Active and Transfer Learning of Grasps by Sampling from Demonstration

Philipp Zech and Justus Piater

Abstract—We guess humans start acquiring grasping skills as early as at the infant stage by virtue of two key processes. First, infants attempt to learn grasps for known objects by imitating humans. Secondly, knowledge acquired during this process is reused in learning to grasp novel objects. We argue that these processes of active and transfer learning boil down to a random search of grasps on an object, suitably biased by prior experience. In this paper we introduce active learning of grasps for known objects as well as transfer learning of grasps for novel objects grounded on kernel adaptive, mode-hopping Markov Chain Monte Carlo. Our experiments show promising applicability of our proposed learning methods.

I. INTRODUCTION

Efficiently learning successful robotic grasps is one of the key challenges to solve for successfully exploiting robots for complex tasks. Considering existing research, grasp learning methods can be grouped into analytic and empirical (or data-driven) methods [1], [2]. Balasubramanian [3] showed that empirical grasp learning grounded upon Programming by Demonstration (PbD) can achieve results superior to planner based, analytic methods.

PbD is a rather simple learning concept constructed from the idea of a robot observing a human demonstrator to then autonomously learn manipulation skills from its observations. Generally, these methods rely on recording hand trajectories. These trajectories then are taken as a basis for either recognizing object and hand shapes (obviously supported by vision), analytic computation of contact points of successful grasps, or a combination of both to learn grasps [1]. In this paper, we propose an alternate approach in that we sidestep the reliance on hand trajectories. Instead, we only require a few user demonstrated grasps as 6D gripper poses. From these, we then learn new grasps by sampling gripper poses relative to a canonical object pose. This ultimately results in a grasp learning method that requires no object specific knowledge.

Treating a grasp as a 6D pose unlocks two key advantages compared to shape-based and analytic methods. First, learned grasps are readily applicable to known objects by just mapping the 6D gripper pose from a canonical object pose to the actual object pose. This requires no further knowledge than the actual object pose. Secondly, acquired grasping skills are easily transferred to novel, as of yet unseen objects, by suitably biasing the learning process. This is by virtue of objects that are similar in shape and size usually have similar grasp affordances. Conversely, shape-based or analytic approaches would require either reconstruction of a shape or computation of new contact points which may easily fail due to clutter, improper segmentation, or missing object information.

Metropolis-Hastings [4] is a popular Markov-Chain Monte Carlo (MCMC) sampler that establishes a Markov chain on a state space $\mathcal{X}$ (e.g., the grasp parameter space) where the stationary distribution of the Markov chain is the target probability density $\pi(x)$ sought-after. By iteratively drawing samples $x_i$ from a proposal distribution $q(x|y)$ one can finally approximate $\pi(x)$.

We propose the application of kernel adaptive, mode-hopping MCMC (Section [II]) for (i) active learning of grasps for known objects and (ii) transfer learning for acquiring grasps for novel objects to learn an object’s grasp density $\pi(x)$ by sampling.

In this work we first introduce active learning of grasps for known objects by combining MCMC Kameleon [5] and Generalized Darting Monte Carlo (GDMC) [6] (Section [IV]). This requires both a rough sketch of the shape of $\pi$ for the former and an initial set of modes (i.e., a set of demonstrated grasps) of $\pi$ for the latter. Given this rough sketch MCMC Kameleon then learns an approximation of $\pi$, while GDMC nudges the proposal generating process to elliptical regions around modes of $\pi$ for efficient mixing between modes. Secondly, we present transfer learning of grasps for novel objects similar in shape and size to already learned objects (Section [V]). This primarily capitalizes on MCMC Kameleon’s learning behavior during a burn-in phase that allows learning of $\pi$ for a novel object (e.g., a soup plate) by approximating it with the Markov chain of a similar object (e.g., a plate). Additionally, we can also reuse demonstrated grasps. This is by virtue of the elliptical regions which for similar objects overlap due to the objects’ similar grasp affordances.

The main contributions of our work thus are:

- the application of kernel adaptive, mode-hopping MCMC for grasp learning,
- active learning of grasps from demonstration without the need for object specific knowledge, and
- transfer learning of grasps for novel objects given a suitable prior by a rough sketch and a few demonstrated grasps of a similar object.

We evaluate our proposed learning methods by a series of carefully designed experiments as presented in Section [VI].

The research leading to these results has received funding from the European Community’s Seventh Framework Programme FP7/2007–2013 (Specific Programme Cooperation, Theme 3, Information and Communication Technologies) under grant agreement no. 610532, SQUIRREL.

Philipp Zech and Justus Piater are with the Faculty of Mathematics, Computer Science and Physics, University of Innsbruck, 6020 Innsbruck, Austria philipp.zech@uibk.ac.at and justus.piater@uibk.ac.at
We conclude in Section [VIII] after discussing our experiments in Section [VII].

II. RELATED WORK

The majority of research in grasp learning from demonstration builds on recording hand trajectories [1], [2]. Given such trajectories, Ekvall and Kragić [7], [8] present a method that uses Hidden Markov Models for classification of a demonstrated grasp, whereas Kjellström et al. [9] and Romero et al. [10], [11], as well as Aleotti and Caselli [12], [13] and Lin and Sun [14] classify demonstrated grasps by a nearest neighbor search among already demonstrated grasps. Zöllner et al. [15] apply Support Vector Machines for classification of demonstrated grasps.

Instead of classifying the demonstrating grasp type and thus learning concrete grasps for specific tasks, another idea is to focus on an object’s or hand’s shape during demonstration. Li and Pollard [16] introduce a shape-matching algorithm that consults a database of known hand shapes for suitably grasping an object given its oriented point representations. Contrary, Kyota et al. [17] represent an object by voxels to identify graspable portions. These portions later are matched against known poses for suitably grasping an object. Herzog et al. [18] learn gripper 6D poses of grasps which are then generalized to different objects by considering general shape templates of objects. Ekvall and Kragić [19], and Tegin et al. [20] extend Ekvall’s and Kragić’s previous work by considering shape primitives which are matched to hand shapes for grasping an object. Also, Aleotti and Caselli [21] extended their work to detect the grasped part of the object, thus enabling generalization of learned grasps to novel objects. Hsiao and Lozano-Pérez [22] segment objects into primitive shapes to map known contact points of grasps to these shapes. They learn contact points from human demonstration. A key feature of shape-based learning methods is that they immediately allow transfer learning of grasps due to the generalization capabilities when only considering the reoccurring parts of an object’s shape.

Yet another approach followed by some researchers is to learn motor skills given trajectories of human demonstrated grasps. Do et al. [23] interpret a hand as a spring-mass-damper system, where proper parameterization of this system allows forming grasps. Kroemer et al. [24] pursue the idea of combining active learning with reactive control based on a damper system, where proper parameterization of this system allows forming grasps. Do et al. [23] interpret a hand as a spring-mass-damper system, where proper parameterization of this system allows forming grasps. Kroemer et al. [24] pursue the idea of combining active learning with reactive control based on a damper system, where proper parameterization of this system allows forming grasps. Do et al. [23] interpret a hand as a spring-mass-damper system, where proper parameterization of this system allows forming grasps. Kroemer et al. [24] pursue the idea of combining active learning with reactive control based on a damper system, where proper parameterization of this system allows forming grasps.

A. Kernel Adaptive Metropolis Hastings

MCMC Kameleon as proposed by Sejdinovic et al. [5] is an adaptive MH sampler approximating highly non-linear target densities $\pi$ in a reproducing Hilbert space. During its burn-in phase, at each iteration it obtains a subsample $z \equiv \{z_i\}_{i=1}^n$ of the chain history $\{x_i\}_{i=1}^\infty$ to update the proposal distribution $q_x(\cdot | x)$ by applying kernel PCA on $z$, resulting in a low-rank covariance operator $C_x$. Using $\nu^2 C_x$ as a covariance (where $\nu$ is a scaling parameter), a Gaussian measure with mean $k(\cdot, y)$, i.e., $\mathcal{N}(f; k(\cdot, y), \nu^2 C_x)$, is defined. Samples $f$ from this measure are then used to obtain target proposals $x^*$. MCMC Kameleon computes pre-images $x^* \in \mathcal{X}$ of $f$ by solving the non-convex optimization problem

$$\arg\min_{x \in \mathcal{X}} g(x),$$

where

$$g(x) = \|k(\cdot, x) - f\|_H^2$$

$$= k(x, x) - 2k(x, y) - 2 \sum_{i=1}^n \beta_i [k(x, z_i) - \mu_x(x)],$$

$$\mu_x = \frac{1}{n} \sum_{i=1}^n k(\cdot, z_i),$$

the empirical measure on $z$, and $y \in \mathcal{X}$. Then, by taking a single gradient descent step along the cost
function $g(x)$ a new target proposal $x^*$ is given by
\begin{equation}
x^* = y - \eta \nabla_x g(x)|_{x=y} + \xi
\end{equation}
where $\beta$ is a vector of coefficients, $\eta$ is the gradient step size, and $\xi \sim \mathcal{N}(0, \gamma^2 I)$ is an additional isotropic exploration term after the gradient. The complete MCMC Kameleon algorithm then is

- at iteration $t + 1$
  1) obtain a subsample $z = \{z_i\}_{i=1}^n$ of the chain history $\{x_i\}_{i=0}^{t+1}$.
  2) sample $x^* \sim q_d(\cdot | x_t) = \mathcal{N}(x_t, \gamma^2 I + \nu^2 M_k \mu H_{k,n}^T)$.
  3) accept $x^*$ with MH acceptance probability $\alpha(x, y) = \min\left\{1, \frac{\pi(y)q(y|x)\mathcal{Z}}{\pi(x)q(x|y)\mathcal{Z}}\right\}$, where $M_k = 2\eta [\nabla_y k(x, z_i)]|_{x=y}, \ldots, \nabla_y k(x, z_i)]|_{x=y}$ is the kernel gradient matrix obtained from the gradient of $g$ at $y$, $\gamma$ is a noise parameter, and $H$ is an $n \times n$ centering matrix.

B. Generalized Darting Monte-Carlo

Generalized Darting Monte Carlo (GDMC) [6] essentially is an extension to classic MH samplers by equipping them with mode-hopping capabilities. Such a mode-hopping behavior is beneficial in case of (i) approximating a highly non-linear, multimodal target $\pi$, and (ii) counterattack the customary random-walk behavior of MH samplers by efficiently mixing between modes.

The idea underlying GDMC is to place elliptical jump regions around known modes of $\pi$. Then, at each iteration, a local MH sampler is interrupted with probability $P_{\text{check}}$, that is, $u_1 > P_{\text{check}}$ where $u_1 \sim U[0, 1]$ to check whether the current state $x_t$ is inside a jump region. If $u_1 < P_{\text{check}}$, sampling continues using the local MH sampler. Otherwise, on $x_t$ being inside a jump region, GDMC samples another region to jump to by
\begin{equation}
P_i = \frac{V_i}{\sum_j V_j}
\end{equation}
where $i$ and $j$ are jump region indices. $V$ denotes the $n$-dimensional elliptical volume
\begin{equation}
V = \frac{\pi^\frac{d^2}{2} e^{\nu^T \lambda_i / 2}}{\Gamma(1 + \frac{d}{2})}
\end{equation}
with $d$ the number of dimensions, $\nu$ a scaling factor, and $\lambda_i$ the eigenvalues resulting from the singular value decomposition of the covariance $\Sigma$ of the Markov chain. $\pi$ denotes the current state of the target density $\pi$. Given this newly sampled region, GDMC then computes a new state $x_{t+1}$ using the transformation
\begin{equation}
x_{t+1} = \mu_{x_{t+1}} - U_{x_{t+1}} S_{x_{t+1}}^{-\frac{1}{2}} S_{x_t}^{-\frac{1}{2}} U_{x_t}^T (x_t - \mu_x)
\end{equation}
where $\mu$ denotes jump regions’ centers (the modes), and $U$ and $S$ again result from the singular value decomposition of the covariance $\Sigma$ of the Markov chain. GDMC accepts the jump proposal $x_{t+1}$ if $u_2 > P_{\text{accept}}$ where $u_2 \sim U[0, 1]$ and
\begin{equation}
P_{\text{accept}} = \min\left[1, \frac{n(x_t)\pi(x_t)}{n(x_{t+1})\pi(x_{t+1})}\right]
\end{equation}
with $n(\cdot)$ denoting the number of jump regions that contain a state $x_t$. If $x_t$ is outside a jump region, it is counted again, i.e., $x_{t+1} = x_t$.

IV. ACTIVE LEARNING OF GRASPS

We formulate a grasp $g$ as a 7D vector $g = (x, y, z, q_w, q_x, q_y, q_z)^\top$, where $x, y, z$ denote the cartesian coordinates of a gripper, and $q_w, q_x, q_y, q_z$ its orientation in quaternion notation about an object. For each grasp, we define a quality measure by the Grasp Wrench Space (GWS) [33] denoted $\mu_{\text{GWS}}$. This measure then allows us to define a target density $\pi(g)$ with $g \in \mathcal{X}$. Observe that $\mu_{\text{GWS}}$ defines a valid density function as $\forall g: \mu_{\text{GWS}}(g) \geq 0$. Further, by introducing the normalization constant $Z$ with $Z = \sum_{g} \mu_{\text{GWS}}(g)$ (where $n$ is the number of known grasps) we have that $\frac{1}{Z} \int \pi(g)dg = 1$.

Our active learning method takes as input a rough sketch of $\pi$ as well as a set of demonstrated grasps. According to Sejdinovic et al. [5] such a rough sketch to initialize MCMC Kameleon does not need to be a proper Markov chain. Instead, it suffices if it provides good exploratory properties of the target $\pi$. We construct such a rough sketch by running a purely random walk MH sampler on the object to be learned. However, we do not take the resulting Markov chain as an initial sketch but instead the set of proposals generated during the random walk, irrespective of whether a proposal was accepted or not. The rationale behind this is that using a purely random MH sampler generally does not result in any learned grasps (Section VII). Hence, the resulting Markov chain essentially would not contain any samples and thus does not inhibit good exploratory properties of $\pi$. On the other hand, the set of proposals as generated during the random walk encapsulates enough information regarding an approximation of the shape $\pi$. Thence, it suffices as a rough sketch to initialize MCMC Kameleon. The random walk MH sampler employed for this uses a Gaussian proposal for the position and a von-Mises-Fisher proposal for the orientation, i.e.,
\begin{align*}
g_{\text{pos}} & = \mathcal{N}(g_{\text{pos}}, \Sigma) \\
g_{\text{ori}} & = C_\kappa(\kappa) \exp(\kappa x_{\text{ori}}^\top x),
\end{align*}
where $\kappa$ is a concentration parameter and $x$ a $p$-dimensional unit direction vector. We use the same probability measures as defined for MCMC Kameleon by the GWS.

In a real-world environment, the set of demonstrated grasps would be established by moving the robot’s gripper towards a position and into a pose, where it can grasp the object. The gripper’s position in cartesian space as well as its orientation about the object in SO(3) are then recorded and treated as a demonstrated grasp. In this work however we only study our grasp learning methods in simulation (Section VII). Thus, we randomly select points on the object’s
surface to then find a grasp by optimizing the gripper’s pose about its orientation [34].

Given a rough sketch of $\pi$ and a set of user demonstrated grasps, the complete learning method then can be sketched as:

- at iteration $t + 1$
  - attempt to perform a jump move according to the procedure as outlined in Section III-B
  - otherwise, sample locally using MCMC Kameleon as outlined in Section III-A

As a kernel $k$ for MCMC Kameleon we chose a Gaussian kernel. Whilst not rigorously applicable in quaternion space, it allows us to model the dependency between a gripper’s position and its orientation. Further, during our experiments we found that a Gaussian kernel works quite well in practice.

V. TRANSFER LEARNING OF GRASPS

Transfer learning fundamentally captures the idea of reusing existing knowledge or already acquired skills to solve problems similar to the original one. For transfer learning of grasps for novel, as of yet unseen objects this ultimately boils down to reusing both the Markov chains constructed when learning to grasp a known object and the respective set of user demonstrated grasps. A crucial factor for the success of this procedure however is that the known and the novel object are similar in shape and size (e.g., a plate and a soup plate).

Given that this constraint is satisfied reusing of Markov chains and grasps is feasible due to both MCMC Kameleon’s learning behavior during a burn-in phase, as well as GDMC’s construction of elliptical regions around known modes. As discussed in Section III-B GDMC samples a new state $x_{t+1}$ by applying the transformation as outlined in equation (6).

As this transformation does not tie a new state $x_{t+1}$ exactly to a mode, but instead into the elliptical region constructed around it, there is a high probability that a new state $x_{t+1}$ is close to a mode of the grasp density $\pi$ for the novel object. Thus, jump moves as done by GDMC are valid in the sense that they again nudge the proposal generating process close to modes of $\pi$. Apart from that, recycling of existing Markov chains and already demonstrated grasps yields substantial time savings by sidestepping both construction of a rough sketch for a novel object and by having a user demonstrate new grasps.

VI. EXPERIMENTAL METHOD

We evaluated our learning methods with 9 different objects as depicted in Figures 1 and 2. In total we performed 5 experiments using RobWork [35], a robotics and grasp simulator. Our first experiment acts as a baseline that allows us to compare the efficiency of our active learning method to a purely random walk (as sketched in Section IV). The next two experiments were designed to evaluate our active learning method. First, MCMC Kameleon was initialized with a random sketch consisting of the trace of a purely random walk MH sampler as discussed in Section IV.

The last two experiments were designed to evaluate our transfer learning method. For initializing MCMC Kameleon for both of these we reused the Markov chains constructed when learning to grasp a similar object. As necessary user demonstrated grasps, in the penultimate experiment we used similar modes, that is, grasps that were demonstrated for a similar object. For the last experiment we used grasps demonstrated on the actual object. This choice of experimental design allows us to evaluate whether our proposed transfer learning method can work with no object specific knowledge at all.

Table I shows our parameterization of MCMC Kameleon and GDMC for our experiments. The values were established during a series of preliminary experiments. For all experiments we used 5 demonstrated grasps.

VII. RESULTS AND DISCUSSION

For all three objects from Figure 1 our active learning method found an additional number of grasps as is shown in Table II. Further, Table III clearly shows that combin-

| Iterations | $\gamma$ | Subsample size | $\nu$ [5] | Burn-in | $P_{check}$ | $\epsilon$ |
|------------|----------|----------------|----------|---------|-------------|----------|
| 1000       | 0.0001   | 100            | $\frac{2.38}{\sqrt{\pi}}$ | 100     | 0.6         | 0.7      |
ing MCMC Kameleon with GDMC drastically outperforms a purely random walk. Also, our active learning method actually works without any knowledge except a few user demonstrated grasps. This is visible from Table II when we did our experiments with MCMC Kameleon initialized with a random sketch. Also visible from Table II, the more complex an object’s shape, the more difficult it is to learn grasps for it (cf. the pitcher with the pan or plate; generally, for the former, fewer grasps were learned). We thus infer that our active learning method for grasping from user demonstration is successful. The top row from Figure 3 shows grasps resulting from our active learning method when applied to the objects from Figure 1.

For transfer learning of grasps for novel, as of yet unseen objects we arrive at the same conclusion as for active learning of grasps. Our learning method again was successful in finding grasps (Table III). Further, as is evident from Table III our learning method generally is able to learn new grasps for novel objects without the need for any user demonstrated grasp for the specific object (cf. pans and plates). However, as can also be seen from the data in Table III our transfer learning method may fail drastically. For both pitchers our learning method failed in learning grasps using similar modes. This is by virtue of the vastly differing sizes and geometries of the pitchers. Obviously, taking the modes and the Markov chain of the pitcher from Figure 1 as a rough sketch as well as initial modes for the grasp density of the tall pitcher from Figure 2 (top row) is a lead balloon. The discrepancy of the size and the geometry of these objects is just too big. The bottom rows from Figure 3 show the outcomes of our transfer learning method when applied to the objects from Figure 2.

To conclude, we state that the combination of MCMC Kameleon and GDMC yields good exploratory properties when searching for feasible grasps in an object’s grasp space by requiring no more input than a few user demonstrated grasps as 6D gripper poses. This is evident from both Tables II and III in that the number of misses generally is substantially smaller than the total number of collisions and grasps where the object slipped out of the gripper.

### VIII. Conclusions

We have presented both a novel method for active learning of grasps as well as a novel method for transfer learning of
grasps, suitably biased by prior experience. We have shown that learning of grasps is feasible without the requirement of object related knowledge. Our learning methods require nothing more than a few demonstrated grasps.

Both our learning methods are grounded on MCMC sampling, more specifically a combination of MCMC Kameleon and GDMC. These algorithms each have advantageous characteristics. MCMC Kameleon allows sampling from highly non-linear distributions, whereas GDMC tackles the issue of properly exploring a multimodal distribution. We found that a combination of both ideally fits the problem of active and transfer learning of grasps. Our results as shown in Tables II and III further undermine our conclusions.

Concerning transfer learning of grasp for novel, as of yet unseen objects, we further want to highlight two observations. First, reusing an existing Markov chain allows boosting of our learning methods by avoiding construction of an initial rough sketch of $\pi$ for an object. Secondly, given that two objects are (i) not too dissimilar in shape and size, and (ii) properly aligned by the same canonical pose, then our transfer learning method is capable of learning grasps for novel objects without any object specific knowledge.

REFERENCES

[1] J. Bohg, A. Morales, T. Asfour, and D. Kragic, “Data-driven Grasp Synthesis — A Survey,” Robotics, IEEE Transactions on, vol. 30, no. 2, pp. 289–309, 2014.
[2] A. Sahbani, S. El-Khouyr, and P. Bidaud, “An Overview of 3D Object Grasp Synthesis Algorithms,” Robotics and Autonomous Systems, vol. 60, no. 3, pp. 326–336, 2012.
[3] R. Balasubramanian, L. Xu, P. D. Brook, J. R. Smith, and Y. Matsuoka, “Physical Human Interactive Guidance: Identifying Grasping Principles from Human-planned Grasps,” IEEE Transactions on Robotics, vol. 28, no. 4, pp. 899–910, 2012.
[4] W. K. Hastings, “Monte Carlo Sampling Methods Using Markov Chains and Their Applications,” Biometrika, vol. 57, no. 1, pp. 97–109, 1970.
[5] D. Sejdić, H. Strathmann, M. L. Garcia, C. Andreu, and A. Grettol, “Kernel Adaptive Metropolis-Hastings,” in Proceedings of the 31st International Conference on Machine Learning, E. P. Xing and T. Jebara, Eds., vol. 32, 2014, pp. 1665–1673.
[6] C. Sminchisescu and M. Welling, “Generalized darting Monte Carlo,” Pattern Recognition, vol. 44, no. 10–11, pp. 2738–2748, 2011.
[7] S. Ekvall and D. Kragic, “Interactive Grasp Learning Based on Human Demonstration,” in IEEE International Conference on Robotics and Automation, ICRA’04, vol. 4. IEEE, 2004, pp. 3519–3524.
[8] ——, “Grasp Recognition for Programming by Demonstration,” in IEEE International Conference on Robotics and Automation, ICRA’05. IEEE, 2005, pp. 748–753.
[9] H. Kjellström, J. Romero, and D. Kragic, “Visual Recognition of Grasps for Human-to-Robot Mapping,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS’08. IEEE, 2008, pp. 3192–3199.
[10] J. Romero, H. Kjellström, and D. Kragic, “Human-to-Robot Mapping of Grasps,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop on Grasp and Task Learning by Imitation, IROS’08, 2008.
[11] ——, “Modeling and Evaluation of Human-to-Robot Mapping of Grasps,” in International Conference on Advanced Robotics, ICAR’09. IEEE, 2009, pp. 1–6.
[12] J. Aleotti and S. Caselli, “Grasp Recognition in Virtual Reality for Robot Pregrasp Planning by Demonstration,” in IEEE International Conference on Robotics and Automation, ICRA’06. IEEE, 2006, pp. 2801–2806.
[13] ——, “Robot Grasp Synthesis from Virtual Demonstration and Topology-preserving Environment Reconstruction,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS’07. IEEE, 2007, pp. 2692–2697.