Active and Transfer Learning of Grasps by Kernel Adaptive MCMC

Philipp Zech, Hanchen Xiong and Justus Piater

***Abstract*— Human ability of both versatile grasping of given objects and grasping of novel (as of yet unseen) objects is truly remarkable. This probably arises from the experience infants gather by actively playing around with diverse objects. Moreover, knowledge acquired during this process is reused during learning of how to grasp novel objects. We conjecture that this combined process of active and transfer learning boils down to a random search around an object, suitably biased by prior experience, to identify promising grasps. In this paper we present an active learning method for learning of grasps for given objects, and a transfer learning method for learning of grasps for novel objects. Our learning methods apply a kernel adaptive Metropolis-Hastings sampler that learns an approximation of the grasps' probability density of an object while drawing grasp proposals from it. The sampler employs simulated annealing to search for globally-optimal grasps. Our empirical results show promising applicability of our proposed learning schemes.**

I. Introduction

Establishing efficient strategies for learning precision grasps is one of the key challenges in robotics research. Currently, a substantial amount of work in this area relies on an object's shape or shape-related information (e.g., surface normals or image gradients) reconstructed from vision [1] to analytically compute object specific contact points [2], [3]. A complementary idea though is to neglect an object's shape and instead utilize its pose to learn feasible grasps by sampling gripper poses relative to an object's pose. This results in grasp learning methods that require very little additional object specific knowledge (given its pose) to guide the search for feasible grasps (cf. Detry et al. [4] who utilized 3D edge information).

Such a *postural* interpretation of a grasp by a gripper's pose relative to an object's pose possesses two key advantages compared to shape-based grasp learning. First, learned grasps can be readily applied to known objects by just aligning a gripper's pose relative to an object's pose. This requires no knowledge other than an object's pose. Secondly, the postural interpretation allows seamless transfer of grasps between similar (e.g., in shape and size) objects by just sampling a new gripper pose suitably biased by an already known gripper pose. Conversely, a shape-based approach in both cases requires reconstructing an object's shape which may easily fail due to clutter or improper segmentation.

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, Hanchen Xiong, and Justus Piater are with the Institute of Computer Science, University of Innsbruck, 6020 Innsbruck, Austria. philipp.zech@uibk.ac.at, hanchen.xiong@uibk.ac.at, and justus.piater@uibk.ac.at

Metropolis-Hastings (MH) [5] is a popular Markov-Chain Monte Carlo (MCMC) algorithm that constructs a Markov chain on a state space $\mathcal{X}$ (e.g., the grasp parameter space) where the stationary distribution of possible states is the target probability density $\pi(x)$. By drawing samples $x_0, x_1, x_2, \dots$ from a proposal distribution $q(x|y)$ one can iteratively approximate $\pi(x)$. We propose the application of kernel adaptive MCMC (Section III) for active learning of grasps for given objects (Section IV) and transfer learning for acquiring grasps for novel objects (Section V) by learning, via approximation, an object's unknown grasp density.

In this paper we first introduce active learning of grasps for given objects by MCMC Kameleon (Section IV). This relies on a rough sketch[1] of the shape of the grasps' probability density π of a specific object. Given this rough sketch, MCMC Kameleon then learns an approximation to π during its burn-in phase for subsequent sampling of grasps. Secondly, we present transfer learning to learn grasps for novel objects similar in *shape and size* to objects already learned (Section V). This capitalizes on MCMC Kameleon's learning behavior during its burn-in phase which allows learning grasps for different objects of the same type (e.g., plates or soup plates) using a rough sketch of the shape and size of a similar object.

A common problem in applying MCMC is deterioration of the sampler, i.e., the repeated sampling of rejected proposals. In our case, this amounts to repeated sampling of gripper poses yielding infeasible grasps and thus zero success probability. To tackle this problem, we utilize the object's rim computed from its point cloud to define a meaningful probability measure for infeasible grasps. This allows us to heuristically nudge the sampler towards nonzero regions of the grasp success probability function. As we only focus on precision grasps, attracting grasps towards rims (in the absence of any graspability information) is an effective heuristic, since rims are very likely to allow precision grasps. This is the only shape information required by our method.

The key contributions of this paper are:

- A heuristic for calculating a pseudo measure for infeasible grasps to overcome the obstacle of applying MCMC for sampling gripper poses from a grasp density function where for infeasible grasps no quality measure can be calculated.
- The application of kernel adaptive MCMC for learning of precision grasps by active sampling.

[1]Observe that Sejdinovic et al. characterize such a rough sketch as just a scheme with good exploratory properties of the target; there is no need for it to result from a converged or even valid Markov chain.

- A transfer learning scheme for learning of precision grasps for novel objects by using suitable prior information by known gripper poses for similar objects, thus facilitating generalization of grasps.

We evaluate our proposed learning methods by a series of carefully designed experiments as presented in Section VI. We conclude our work in Section VII.

II. Related Work

Traditionally, grasp learning methods rely on vision for both (i) finding graspable parts of the object and (ii) evaluating the learned grasps. Detry et al. [4] learn grasp affordance densities by (i) establishing a grasp affordance model for an object, and (ii) training this model by autonomous exploration, e.g., sampling, where grasp affordance densities are modeled by Kernel Density Estimation. Based on early visual cues an initial density is created which is then trained and finally yields the empirical density of the object's grasps. Detry et al. [6] study transfer learning of grasping strategies. Their method is based on learning object shape prototypes to generalize grasping strategies among different objects. Kopicki et al. [7] propose to learn grasps by computing a gripper's shape as to a specific grasp relative to an object's shape. Their method allows transfer of grasps by matching the gripper's shape to shapes of novel objects. Fischinger et al. [8] investigate grasping objects from cluttered scenes on the grounds of a point cloud of the scene. Using Symmetry Height Accumulated Features their system is trained by Support Vector Machines yielding grasp classifiers which subsequently allow a robot to decide on an optimal 6D gripper pose relative to its environment. The work of Kroemer et al. [9] combines the two notions of active learning and reactive control into a grasp learning framework. Their system implements a hybrid architecture, where, on the high-level side, a reinforcement learner (the active learner) determines grasps and, on the low level, a reactive controller is responsible for action execution. Further, vision is used to incorporate geographic scene information for optimal grasp learning. Recently, Lenz et al. [10] applied deep learning for learning grasps from an RGB-D view of a scene. Their method ultimately implements rectangle-based grasp detection [11] using two neural networks, i.e., (i) for detecting potential candidate rectangles on the object, and (ii) to extract top rectangles from these candidate rectangles. The top rectangles then represent optimal grasps with respect to a gripper's pose. Rodriguez et al. [12] suggest early abort and retry to reduce the time to learn a grasp. In their work, grasp signatures are used to establish probabilistic models to track the instantaneous probability of a grasp to succeed. Given that the model suggests that a grasp may not succeed, it is aborted early and retried using slightly modified parameters. Abort and retry are modeled as a Markov chain which subsequently can be used to minimize the time for learning grasps. Saxena et al. [13] introduce a vision-based grasping system which learns grasping points for images of cluttered scenes. By supervised learning, their system learns visual features for identifying a 3-D point and an orientation at which to grasp the object. Given this information, a path planner then calculates the optimal trajectory to reach the object and apply a grasp. Stulp et al. [14] investigate learning grasps with a special emphasis on uncertainty. Their key idea in learning optimal grasp poses is to sample actual object poses from a distribution that represents the state estimation uncertainty. Dynamic movement primitives necessary to reach the object are learned by reinforcement learning.

Our learning methods mainly differ in that they require only weak information from object models (such as point clouds). Except for Detry et al. [6] and Kopicki et al.[7], most other research does not address transfer learning. Our approach to transfer learning differs in that we do not rely on precise shape information to generalize grasping strategies but instead only on object poses. The sole purpose of the object's rim points is to keep the sampler from deterioration. Our method would still find grasps without them, but it would take longer, as the sampler would degenerate to a purely random walk. In terms of heuristic search strategies, the work of Stulp et al. [14] is related to ours. Yet, contrary to our grasp learning method, Stulp et al. rely on an initial, feasible grasp and preshape posture for their method to work. As Table II shows, our grasp learning method, in contrast, also learns feasible grasps from random initialization.

III. Kernel-Adaptive MCMC

MCMC Kameleon as proposed by Sejdinovic et al. [15] is an adaptive MH sampler approximating highly non-linear target densities π. During its burn-in phase, at each iteration it obtains a subsample $\mathbf{z} = \{z_i\}_{i=1}^{n}$ of the chain history $\{x_i\}_{i=0}^{t-1}$ for updating the proposal distribution $q_{\mathbf{z}}(\cdot \mid x)$ by applying kernel PCA on $\mathbf{z}$, resulting in a low-rank covariance operator $C_{\mathbf{z}}$. Using $\nu^2 C_{\mathbf{z}}$ as a covariance (where ν is a scaling parameter), a Gaussian measure with mean $k(\cdot, y)$, i.e., $\mathcal{N}(f; k(\cdot, y), \nu^2 C_{\mathbf{z}})$, is defined. Samples f from this measure are subsequently used to obtain target proposals x^*.

MCMC Kameleon computes pre-images $x^* \in \mathcal{X}$ of f by considering the non-convex optimization problem

$$\arg\min_{x \in \mathcal{X}} g(x), \tag{1}$$

where

$$\begin{aligned} g(x) &= \| k(\cdot, x) - f \|_{\mathcal{H}_k}^2 \qquad (2) \\ &= k(x,x) - 2k(x,y) - 2\sum_{i=1}^{n} \beta_i \left[k(x, z_i) - \mu_{\mathbf{z}}(x) \right], \end{aligned}$$

$\mu_{\mathbf{z}} = \frac{1}{n}\sum_{i=1}^{n} k(\cdot, z_i)$, the empirical measure on $\mathbf{z}$, 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

$$x^* = y - \eta \nabla_x g(x)|_{x=y} + \xi \tag{3}$$

where β is a vector of coefficients, η the gradient step size, and $\xi \sim \mathcal{N}(0, \gamma^2 I)$ an additional isotropic exploration term after the gradient. The complete MCMC Kameleon algorithm then is

- at iteration $t + 1$

1) obtain a subsample $\mathbf{z} = \{z_i\}_{i=1}^{n}$ of the chain history $\{x_i\}_{i=0}^{t-1}$,
2) sample $x^* \sim q_{\mathbf{z}}(\cdot \mid x_t) = \mathcal{N}(x_t, \gamma^2 I + \nu^2 M_{\mathbf{z},x_t} H M_{\mathbf{z},x_t}^T)$,
3) accept x^* with MH acceptance probability $\alpha(x,y) = \min\left\{1, \frac{\pi(y)q(x|y)}{\pi(x)q(y|x)}\right\}$,

where $M_{\mathbf{z},y} = 2\eta\left[\nabla_x k(x,z_1)|_{x=y}, \ldots, \nabla_x k(x,z_n)|_{x=y}\right]$ is the kernel gradient matrix obtained from the gradient of (**??**) at y, γ is a noise parameter, and H is an $n \times n$ centering matrix.

IV. Active Grasp Learning

We represent a grasp g as a 7D vector, i.e., $g = (x, y, z, q_w, q_x, q_y, q_z)^T$, 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 relative to an object. To each grasp g is associated a measure μ_{GWS} based on the Grasp Wrench Space (GWS) [16], indicating its quality. This measurability then allows us to define a target density $\pi(g), g \in \mathcal{X}$.

A. Metropolis Criteria

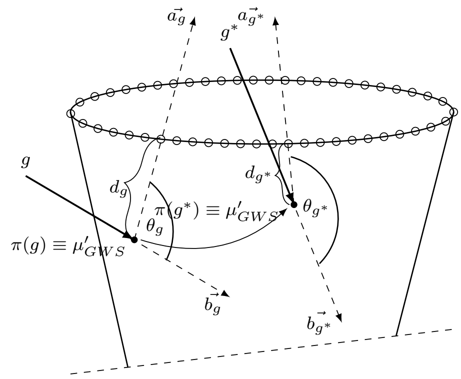
The GWS is only defined for a feasible grasp g; i.e., if a grasp g cannot be applied to an object, μ_{GWS} cannot be calculated. Therefore, a heuristic is needed to calculate approximate measures for samples g that do not represent a feasible grasp. This stems from MCMC's need for continuous probability densities to efficiently converge to the global optimum by the MH acceptance criterion.

1) Feasible Grasps: In the case of a feasible grasp g, the probability measure is calculated in terms of the GWS. Thus, $\pi(g) \propto \mu_{GWS}$, and the acceptance probability of the proposed sample g is calculated by the MH acceptance criterion. The likelihood of a new proposal g^* conditioned by g is given by the proposal distribution q's density for g^*, conditioned by g, i.e., $q(g^* \mid g)$. Conversely, the likelihood of g conditioned on g^* is given by $q(g \mid g^*)$.

2) Infeasible Grasps: In the case of an infeasible grasp g, μ_{GWS} is zero, implying that the MH acceptance probability would always be zero. This eventually robs the algorithm of any clue whether the current search direction is promising or not. To overcome this problem, we apply a heuristic to learn whether an infeasible proposal g^* points into a direction where a feasible grasp could be found or not. Figure 1 illustrates how we calculate this heuristic quality measure μ'_{GWS}.

The idea underlying the heuristic quality measure μ'_{GWS} is based (i) on the angle θ_g between the gripper's pose and the vector between the gripper direction and a closest rim point as shown in Figure 1 by the vector $\vec{a}_g$, and (ii) the distance of the gripper position and a closest rim point indicated by d_g. Setting the constant min_GWS to 0.01, i.e., the lowest tolerable quality measure for a grasp g, and using μ'_{GWS} as defined in Figure 1, we move the search for grasps g^* towards the rims of an object.

To detect rim points of objects by their point clouds, first, for each point o, we construct a spherical neighborhood N around o. Next, we connect o to each of its neighboring



$\mu'_{GWS} = \frac{min\_GWS}{1+(d*(\pi-\theta))}$
$\pi(g^*) \equiv \mu'_{GWS}$
$\pi(g^*) \geq \pi(g)$

Fig. 1: Illustration of our heuristic used to calculate μ'_{GWS}. Here, π denotes the mathematical constant and $\pi(\cdot)$ a probability density.

points $i \in N$ to obtain vectors $\vec{p}_i$. Then, if o is a non-rim point and N is flat, $\left\|\sum_{i \in N} \vec{p}_i\right\|^2 \approx 0$. However, if o is a rim point, then $\left\|\sum_{i \in N} \vec{p}_i\right\|^2 > \zeta$, where ζ is a threshold that is tuned based on the density and noise of the point clouds. For all other non-rim points, $0 < \left\|\sum_{i \in N} \vec{p}_i\right\|^2 \leq \zeta$. Figures 2 and 3 show the results of the rim detection applied to the objects used in our experiments (Section VI).

Observe that both quality measures, i.e., μ_{GWS} and μ'_{GWS}, are valid density functions in that (i) their values are always greater or equal zero and (ii) by introducing some normalization constant Z, i.e., $Z = \sum_{i=1}^{n} \mu_{GWS}^{i}$ (where n is the number of known, feasible grasps for an object), applied to μ_{GWS}, i.e., $\frac{1}{Z}\mu_{GWS}$, we have that $\int \pi(g)\mathrm{d}g = 1$ (similarly, this also holds for μ'_{GWS}).

B. Simulated Annealing

Using a plain Metropolis-Hastings (MH) acceptance criterion $\alpha(g^*, g) = \min\left\{1, \frac{\pi(g^*)q(g^*|g)}{\pi(g)q(g|g^*)}\right\}$, MCMC Kameleon considers the whole search space $\mathcal{X}$, i.e., the Markov chain will likely visit bad samples as well as good ones. In learning to grasp objects, however, we are only interested in good samples, i.e., feasible grasps. To tackle this problem we apply simulated annealing (SA) [17]. The idea is to equip the sampler with an initial temperature $T > 0$ which decreases over the sampling process with the effect of gradually decreasing the probability of accepting poor samples while exploring the state space $\mathcal{X}$. Consequently, while traversing the Markov chain the sampler more likely moves in regions most likely containing the global optima. We thus extend the plain MH acceptance criterion by raising it to the power of T, where T is the current system temperature, i.e.,

$$\alpha(x,y) = \min\left\{1, \frac{\pi(g^*)q(g|g^*)}{\pi(g)q(g^*|g)}\right\}^T \tag{4}$$

$$T = \max\left\{T_N, \frac{T_N}{T_0}^{\frac{j}{N}}\right\} \tag{5}$$

where T_0 is the initial temperature, T_N the final temperature, N the number of iterations, and j the current iteration. Using (4) we slowly decrease the acceptance probability of bad, i.e., of low quality, grasp proposals g^*. Equation (4) slowly decreases the temperature T over time.

C. Complete Learning Method

We use a Gaussian proposal for both position *and* orientation to capture the relation between gripper position and orientation relative to an object. Further, as a kernel k for MCMC Kameleon we use a Gaussian kernel. Although Gaussian proposals and kernels are not rigorously applicable in quaternion space, this choice allows us to easily model the dependencies between gripper positions and orientations, which is crucial for our method to perform well.

Our complete learning algorithm then is

- at iteration $t+1$
  1) obtain a subsample $\mathbf{z} = \{z_i\}_{i=1}^n$ of the chain history $\{g_i\}_{i=0}^{t-1}$,
  2) sample $g^* \sim q_{\mathbf{z}}(\cdot \mid g_t) = \mathcal{N}(g_t, \gamma^2 I + \nu^2 M_{\mathbf{z},g_t} H M_{\mathbf{z},g_t}^T)$,
  3) calculate $\pi(g^*)$ using either μ_{GWS} in the case of a feasible grasp or μ'_{GWS} otherwise
  4) accept g^* with MH acceptance probability (4).

The chain history $\{g_i\}_{i=0}^{t-1}$ is initialized by samples as retrieved[2] from a random walk (RW) MCMC sampler using a Gaussian proposal for the position and a von Mises-Fisher proposal for the orientation, i.e.,

$$g_{pos}^* = \mathcal{N}(g_{pos}^t, \Sigma)$$
$$g_{ori}^* = \mathcal{C}_4(\kappa) \exp(\kappa {g_{ori}^t}^\top \mathbf{x}),$$

where κ is the concentration parameter and $\mathbf{x}$ a p-dimensional unit direction vector. We use the same probability measures as defined for MCMC Kameleon.

V. Transfer Learning

Humans apply transfer learning on a daily basis by reusing acquired knowledge and applying it to solve a problem similar to the one the knowledge originally was learned for. Similarly, here we can reuse a chain history $\{g_i\}_{i=0}^{t-1}$ that was learned for some object as a rough sketch from which to learn grasps g for similar objects by properly biasing, i.e., initializing, MCMC Kameleon with such a chain history. This is feasible thanks to the burn-in phase of MCMC Kameleon where it learns an approximation of the target density π. Given that two objects a and b are of similar shape and size (e.g., a plate and a soup plate), the algorithm from Section IV-C can be used for transfer learning of grasping a novel object b given a chain history for object a. Since producing an initial chain history for a novel object is expensive, transfer learning by avoiding the burn-in phase can result in substantial savings.

[2]These samples amount to the "rough sketch" (Sec. I) of the shape of the grasps' success probability.

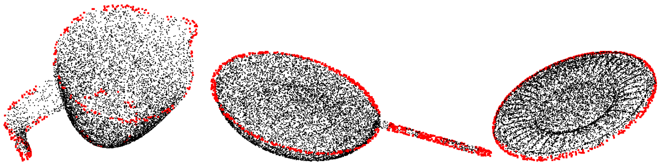


Fig. 2: Object set used for grasp learning with object rims depicted in red (best viewed in color).

TABLE I: Parameters for MCMC Kameleon used during our experiments.

Iterations (N)	γ	Subsample size	ν	Burn-in	T_0	T_N
5000	0.0001	200	$\frac{2.38}{\sqrt{6}}$	1000 2000	1.0	0.05

VI. Experiments

In the following, we discuss (i) our results on learning grasps for a given object (Section VI-A), and (ii) our results on transfer learning of learning grasps for previously unseen objects (Section VI-B). Both our learning algorithms were implemented in Python. As a simulation environment for executing a sampled grasp g and subsequently calculating the GWS we used RobWork [18]. Inside RobWork, all grasps were executed with a Schunk SDH gripper.

A. Grasp Learning

Figure 2 depicts the object set we used for learning grasps of given objects. The experiments were done in two steps: First, we ran a RW MCMC sampler for 5000 iterations to produce initial chain histories $\{g_i\}_{i=0}^{t-1}$. Secondly, we used these chain histories (as "rough sketches") to initialize MCMC Kameleon for learning an approximation of the target density π during burn-in and subsequent grasp sampling (after burn-in). Apart from the simulations just mentioned, we further did a series of experiments where MCMC Kameleon was initialized with a randomly generated chain, without using prior knowledge.

Table I shows our parametrization of MCMC Kameleon used during our experiments[3]. Apart from the value of the scaling parameter ν which was chosen following the suggestion by Sejdinovic et al. [15], the parameters were tuned by a series of preliminary experiments.

For each of the objects from Figure 2 we thus conducted one initial random walk for establishing a chain history $\{g_i\}_{i=0}^{t-1}$, and then four runs using MCMC Kameleon for learning grasps with either a randomly initialized chain (no prior information), or using chain histories from the random walks (using prior information), and differing burn-in durations (Table I). Table II shows our results. Figure 4 (top row) shows sampled gripper poses for the objects from Figure 2.

As is evident from Table II, the complexity of an object's shape clearly drives the success of learning. Generally, the algorithms we used (RW MCMC and MCMC Kameleon)

[3]Our RW MCMC sampler offers two parameters, viz. κ and Σ which were set to 5 and $0.01 I_3$ for all experiments.

TABLE II: Experimental results: number of sampled gripper poses yielding feasible grasps during grasp learning. The numbers in square brackets in the leftmost column denote the burn-in duration and the use of prior information (p) or not (np).

	Pitcher	Pan	Plate
RW MCMC	12	3	20
MCMC Kameleon [1000,np]	12	10	35
MCMC Kameleon [1000,p]	2	34	20
MCMC Kameleon [2000,np]	16	54	95
MCMC Kameleon [2000,p]	14	115	81

learned more gripper poses yielding feasible grasps for the pan and the plate than for the pitcher (except for the RW MCMC and one run of MCMC Kameleon). This is due to the geometry of the pitcher, which is more complex than the pan or the plate. Table II also shows that using prior information generally fosters learning of grasps. MCMC Kameleon generally outperformed RW MCMC thanks to burn-in where it learns an approximation of $\pi(g)$.

Another interesting aspect of MCMC Kameleon that is evident from Table II is the choice of the burn-in duration. Using a longer burn-in results in better learning fo π, which generally results in a substantially larger number of grasps found. Here, we kept the total number of iterations fixed to 5000: If the burn-in duration lasted for 1000 iterations, then MCMC Kameleon had a budget of 4000 iterations for sampling grasps; if the burn-in lasted for 2000 iterations, it had a budget of 3000 iterations.

B. Transfer Learning

For learning of grasps for novel (as of yet unseen) objects we consider two methods for biasing MCMC Kameleon with prior information (representing a rough sketch of the target grasp density $\pi(g)$). First, we reuse complete chain histories $\{g_i\}_{i=0}^{t-1}$ as generated by MCMC Kameleon for similar objects. Secondly, we investigate using subsamples $\mathbf{z}$ as computed by MCMC Kameleon at the last iteration of the burn-in when sampling for a similar object. In the latter case we skip the burn-in. This amounts to a precast covariance operator $C_{\mathbf{z}}$, preventing MCMC Kameleon from learn a better approximation of $\pi(g)$, limiting it to just using what it is provided with. Figure 3 shows our object set for this purpose.

We ran MCMC Kameleon with the same parameters as discussed in Section VI-A (Table I), except for the number of iterations. In the case of using a subsample $\mathbf{z}$ for transfer learning we reduced the default number of 5000 iterations, i.e., the number of iterations was set to 3000 and 4000 respectively, depending on the burn-in duration of the run that produced $\mathbf{z}$. Both chain histories and subsamples for transfer learning are a result of the experiments from Section VI-A.

Hence, for each object from Figure 3 we did (i) two runs with MCMC Kameleon initialized by a chain history as established during an earlier burn-in phase and a randomly selected starting point from that chain, and (ii) two runs with MCMC Kameleon initialized with a kernel $\mathcal{K}$ where the starting point is chosen to be the last element of the kernel $\mathcal{K}$. Observe that in the case of transfer learning, the objects from Figure 3 first were aligned to a canonical pose to be in line with the objects from Figure 2. This is necessary as chain histories and kernels $\mathcal{K}$ are established relative to an object's orientation and size. Table III shows our results. Figure 4 (middle and bottom rows) shows sampled gripper poses for the objects from Figure 3.

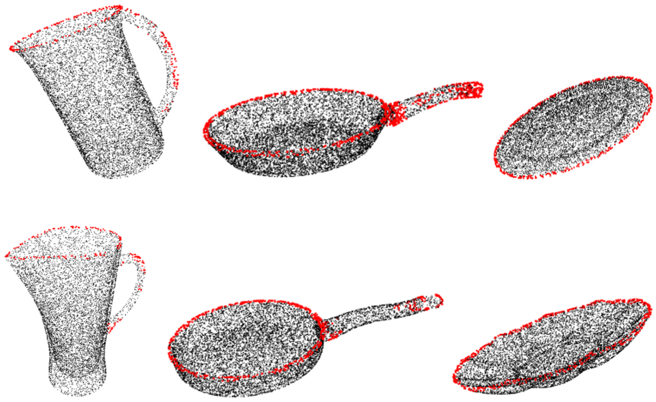


Fig. 3: Object set used for transfer learning with object rims depicted in red (best viewed in color).

TABLE III: Experimental results: number of sampled gripper poses yielding feasible grasps, for transfer learning of grasps. The second and fourth row correspond to runs with 4000 iterations, the sixth and eighth to runs with 3000 iterations. The arrangement of results corresponds to Figure 3. The numbers in square brackets in the leftmost column denote the burn-in duration and the usage of a chain (c) or a subsample ($\mathbf{z}$) to initialize MCMC Kameleon.

	Pitcher	Pan	Plate
MCMC Kameleon [1000,c]	3	53	117
MCMC Kameleon [0,$\mathbf{z}$]	0	0	0
MCMC Kameleon [2000,c]	2	19	7
MCMC Kameleon [0,$\mathbf{z}$]	4	68	0
MCMC Kameleon [1000,c]	3	29	45
MCMC Kameleon [0,$\mathbf{z}$]	0	0	0
MCMC Kameleon [2000,c]	1	6	62
MCMC Kameleon [0,$\mathbf{z}$]	0	41	0

What first leaps to the eye in Table III are the rather poor results we achieved for the pitcher from Figure 3. It turns out that this is due to object misalignment, i.e., the pitchers from Figure 3 were not properly aligned with the pitcher from Figure 2. Apart from that, the similarity in dimensions also plays a key role for transfer learning. Clearly, the two pitchers from Figure 3 are taller than the one from Figure 2, rendering learning with MCMC Kameleon difficult as the initial rough sketch of the shape of the object's grasp density does not fit well. Obviously, misalignment of objects further exacerbates this mismatch.

Our results from Table III show that if MCMC Kameleon is initialized with a subsample $\mathbf{z}$ of an earlier run it may

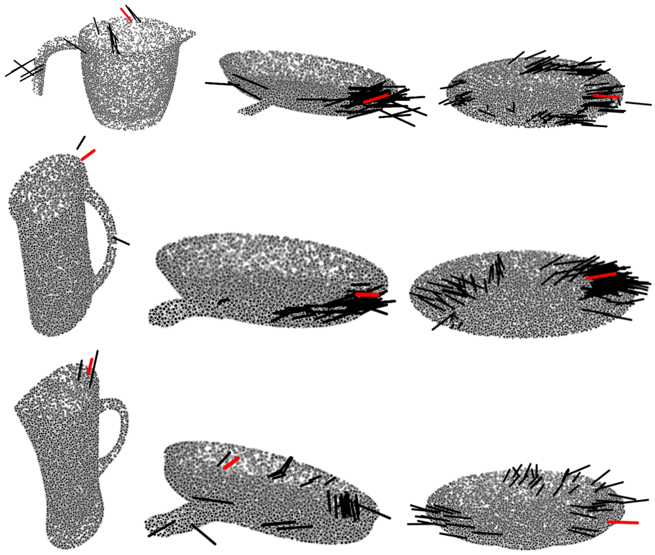


Fig. 4: Results for learning grasps for the objects from Figure 2 (top row) and for transfer learning of grasps for corresponding objects from Figure 3 (middle and bottom rows) with a burn-in duration of 1000. Red pointers indicate optimal grasps (best viewed in color). Observe that grasps are rather unevenly distributed; this results from both running MCMC Kameleon for only 5000 iterations and the use of SA which at some point locks the sampler to a mode of π.

drastically fail. We conclude that this is because a subsample $\mathbf{z}$ is already too specific a sketch of an object's grasp density, i.e., it does not provide the necessary diversity to "recognize" similar objects. Clearly, by again permitting MCMC Kameleon to employ a burn-in phase, such a subsample $\mathbf{z}$ could be adapted to an object. In contrast, our results in Table III demonstrate that initializing MCMC Kameleon by reusing a chain history $\{g_i\}_{i=0}^{t-1}$ is very effective. We thus conjecture that reusing a chain history, i.e., prior experience, together with a burn-in phase for transfer learning results in a good number of feasible grasps for previously unseen objects. Moreover, reusing such chain histories for transfer learning successfully eradicates the need for performing random walks to obtain a rough sketch of the object, thus reducing the computational burden incurred by our learning methods.

To sum up, we can state that the learning methods introduced in this paper have proven successful both in learning of grasps for given objects (Section IV) and in learning of grasps for previously unseen objects, i.e., applying transfer learning for learning of grasps (Section V).

VII. Conclusion

In this paper we have introduced (i) active learning of grasps for given objects and (ii) transfer learning for learning grasps for novel objects. Both our learning methods build upon MCMC Kameleon, an adaptive MH sampler that learns an approximation of a target density π during its burn-in phase.

The experimental evaluation shows that the application of an adaptive MH sampler, e.g., MCMC Kameleon, is promising for grasp learning tasks as discussed in this paper. Our results show that our learning methods allow learning grasps from no knowledge at all, as Table II shows, when MCMC Kameleon is initialized with a random chain, i.e., it uses no prior information. Table II however also shows that our learning methods can be easily boosted if MCMC Kameleon is initialized with suitable prior experience, i.e., a rough sketch of the shape of the grasps' probability density associated with an object. We successfully exploit this aspect of biased initialization in transfer learning (Table III) of grasps for novel objects.

References

[1] J. Bohg, A. Morales, T. Asfour, and D. Kragic, "Data-Driven Grasp Synthesis - A Survey," *IEEE Transactions on Robotics*, vol. 30, no. 2, pp. 289–309, 2014.

[2] K. B. Shimoga, "Robot Grasp Synthesis Algorithms: A Survey," *The International Journal of Robotics Research*, vol. 15, no. 3, pp. 230–266, 1996.

[3] A. Bicchi and V. Kumar, "Robotic Grasping and Contact: A Review," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2000, pp. 348–353.

[4] R. Detry, D. Kraft, O. Kroemer, L. Bodenhagen, J. Peters, N. Krüger, and J. Piater, "Learning Grasp Affordance Densities," *Paladyn*, vol. 2, no. 1, pp. 1–17, 2011.

[5] W. K. Hastings, "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.

[6] R. Detry and J. Piater, "Unsupervised learning of predictive parts for cross-object grasp transfer," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2013, pp. 1720–1727.

[7] M. Kopicki, R. Detry, F. Schmidt, C. Borst, R. Stolkin, and J. Wyatt, "Learning dexterous grasps that generalise to novel objects by combining hand and contact models," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, May 2014, pp. 5358–5365.

[8] D. Fischinger, M. Vincze, and Y. Jiang, "Learning Grasps for Unknown Objects in Cluttered Scenes," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2013, pp. 609–616.

[9] O. Kroemer, R. Detry, J. Piater, and J. Peters, "Combining Active Learning and Reactive Control for Robot Grasping," *Robotics and Autonomous Systems*, vol. 58, no. 9, pp. 1105–1116, 2010.

[10] I. Lenz, H. Lee, and A. Saxena, "Deep Learning for Detecting Robotic Grasps," in *Proceedings of Robotics: Science and Systems*, Berlin, Germany, June 2013.

[11] Y. Jiang, S. Moseson, and A. Saxena, "Efficient Grasping From RGB–D Images: Learning Using a New Rectangle Representation," in *IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2011, pp. 3304–3311.

[12] A. Rodriguez, M. T. Mason, S. S. Srinivasa, M. Bernstein, and A. Zirbel, "Abort and Retry in Grasping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2011, pp. 1804–1810.

[13] A. Saxena, L. Wong, M. Quigley, and A. Y. Ng, "A Vision-based System for Grasping Novel Objects in Cluttered Environments," in *Robotics Research*. Springer, 2011, pp. 337–348.

[14] F. Stulp, E. Theodorou, J. Buchli, and S. Schaal, "Learning to Grasp Under Uncertainty," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 5703–5708.

[15] D. Sejdinovic, M. L. Garcia, H. Strathmann, C. Andrieu, and A. Gretton, "Kernel–Adaptive Metropolis-Hastings," *CoRR*, 2013.

[16] A. T. Miller and P. K. Allen, "Examples of 3D Grasp Quality Computations," in *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, vol. 2. IEEE, 1999, pp. 1240–1246.

[17] P. J. Van Laarhoven and E. H. Aarts, *Simulated Annealing: Theory and Applications*. Springer, 1987.

[18] L.-P. Ellekilde and J. A. Jorgensen, "RobWork: A Flexible Toolbox for Robotics Research and Education," in *41st International Symposium on Robotics (ISR) and 2010 6th German Conference on Robotics (ROBOTIK)*. VDE, 2010, pp. 1–7.