_t,s_{t+1}) we thus must come up with a probability of transitioning from state s_t to s_{t+1}. These probabilities can of course only be estimated. The most important aspect of defining these matrices $Q(a_t)$ is to avoid eliminating possible transitions (i.e. by setting their probability equal to 0).

Figure 2 provides a visualization of some possible transitions between different states. Note that for all states, one possible transition is to 'no contact' (state 14). As an example transition model for the state 'Link 2 outside roll' (state 10), if the action a_t is 'move a fixed distance to the left' one might assign probabilities {0.25; 0.25; 0.2; 0.1; 0.1; 0.1} to the states {10; 14; 9; 6; 7; 8}, respectively.

III. The Observation Model

For the Observation Model, we need to calculate the probability of the observation, P(o|s), for every state s and for every observation. Since the observations are continuous and 2 dimensional $o=(\theta_1,\theta_2)$, P(o|s) is a 3 dimensional probability density function. Intuitively we can say that for example if the state s = 'Link 2 Inside Slip Up' (state 3), a configuration with $\theta_1 > \varphi_1$ and $\theta_2 > \varphi_2$ must have probability 0. Also, we can see

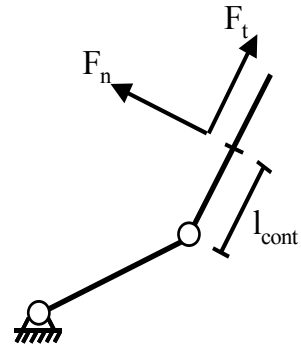


Fig. 3. Force and length convention

that for this state a configuration with $\theta_2 = 0$ yields an infinite force component in the direction of the distal link, which is impossible.

To systematically approach this problem we first derive the equations that give the relationship between the force acting on the distal link and the measured observation (θ_1,θ_2). These equations will give us insights in what we can infer from knowledge of the joint angles θ_1 and θ_2. Finally a method is described for how to calculate P(o|s) for every state *s*.

A. Governing Equations

To derive the relationship between the joint angles, joint stiffnesses, and the contact forces on the finger, the following convention is used (Fig. 3): The tangential force F_t is taken positive in the direction of link 2, out. The normal force F_n is taken positive to the inside. The length where the object makes contact is l_{cont}. If contact is on the tip of the outer link, $l_{cont} = a_2$. A simple force-torque balance gives us the relationship between the angular deflections, the contact length and the force components (forward and inverse):

$$\begin{cases} \theta_1 = \varphi_1 + \dfrac{1}{k_1}\left(F_t l_1 \sin\left(\varphi_2 + \dfrac{F_n l_{cont}}{k_2}\right) \right. \\ \qquad \left. + k_2 \dfrac{F_n l_{cont}}{k_2} + F_n l_1 \cos\left(\varphi_2 + \dfrac{F_n l_{cont}}{k_2}\right)\right) \\ \theta_2 = \varphi_2 + \dfrac{F_n l_{cont}}{k_2} \end{cases}$$

$$\begin{cases} F_n = \dfrac{k_2(\theta_2 - \varphi_2)}{l_{cont}} \\ F_t = \dfrac{k_1(\theta_1 - \varphi_1) - k_2(\theta_2 - \varphi_2) - F_N l_1 \cos(\theta_2)}{l_1 \sin(\theta_2)} \end{cases}$$

Note that the nomenclature used in this paper is summarized in Table II. Additionally, note that these equations are valid only for contact on link 2, as tangential force on link 1 cannot be known based on the available information.

Each of these two equations are in terms of three unknowns F_n , F_t and l_{cont}. Therefore, after measuring θ_1 and θ_2, some further information relating to the state must be assumed in order to solve for the three remaining state variables F_n , F_t

and l_{cont}. For example, if we know there is contact on the tip of link 2 ($l_{cont} = a_2$), we can infer both components of the force from measuring the angular deflections. Also, if we know for example that the system is in state 11, 'Link 2 Outside Slip Down', then we know the relation between the force components $F_t=\mu F_n$ and can make inferences about the three state variables by assuming some bounds on the coefficient of friction, μ, based upon knowledge of our finger coverings, for instance.

B. Derivation of the Observation model

To generate the observation model, we must find the probability of observing $o=(\theta_1,\theta_2)$ for each of the fourteen states at some predetermined sampling density and range of the observations θ_1,and θ_2. For example, suppose we are analyzing state 11, 'Link 2 Outside Slip Down'. What can we then infer for the probability of $o=(\theta_1,\theta_2)$? To make judgments regarding this probability, we can assume some reasonable bounds on our unknowns. For example:

- $\mu_{\min} \le \frac{F_t}{F_n} \le \mu_{\max}$
- $F_n \ge 0$ (for most objects and finger coverings)
- $\left\|\vec{F}\right\| < F_{\max}$
- $0 \le l_{cont} \le a_2$

These limits can be set based on some *a priori* knowledge of the specific task. For instance, if you know you are attempting to grasp an empty glass, one might put F_{max}=10N, above which the glass will almost certainly be pushed away and will no longer be in contact with the gripper. Additionally, a reasonable limit on the coefficient of friction could be assumed (e.g. $0 \le \mu \le 3$).

With these constraints specified, one can now construct the probability density function P(o|s), for *s* = 'Link 2 Outside Slip Down' as follows: For every combination of θ_1 and θ_2, calculate the contact lengths, l_{cont} , that satisfy all the above constraints by using the formulae that relate the angular deflections to the forces. For example, one might find that for $l_1<l_{cont}<l_2$ all the above constraints are satisfied for that specific combination of θ_1 and θ_2. The probability at (θ_1 , θ_2) might then be chosen to be proportional to the ratio of the possible link locations to the total link length, or $(l_2-l_1)/l_2$. The proportionality constant is determined such that it normalizes the 3D probability distribution.

For the case of state 9, 'Link 2 Outside Slip Up', we can use the previous derivation after changing the bound for the friction factor to $-\mu_{max} \le F_t/F_n \le -\mu_{min}$ and the bound for the normal force $F_n \le 0$. When the state is state 10, 'Link 2 Outside Roll', the friction factor bound becomes $-\mu_{max} \le F_t/F_n \le \mu_{max}$.

The derivation of the probability functions for the states on the inside of link 2 are similar, with the only difference that here we can improve our estimation of the probability function if we would know something about the local object geometry (e.g. a circle with r >0 cannot touch link 2 with l_{cont} = 0 without touching link 1). In these instances, we can change $0 \le l_{cont} \le l_2$ to $l_{low} \le l_{cont} \le l_2$, with l_{low} determined by the lower bound of the local geometry.

If the given state is contact on the tip of link 2, we cannot 'scan' the contact length anymore because this is now given ($l_{cont}=l_2$). Since θ_1, θ_2 and l_{cont} are known, the two components of the force at the tip are known. However, the direction of the normal force relative to the object is now unknown (the normal of the local object geometry at the contact point is unknown). Instead of 'scanning' the contact length we can now do the same thing, but with the direction of the normal force with respect to the object For the given states on link 1, similar principles can be used to derive P(o|s).

IV. Conclusions and Future Work

In this paper we formalized the problem of determining the contact state of a simple compliant gripper with an unknown target object in the framework of a partially-observable Markov decision process (POMDP) utilizing only information about the kinematics of the gripper (i.e. joint angles). Specifically, we proposed a set of states, a set of actions and an observation model that can be used in setting up this POMDP.

A methodology was developed to construct the probability distribution functions that belong to the different states. These probability distribution functions reflect the probability of observing a certain pair of angles, without assuming anything about the environment. This observation model has the property that it gets more accurate if more about the environment is known (e.g. bounds on the friction coefficient or maximal force).

An immediate way of improving the performance of this POMDP is to divide the links in sub-regions and increase the set of states accordingly. In this way, the observation model can 'eliminate' some states more effectively (e.g. by giving the regions at the end of the distal link higher probabilities than the regions closer to tip 1). In combination with a respective change in the reward function this can substantially improve the performance. For example, one might want to perform a different set of actions when there is contact at the end of the distal link than when contact is closer to the joint.

Subsequent future work includes the implementation of the suggested POMDP. Doing this will require specifying a Markov transition matrix for every action as well as a reward for every combination of action and state. Additionally, we will need to construct a simulation to model the physical interaction between the grasper and object, including the grasping process.

Ultimately we would like to use the POMDP framework to evaluate the tradeoffs between different sensory suites available to the grasper. For instance, how much better would performance be if contact sensor information was available in addition to joint angle information? Does the inclusion of

fragile sensors such as force transducers warrant the added expense and unreliability? Furthermore, the POMDP framework could be extended to grasp stability analysis. How reliably can "stability state" be determined under various observation sets (i.e. sensory information)? These types of questions might be rigorously undressed under the POMDP framework to provide valuable information to both designers and end-users of robot hands.

REFERENCES

[1] R.D. Smallwood and E.J. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Operations Research,* vol. 21(5), pp.1071–1088, 1973.

[2] K. Hsiao, L.P. Kaelbling, and T. Lozano-Perez, "Grasping POMDPs," proceedings of the 2007 International Conference on Robotics and Automation, pp. 4685-4692, 2007.

[3] P. Dupont, T. Schulteis, P. Millman, and R. D. Howe, "Automatic Identification of Environment Haptic Properties," *Presence*, vol. 8(4), pp. 392-409, 1999.

[4] S. Haidacher and G. Hirzinger, "Contact Point Identification in Multi-Fingered Grasps Exploiting Kinematic Constraints," proceedings of the 2002 IEEE International Conference on Robotics and Automation, pp. 1597-1603, 2002.

[5] S. Hirose and Y. Umetani, "The Development of Soft Gripper for the Versatile Robot Hand," Mechanism and Machine Theory, vol. 13, pp. 351-359, 1978.

[6] M. Higashimori, M. Kaneko, A. Namiki, M. Ishikawa, "Design of the 100G Capturing Robot Based on Dynamic Preshaping," The International Journal of Robotics Research, vol. 24 (9), pp. 743-753, 2005.

[7] A. M. Dollar and R. D. Howe, "Towards grasping in unstructured environments: Grasper compliance and configuration optimization," Advanced Robotics, vol. 19 (5), pp. 523-544, 2005.

[8] A.M. Dollar and R.D. Howe, "Joint Coupling Design of Underactuated Grippers," proceedings of the ASME 30th Annual Mechanisms and Robotics Conference, 2006 International Design Engineering Technical Conferences (IDETC), Philadelphia, PA, Sept. 10-13, 2006.

[9] A.M. Dollar and R.D. Howe, "Simple, Robust Autonomous Grasping in Unstructured Environments," proceedings of the 2007 IEEE International Conference on Robotics and Automation (ICRA), Rome, Italy, April 10-14, 2007.