Learning Grasp Phases from Human Demonstrations

Oliver Kroemer, Gerhard Neumann, Herke van Hoof, and Jan Peters

I. Introduction

Manipulation tasks can usually be decomposed into multiple discrete phases [3], [7]. For example, picking up an object generally involves three phases: reaching, loading and lifting. In the reaching phase, the hand is moving in free space to the object. In the loading phase, the hand has made contact with the object and it is applying forces to the object. In the lifting phase, the hand has lifted the object from the table and can freely move it around. The transitions between phases often occur when there is a change in the contacts, e.g., when the fingers make contact with the object, or the object breaks contact with the table.

A defining characteristic of a phase is the effects of actions. For example, flexing the fingers will cause the fingers to move when reaching and apply more force to the object in the loading phase. Therefore, in order to apply the desired manipulation to an object, the agent must first transition to a suitable phase. As a result, the conditions for transitioning between phases represent subgoals of the manipulation task.

In this paper, we present our ongoing work on learning the phases of manipulation tasks from human demonstrations. In our previous work [4], we focused on learning phases from purely exploratory actions, by observing the effects of actions. However, humans also select which action to perform based on the current phase [3]. Therefore, by observing human demonstrations of the manipulation task, additional information regarding phase transitions can be obtained. We present a probabilistic model for representing the demonstrations, wherein the phases are represented as hidden variables. The experimental setup for collecting grasping demonstrations is described in Section II. The probabilistic model is explained in Section III.

II. Grasping Demonstrations

Experiments were performed using the Darias robot shown in Fig. 1. The robot consists of two Kuka light weight robot arms, and two DLR five-fingered robot hands. Both the arms and the hands possess active compliance, which makes them suitable for kinesthetic teaching. The fingers are controlled using joint impedance control. Although the hands currently do not possess tactile sensors, each joint is equipped with a torque sensor, which provides information on the forces involved in the manipulation task. This information is however not available when performing kinesthetic teaching, as the forces applied by the human are cancelled out by the forces applied by the object interaction.

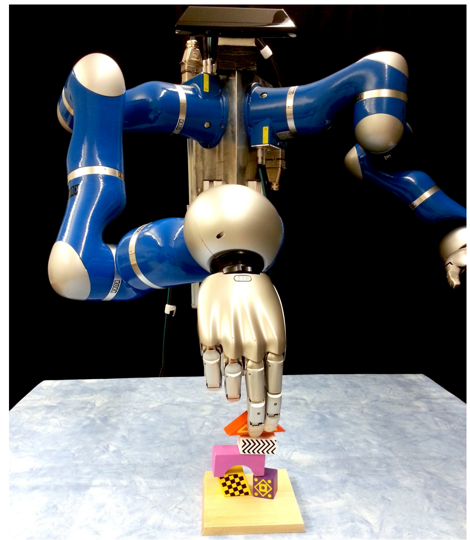

Fig. 1. The bimanual Darias robot used in the grasping experiments.

Rather than performing the kinesthetic teaching directly, we employed a teleoperation approach, as shown in Fig. 2. The human demonstrates the task using the left arm and hand, which is then immediately used to control the right arm and hand. The haptic information from the right hand is also sent back to the left in order to give the human additional force feedback. Controlling fingers individually is not a trivial task for a human demonstrator. The demonstrations were therefore performed in two stages. In the first stage, the robot was demonstrated a set of grasps, which were used to learn a set of eigengrasps [2]. For the second stage, the joint configuration of the left hand was projected into the space of the eigengrasps before being used to define the desired hand configuration of the right hand. This dimensionality reduction greatly simplifies the task for the human demonstrator, leading to faster and smoother demonstrations in the second stage of demonstration. For each demonstration, the joint trajectories and torques of the fingers were recorded. The 3D poses of the hand and objects were also tracked. The objects were tracked at 90Hz using an Optitrack system.

III. Probabilistic Model of Manipulation Phases

In order to learn the phases and phase transitions of a grasping task, we use a modified auto-regressive hidden Markov model. The graphical model is shown in Figure 3. At time step t, the phase is denoted by ρ_t and modeled as

All of the authors are members of the Intelligent Autonomous Systems group at the Technische Universitaet Darmstadt, Germany. Jan Peters is also a member of the Max Planck Institute for Intelligent Systems
{kroemer,neumann}@ias.tu-darmstadt.de
{hoof, peters}@ias.tu-darmstadt.de

a hidden variable, as it cannot be directly observed. The current observed state is denoted by s_t. Given the current phase ρ_t and state s_t, the distribution over actions a_t is given by the policy distribution $p(a_t|s_t,\rho_t)$. The policy distribution is modeled as a separate linear feedback controllers of the form $a_t = K_\rho s_t + \varepsilon_\rho$, where ε_ρ is Gaussian noise.

The state then transitions to a new state s_{t+1} according to the state-transition distribution $p(s_{t+1}|s_t,a_t,\rho_t)$. The state-transition distribution is represented by a separate linear Gaussian model $s_{t+1} = A_\rho s_t + B_\rho a_t + \varepsilon_\rho$ for each phase. It should be noted that both the policy and the state-transition distributions depend on the current phase. Hence, the phases can be inferred from the observed states and actions.

The phase-transition distribution $p(\rho_{t+1}|s_{t+1},\rho_t)$ defines the change in phase over time. One of the key differences to standard hidden Markov models for segmenting demonstrations [1], [5] is the dependence of the phase-transition distribution on the observed state. This additional dependence is important for learning when transitions between phases are more likely to occur. For example, the state dependence allows the model to learn that a transition from the reaching phase to the loading phase is more likely to occur when the hand and object are in close vicinity to each other. The learned phase-transition distribution thus represents the subgoals of the phase. The phase-transition distributions are modeled using logistic regression.

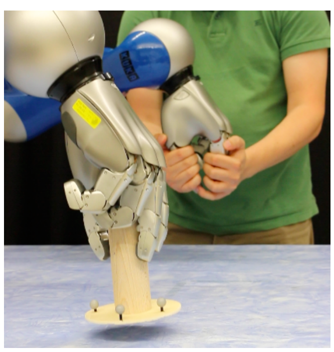

Fig. 2. The teleoperation system used for demonstrating grasps to the robot.

Given a set of human demonstrations, the robot must learn the parameters of the model. However, these distributions depend on the unobserved phases $\rho_{1:N}$. We therefore use an Expectation Maximization (EM) approach to learn the parameters. The EM algorithm is an iterative procedure for learning maximum-likelihood parameters for scenarios with hidden variables. The algorithm consists of two steps, as detailed below.

In the E-step, we infer the joint probabilities $p(\rho_{t-1},\rho_t|s_{1:N+1},a_{1:N})$ and $p(\rho_t|s_{1:N+1},a_{1:N})$ by marginalizing out the phase variables for all other timesteps. The inference is based on the current estimates of the distribution parameters. Given the structure of the graphical model, the marginalization process can be performed efficiently using a message passing algorithm [6].

In the M-step, we use the probabilities of the the phases to determine the maximum likelihood estimates of the distribution parameters. The parameters of the state-transition distributions and the policy distributions can be learned using weighted linear regression, where the weights of each sample are given by the corresponding phase probabilities $p(\rho_t|s_{1:N+1},a_{1:N})$. The state-dependent transition probabilities between phases are modeled using weighted logistic

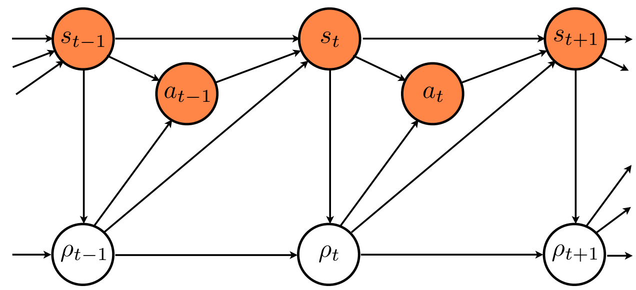

Fig. 3. The graphical model used to learn the phases. The filled circles indicate observed variables, and the white circles represent unobserved variables.

regression, where sample weights are given by the respective joint probabilities$p(\rho_{t-1},\rho_t|s_{1:N+1},a_{1:N})$. The EM algorithm continues to iterate between these two steps until the model has converged.

IV. Conclusion

The proposed probabilistic model allows the robot to learn three key components of the manipulation task from human demonstrations. The learned state-transition distributions $p(s_{t+1}|s_t,a_t,\rho_t)$ model the effects of actions within each phase, and can therefore also be used to recognize the current phase. The phase-transition distributions $p(\rho_{t+1}|s_{t+1},\rho_t)$ define the states wherein transitions between the phases are more likely to occur. These states represent subgoals of the demonstrated manipulation task. For other manipulation tasks, these states may be regarded as task constraints. For example, if an object should not be moved, the robot can try to avoid phases that can affect the object. The final learned component is the policy $p(a_t|s_t,\rho_t)$. The policy of each phase represents a guarded movement when combined with the phase-transition distribution.

V. Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7-ICT-2009-6) under grant agreement 270327 (ComPLACS), 610967 (TACMAN), and 610878 (3rdHand).

References

[1] J. Butterfield, S. Osentoski, G. Jay, and O.C. Jenkins. Learning from demonstration using a multi-valued function regressor for time-series data. In *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, pages 328–333, Dec 2010.

[2] Matei Ciocarlie and Peter Allen. Hand posture subspaces for dexterous robotic grasping. *The International Journal of Robotics Research*, 28:851–867, 07/2009 2009.

[3] J. R. Flanagan, M. C. Bowman, and R. S. Johansson. Control strategies in object manipulation tasks. *Curr Opin Neurobiol*, 16(6):650–659, December 2006.

[4] O. Kroemer, H. van Hoof, G. Neumann, and J. Peters. Learning to predict phases of manipulation tasks as hidden states. In *Proceedings of IEEE International Conference on Robotics and Automation*, 2014.

[5] Scott Niekum, Sachin Chitta, Bhaskara Marthi, Sarah Osentoski, and Andrew G Barto. Incremental semantically grounded learning from demonstration. In *Robotics: Science and Systems 2013*, 2013.

[6] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, February 1989.

[7] Joseph Romano, Kaijen Hsiao, Günter Niemeyer, Sachin Chitta, and Katharine J. Kuchenbecker. Human-inspired robotic grasp control with tactile sensing. *IEEE Transactions on Robotics*, 27:1067–1079, 2011.