Robotica (2015) volume 33, pp. 1131–1146. © Cambridge University Press 2014
doi:10.1017/S0263574714001192

On transferability and contexts when using simulated grasp databases

Jimmy A. Rytz†*, Lars-Peter Ellekilde†, Dirk Kraft‡, Henrik G. Petersen† and Norbert Krüger†

†*University of Southern Denmark, Maersk MC-Kinney Møller Institute*
‡*Syddansk Universitet, Mrsk Mc-Kinney Mller Instituttet*

(Accepted April 17, 2014. First published online: May 22, 2014)

SUMMARY
It has become a common practice to use simulation to generate large databases of good grasps for grasp planning in robotics research. However, the existence of a generic simulation context that enables the generation of high quality grasps that can be used in several different contexts such as bin-picking or picking objects from a table, has to our knowledge not yet been discussed in the literature.

In this paper, we investigate how well the quality of grasps simulated in a commonly used "generic" context transfers to a specific context, both, in simulation and in the real world.

We generate a large database of grasp hypotheses for several objects and grippers, which we then evaluate in different dynamic simulation contexts e.g., free floating (no gravity, no obstacles), standing on a table and lying on a table.

We present a comparison on the intersection of the grasp outcome space across the different contexts and quantitatively show that to generate reliable grasp databases, it is important to use context specific simulation.

We furthermore evaluate how well a state of the art grasp database transfers from two simulated contexts to a real world context of picking an object from a table and discuss how to evaluate transferability into non-deterministic real world contexts.

KEYWORDS: Automation; Grasping; Robot dynamics; Control of robotic systems; Pose estimation and registration.

1. Introduction

For more than a decade, data-driven grasp planning approaches have been used in the research community and the main focus has been on how to online select good grasps from a grasp-database generated using heuristics or simulation.[1–5] Only little attention has been payed to the simulated context in which these databases were generated and how well the generated grasps actually perform in contexts which are different from the simulated context, e.g., grasps that have been generated in a free floating environment and then are applied to an object placed on a table.

Furthermore, object pose uncertainties due to model and sensor limitations may negatively influence the grasp execution. Pose uncertainties can be compensated for in the offline calculation of the grasps[6–8] by evaluating either the quality of neighboring grasps or the quality of the neighborhood grasp contacts, where a neighboring grasp can be simulated by applying a small displacement to the object before executing the target grasp in simulation. The basic idea is that a high average quality of grasps in the neighborhood of the target grasp reflects a high robustness toward uncertainties in the execution of the target grasp.

The above approach of calculating the robustness of a grasp target by evaluating the outcome of the neighborhood using small perturbations, assumes that the calculated neighborhood of

* Corresponding author. E-mail: jimali@mmmi.sdu.dk, jimmy.alison.jorgensen@gmail.com

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

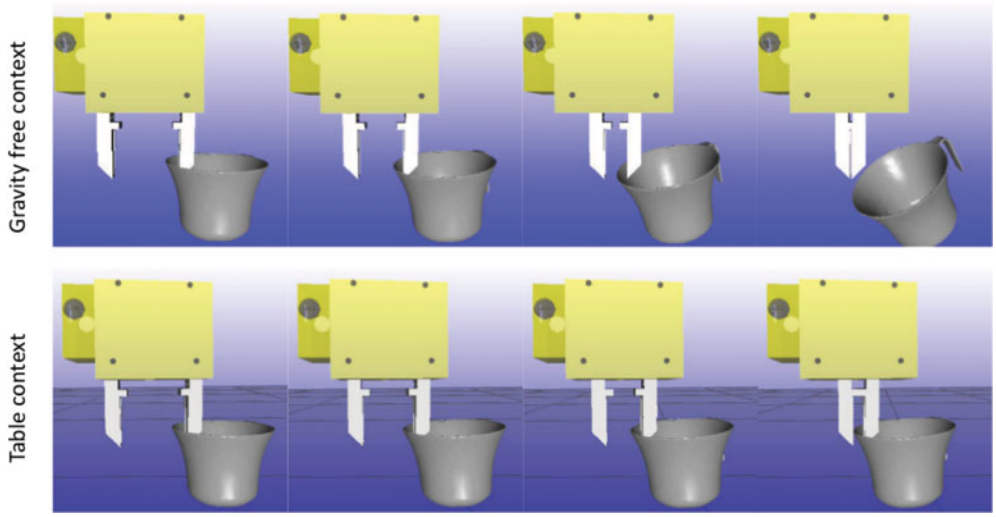


Fig. 1. Two simulated grasp sequences of the same grasp but in two different contexts. The top sequence fails to grasp the cup in the gravity free environment. In the bottom sequence the cup is successfully grasped when placed on top of a table.

grasps—whether done in simulation or with another heuristic—captures the same neighborhood as if the grasps were executed in the real world. This is problematic since the success of a grasp executed in the real world can be dependent on the environmental context, e.g., when the object is standing on a table, lying in water or leaning against a wall. This is illustrated in Fig. 1 where the same grasp is executed in two different environments, in one it succeeds and in the other it fails. It is obvious that changes in the environment can result in collisions between gripper and environment and therefore some grasps that are successful in one environment will fail in another due to collision. This specific form of grasp failure can be effectively handled during the online selection of a grasp by testing gripper geometry for collision against a virtual model or the sensor information available. Thus, changes in grasp success space due to collisions are not of great importance in respect to predicting grasp success outcome.

However, the interaction forces between the object to be grasped and the environment may differ substantially in different contexts, which may result in different outcomes of attempting the same grasp. The reason is that pose uncertainties imply that the fingers will not be synchronously placed at the target contacts on the object, which will cause the object to move during the grasp. The movement will be constrained by the environment and hence different environments or contexts may influence the success of a specific grasp attempt.

In this paper, we investigate the transferability of grasp successes from both a simulated context to another simulated context as well as the transfer from a simulated context to a real world context. Our motivation is not only to understand how the simulated context affects the success of a given grasp, but also to understand how the success neighborhood of a grasp is affected when transferred to a new context. The influence on the neighborhood when transferring a grasp is important for several reasons:

1. *Grasp quality estimation:* the neighborhood is used to calculate the robustness quality of a grasp, which describes how much pose uncertainties affect the success likelihood of a grasp.
2. *Continuous grasp representation:* some grasp database representations such as grasp densities[9] represent not a single successful grasp, but a complete continuous grasp success space which include the grasp neighborhood successes.
3. *Uncertainties in execution:* when executing a grasp from a grasp database, then the actual executed grasp will likely be a grasp from the neighborhood of the selected grasp due to pose uncertainties.

We investigate the transferability of the grasp neighborhood by randomly generating grasp configurations that are simulated in different environments. Comparing the outcome of hundreds of thousands of grasps simulated in different environments enables us to get insight into the nature

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

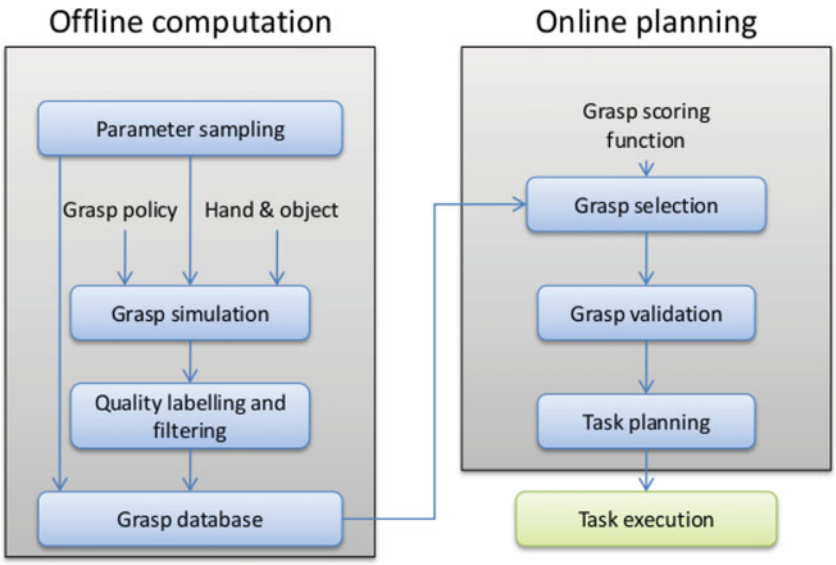


Fig. 2. Typical data-driven grasp planning approach.

of the transferability of a grasp database, which allows us to qualitatively show the importance of using context specific simulation when computing grasp databases for data-driven grasp planners. In addition, we will present a quality measure for the context transferability of a grasp database and use it to compute the transferability for several chosen objects and contexts.

We furthermore provide real world results that show that context-aware generated grasp databases have a better grasp success transferability/predictability than context-unaware databases. We discuss and present a quality measure for transferability to real world contexts which differ from simulated contexts by being non-deterministic and noisy.

In Section 2, we introduce related work within data-driven grasping and how it is commonly used in the community. Then in Section 3, we discuss transferability and how we can measure it, firstly when considering transfer between simulated contexts and secondly when considering transfer between simulated and real contexts. We introduce the setup, objects, grippers and control strategies in Section 4 which is followed by a description of how we compute the grasps in Section 5. In Section 6 the results are presented and a analyzed.

2. Related Work

State of the art grasp planning algorithms often rely on offline computed databases of feasible grasps generated using kinematic and/or dynamic simulation.[1–5] These grasp databases are used in online grasp/motion planning where the best grasp is chosen to be used in a specific context. The main steps in the generation of a grasp-database and its usage is outlined in Fig. 2. In the following, we first address the offline phase depicted on the left side of Fig. 2 followed by the online phase depicted to the right.

In the offline computation a sampling strategy is chosen to sample the parameter space of the system. This is defined by at least the pose of the gripper, the configuration of the gripper fingers (the grasp preshape) and the maximum joint torques. Additional parameters such as max finger velocity or the trajectory of the base of the gripper can also be used depending on the policy used to perform the actual grasping execution. The sampling can be selected arbitrarily, it can use grasp preshapes[10] and directly sample the base of the gripper in $SE(3)$ or use more biased sampling approaches that generate or plan samples based on the object geometry and/or gripper kinematics.[11,12]

The sampled parameters are used to initiate the grasp simulation which consist of the gripper and an object. When a single grasp has been generated, the simulator is started and a grasping policy is used to control the fingers into a grasp. This policy is typically simple and often dependent on which type of simulation is used e.g., kinematic or dynamic.

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

The simulation may be terminated due to a number of different criteria, which themselves involve the determination of whether the grasp was successful or not. In a kinematic simulation, analytic stability measures based on the Grasp Wrench Space (GWS),[13] are computed based on contact points and used to describe if a grasp is stable. These measures also apply to dynamic simulations, but here the stability might also be determined by performing a lifting motion with the gripper and object as in ref. [14]. If the object is not dropped after moving the gripper, then a stable grasp has been computed. Notice, that in a dynamic simulation, the object may be lost during lifting due to forces from the environment, e.g., gravity or contacts with other objects.

After simulating sampled grasps the grasps are labeled and filtered before they are added to the grasp database, see bottom left in Fig. 2. Intuitively, only successful grasps of a certain quality should be added to the grasp database. However, for some applications it is important to include low quality grasps in order to maintain enough coverage of the object.[15] The grasp quality might be based on a grasp wrench space (GWS) analysis,[13] on tactile based quality metrics[16] or any other form of quality measure[17] that can be computed offline. The main importance is to store grasps that will maximize the likelihood of success when used in an actual application.

In the online computation (the right block in Fig. 2), the context is changed. We have a robot with a limited reach, we have obstacles such as the supporting table or virtual constraints that define the workspace which limit the grasping possibilities, sensors that are imperfect and other constraints imposed by the task, that we wish the robot to perform. The main problem is then to select a grasp from the grasp database which is feasible in spite of these limitations and constraints. In general the online selection of a grasp from a grasp database should maximize the likelihood of success. In practice, grasps with highest offline computed qualities are typically used. A more advanced online grasp selection approach was presented in ref. [2], were an online scoring function that favors grasp configurations that are furthest away from obstacles was used. The scoring function reduced unwanted collisions with other objects in the environment enabling safer grasping in cluttered environments.

Several learning approaches to grasping[18,19] use simulated grasp databases as training data. The main goal is typically to use learning to infer a feasible grasp when presented with some form of input, which could be point cloud data, images, or simply the geometry of an object.

In ref. [4], a grasp database "the Columbia Grasp Database" is presented and it is used for grasp planning in ref. [20]. The database is generated using kinematic simulation of grasps in an obstacle free environment and the grasp configurations where calculated by the authors' eigen-based grasp planner.

In ref. [1], the grasp database is generated based on a kinematic simulation of closing the gripper fingers around the object in an obstacle free environment. The authors do not comment on the context independence of the database but point out that a gripper with tactile feedback is necessary to execute grasps generated with their approach. In ref. [2], the same approach is used to generate the grasp database, but here the authors do not mention the use of tactile feedback in their experimental setup.

The use of tactile sensors to online guide the grasp stability has been investigated in several papers[21–23] and shows great promise. However, the performance of tactile sensor hardware is far from ideal and typical issues are drift, low sensitivity range, low durability, detection of sliding, detection of measuring normal and measuring sheer forces. Reliable and general reactive grasping therefore still remains a challenge.

However, all above approaches rely on databases generated in a context independent simulation. To rely on the quality of those grasps it is necessary to have either a perfectly calibrated setup with no uncertainties or an online grasp execution that rely on tactile feedback to correct for uncertainties. Both of these properties in a setup are non-trivial and the latter is still not solved and is actively researched.

In this work we use the dynamic rigid body grasp simulator in RobWork.[24] For the simulated-to-simulated transfer experiments done in this paper, no quality other than success and failure is used. These are defined by trying to move the object after it is grasped. If the grasp can withstand the accelerations during movement then the grasp is successful.

This paper is an extension on the conference paper[25] presented at the conference RAAD13. The original paper addressed the issues concerning the transfer of grasps that are computed in contexts that differ from the contexts they are used in. Simulated experiments where provided to document this transfer issue. In this paper we have extended the simulated experiments and provided results on transferring grasps from two different simulated contexts to a real world context. We furthermore

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

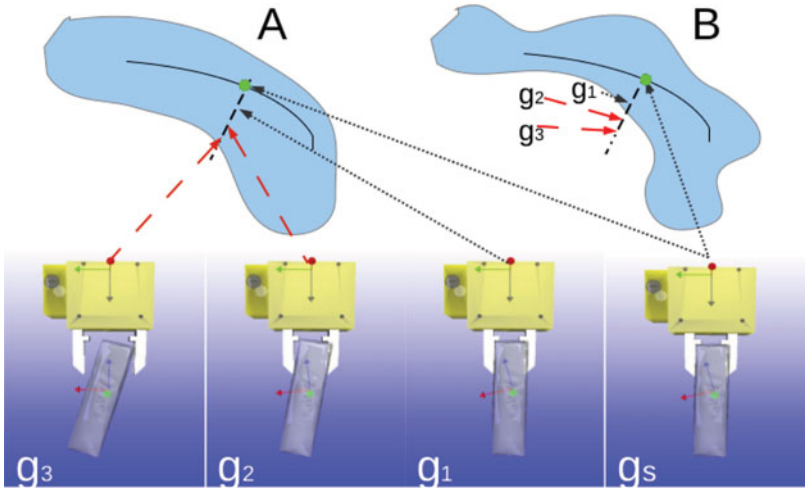


Fig. 3. Grasp sequence demonstrating the difference between successful grasp configurations (the blue area) and stable grasp configurations (grasps on the black area). The two blue areas, depict the artificial grasp success space in two different environments A and B. The grasp sequence demonstrates the convergence of a successful grasp configuration into a stable grasp configuration in environment A. The success space of environment B suggests that grasps g_3 and g_2 should result in failure in environment B and are therefore not transferable from A to B.

show that context-aware generated grasp databases have a significantly better transferability than context-unaware databases.

3. On Quantifying Transferability

In this section we discuss how to quantify transferability both from simulated to simulated contexts and from simulated to real world contexts. Simulation to simulation quantification is simpler due to ground truth knowledge of the virtual world and the deterministic way of simulation. In Section 3.1 we define a measure that exploits the deterministic behavior of simulation to evaluate the same grasps in two different contexts. Quantifying transferability from simulation to real world is more difficult due to the non-deterministic way of sensor and process noise in the real world. Hence, in Section 3.2 we describe a quantification method that is based on a continuous success evaluation that use the outcome statistics generated from executing the same grasp multiple times.

3.1. Quantifying transferability in simulation

The transferability measure that we present in this section requires a set of grasps that have been executed in multiple contexts. The binary outcomes (success, failure) of these grasp experiments then define the transferability.

This is illustrated in Fig. 3 where the neighborhood (blue area) of successful grasp experiments in one context A is different from the neighborhood of successful grasp experiments in another context B. The green point illustrates the target grasp g_s which belongs to the set of stable grasps G_{stable} (black line/area) and the arrows illustrate three grasps (g_1, g_2, g_3) chosen from the neighborhood of g_s. In context B grasps (g_2, g_3) are no longer part of the neighborhood of successful grasps (indicated as red arrows) and therefore fail. We define the transfer of a grasp from one context to another to be successful in the case of g_1, but unsuccessful in the case of (g_2, g_3).

If we think of transferability as a measure that determines how well we can predict outcomes of grasp experiments in one context, given the knowledge of the outcomes of the same experiments from another context, then we can borrow metrics from the machine learning literature and pose the measure of transferability as a classification problem. If we consider the execution of the same grasp in two different contexts then we may have four possible outcomes: (success, success), (failure, success), (success, failure) and (failure, failure). These are illustrated in a confusion matrix in Table I as true positive (TP), false negative (FN), false positive (FP) and true negative (TN). In the above example g_1 thus belongs to TP and g_2, g_3 belong to FP.

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

Table I. A confusion matrix element that illustrates the outcome of transfer between two contexts A and B. High values in TP and TN indicate high transferability of grasp results from one scenario to the other.

		Results in A	
		success	failure
Results in B	success	TP = 74 (true positives), the number of grasp simulations that where successful in both scenarios	FN = 113 (false negatives), the number of grasp simulations that failed in A but succeeded in B
	failure	FP = 539 (false positives), the number of grasp simulations that succeeded in A but failed in B	TN = 2507 (true negatives), the number of grasp simulations that failed in both A and B

The probability that a selected successful grasp from A will also be successful in B is then $TP/(FP + TP)$.

We use the Matthews correlation coefficient (MCC)[26] as the transferability measure. It is based solely on the confusion matrix and produces a value between -1 and 1, where 1 indicate perfect prediction, 0 indicate a random prediction and -1 indicate an inverse prediction of a binary classification. It can be calculated directly from the confusion matrix as in Eq. (1).

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{1}$$

3.2. Quantifying transferability in real experiments

As discussed in the previous section we think of transferability as a measure that determines how well we can predict the outcome of a grasp executed in one context, based on the knowledge of the outcome of a grasp executed in another typically simulated context. In the real world the outcome of performing the same grasp may vary due to uncertainties in the system, eg. sensor data noise, calibration, lighting conditions, etc.. Instead of directly comparing outcomes as was done in the transferring between simulations, we compare the success probability generated from outcomes of executing the same grasp. In the real world we can compute the success probability directly from the outcome of multiple executions of the same grasp. In simulation we can also evaluate the success probability but we need to model the uncertainties. This is done by perturbing each grasp target slightly where the perturbation is drawn from a distribution that models the uncertainty. Finally we can evaluate the grasp transferability by calculating the correlation between simulated success probability and real world success probability.

There exist several grasp quality metrics that describe the quality of a grasp as a continuous value. Some try to estimate success probability of a grasp by evaluating successes or quality over the neighborhood of grasps,[6,7,15] others use heuristics typically based on grasp stability in relation to the grasp wrench space[13,17] to determine grasp quality. Common for all quality metrics should be that a higher grasp quality correlates with a higher success probability. The Spearman rank correlation coefficient is capable of capturing this relationship between two variables. It essentially measures if the increase of the value on one variable does also reflect an increase (any increase) in the other variable. It is defined as the Pearson correlation coefficients over the ranked variables:

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}} \tag{2}$$

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press



Fig. 4. Object from the KIT object database. From left to right: corny object, cup-object and tomato object.

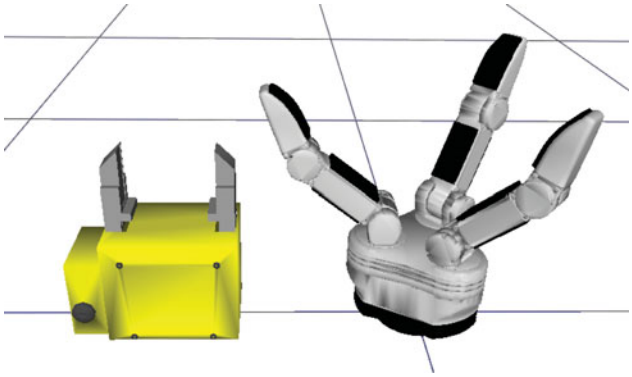


Fig. 5. PG 70 gripper (left) and the SDH-2 (right).

where x_i, y_i are the ranked values and $\bar{x}$, $\bar{y}$ are the mean of the rank variables. If there are no correlation then ρ yields 0. If there is a increasing monotonic trend in the data then ρ yields 1 and when there is a decreasing monotonic trend then ρ yields -1.

For quality metrics that directly estimate success probability, we would expect a linear relationship between simulated grasp quality and the real world success probability. It would therefore be more accurate to use the Pearson correlation coefficient in such scenarios since this captures the linear dependency between two variables. The Pearson correlation coefficient is defined in Eq. (2) but with x_i, y_i as sampled values and $\bar{x}$, $\bar{y}$ as the sampled means.

4. System Description

In this section we present the objects, grippers and the different simulated and real contexts used in this paper.

4.1. Objects and grippers

Three objects (see Fig. 4) were selected from the KIT Object-Models Web Database[1]. The objects are common household objects which are sufficiently different to provide interesting comparisons.

The grippers used to grasp these objects are the Schunk parallel gripper (PG 70) and the Schunk Dexterous Hand (SDH-2), which are both shown in Fig. 5. The PG 70 gripper has two fingers coupled into one Degree Of Freedom (DOF), that is, one DOF moves both fingers. The fingers can move up to 6.8 cm apart and the contact surface is approximately 2×3 cm and covered by rubber.

The SDH-2 is a 3-fingered dexterous hand with two DOF per finger and one coupled DOF to control the base rotation of two of the three fingers. The SDH-2 has six contact surfaces covered with rubber, each of them measuring approximately 2×3 cm. However, for precision grasps only the three contact surfaces on the distal joints are normally used.

For the PG 70, we choose only one preshape with the maximum distance of 6.8 cm between its jaws. Four preshapes were chosen for the SDH-2 which are shown in Fig. 6. These different preshapes enable different grasping options, and as such are important when characterizing the grasp affordances of the gripper.

[1] http://wwwiaim.ira.uka.de/ObjectModels

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

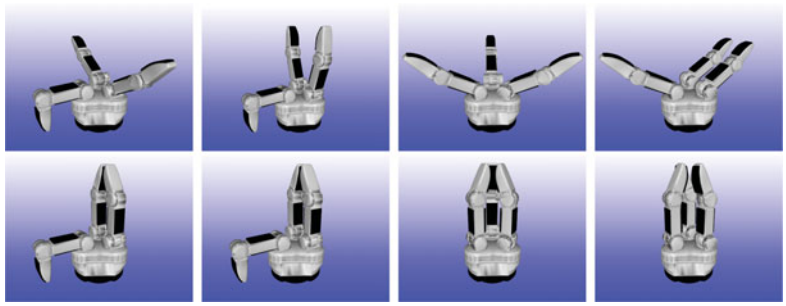


Fig. 6. Four preshapes used with the SDH-2 dexterous hand. From left to right: c_{par}, $c_{parsmall}$, c_{ball}, c_{cyl}. The top row is the initial preshape before grasping, the bottom row is the target configuration that the hand will move to during grasping.

We use preshapes because:

- Using preshapes is a simple and direct way of providing multiple ways of grasping an object. The grasping process reduces to the sequence: *(open hand, move towards object, close hand)*, which is already supported by the software of the SDH-2.
- Preshapes do not require additional parametrization. This property is important when sampling random grasps since the dimensionality of the search space is not extended by any gripper parameters such as the individual joint configurations.

4.2. Contexts

The context describe the environment and task in which grasps are being executed. We have evaluated grasps in four different contexts where the first three are simulated environments and the last is a real environment:

- *float* - the object is placed in a simulated gravity free environment with no obstacles.
- *table* - a simulated environment with gravity is used and the object is placed on a planar surface in one of its canonical poses.
- *full* - the object is placed randomly on a table in an environment that closely models the real environment and task. Compared to the table context, the full context also include simulated sensor data, the robot arm and the sensor processing (pose estimation) that is also performed in the real context.
- *real* - the object is placed randomly on a table in a real world environment and picked up by a robot with a dexterous gripper.

5. Computation of Grasp Affordances

We compute complete grasp affordances by evaluating randomly sampled gripper poses in the neighborhood of an object. Each of these sampled poses, combined with a preshape of the gripper, represents a grasp hypothesis, which can be evaluated in simulation. The outcome of the simulation may be one of $\{$*success, failure or collision*$\}$, where *success* represents a successful grasp of the object (indicated by the fact that the gripper is still in contact with the object after grasping it) and *failure* represents a grasp where the gripper has no contact with the object after trying to grasp or lift it; *collision* represents a grasp, where the gripper is in collision with the object or the environment in the initial state of the simulation.

The evaluation of a grasp hypothesis is performed using a dynamics grasp simulator from RobWork.[24] This simulator is based on the Open Dynamics physics engine where the contact resolution and generation mechanism have been modified to improve accuracy and robustness when simulating complex triangle meshes. The main simulation process of a single grasp is:

1. Set the initial scene configuration, e.g., gravity, friction, poses of obstacles and objects.
2. Place the gripper in a sampled pose relative to the object and set the gripper configuration to one of the preshapes.
3. Test if the gripper collides with object or environment.

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

4. Start the simulation and set the target gripper configuration (preshape dependent) of the gripper controller.
5. If a grasp is obtained, lift the object and compute the success criteria.

We shall now discuss these steps in more detail.

5.1. Initial scene configuration

The initial scene configuration for generating the grasp-databases use a free floating environment, in which no gravity or obstacles are present. In such an environment the object can freely slide into or out of a grasp.

To investigate the transfer of grasp affordances of an object from the free floating environment we repeat the grasp simulation, but with the object placed on a table. Two canonical poses of each object are used to create two different table environments per object.

5.2. Object specific sampling strategy

The sampling of the gripper configuration and the choice of sampling strategy necessarily influence the resulting set of grasp affordances and the overall success probability. Typically, the primary goal of grasp planning is to maximize the grasp success probability by exploiting knowledge of gripper, object and environment. This tends to generate grasp databases that only represent a small subset of the complete set of successful grasps.

In this work, we need an approximation of the complete set of grasp affordances that are possible on a specific object. Thus, an unbiased sampling strategy that explores $SE(3)$ fully is preferred. Uniform random sampling in $SE(3)$ would therefore be ideal but it is also impractical because of the large number of simulations necessary to cover $SE(3)$. Instead a sampling strategy that is biased toward the object geometry is used. This effectively reduces the number of required simulations without reducing the success space too much.

Beside geometric models of the object the sampling also requires the approach vector of the gripper to be placed in the same direction as the positive z axis of the gripper Tool Center Point (TCP) frame. The sampling effectively encapsulates the idea that the gripper needs to point toward some part of the object geometry before it is able to successfully grasp the object.

First a random point $\mathbf{p}$ is selected from a uniform distribution over the surface of the object. Then an orientation $\mathbf{R}$ is selected from a uniform distribution in $SO(3)$ and used to define the temporary target pose $(\mathbf{p}, \mathbf{R})$. The pose is then translated along the z axis by a randomly generated value d in the interval $[-0.04\,\text{m}; 0.04\,\text{m}]$. 4 cm was determined from the size of the fingers of the biggest gripper. If the tool center point was placed more than 4 cm away from the triangle surface then the chance of generating a valid grasp configuration is very low. Hence, a larger value would only increase the computation time without increasing the number of generated grasps. The final pose is therefore:

$$\mathbf{T}_{\mathbf{pose}} = (\mathbf{p} - (\mathbf{R} \cdot [0, 0, 1]^T) \cdot d, \mathbf{R}) \tag{3}$$

5.3. Collision filtering and labeling

The grasp hypothesis sampled in the previous step are generated using a heuristic which does not include collision testing. It is possible that the pre-grasps pose of a generated target is actually in collision with the object or the environment. If the gripper collides with the object in its pre-grasp pose then it will do so in all contexts since the object to gripper pose is the same for all contexts. Hence, grasp hypotheses where the gripper collides with the object in its pre-grasp pose are discarded. However, if the gripper collides with the environment we cannot discard the grasp because the environment may change between contexts. Hence, we add the environment colliding grasp hypotheses to the database but label them *Colliding*.

When generating a grasp database for a real application traditionally only high quality grasps and non-colliding grasps are added to the database. However, for this work we need to keep track of colliding grasps in order to avoid comparing grasps that collide in one context but are successful in another. In effect, when computing the transfer statistics between two databases A and B then all grasps that are labeled colliding in either will be left out of the computation.

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

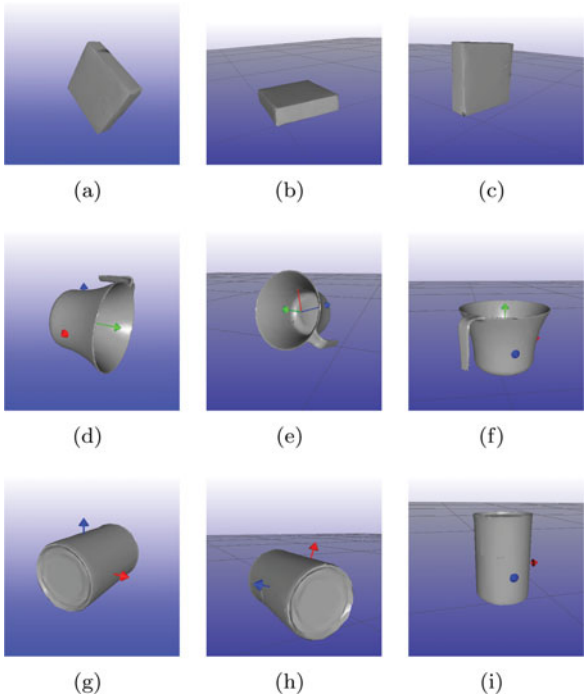


Fig. 7. The floating pose is illustrated on the left, followed by two canonical poses on the table (side and upright). Top row: Corny object. Middle row: Cup object. Bottom row: Tomato object.

5.4. Grasp simulation

The final step before adding a non-colliding grasp to the database is the grasp simulation. Both pose and configuration of the gripper have been sampled and no collisions with environment or object were detected. The simulation is initialized with the scene configuration and the sampled parameters. When the simulation starts, a penalty based grasp controller guides the fingers toward a closing configuration. Closing configurations for the four preshapes of the SDH-2 are shown in Fig. 6. A grasp is successful if the object gets caught between the fingers and if it stays there while the hand is moved 10 cm with a maximum acceleration of 3 m/s^2 in the positive z-axis of the world coordinate frame (gravity works in the direction of the negative z-axis). This is termed a lifting operation, which only makes sense in the table environments where the objects need to be lifted free of the table.

6. Experiments

In this section, we first provide experimental results on the transfer between two simulated contexts in Section 6.1 and discuss it in Section 6.2. Secondly, we provide results on the transfer from simulated contexts to a real world context in Section 6.3 and discuss these in Section 6.4.

6.1. Simulation results

We provide simulation results in terms of complete sets of grasp affordances for three objects (see Fig. 4) using the PG 70 parallel gripper and the SDH-2 hand in two different contexts, namely the rather artificial free-floating context shown in the left column of Fig. 7 and the more application-oriented context, where the object is placed on a table (see the middle and right column of Fig. 7). The table environments have a gravity of 9.81 m/s^2 and the viscous Coulomb friction between table and object is set to 0.3 $N/(m/s)$. All simulated environments for all objects are illustrated in Fig. 7.

Figure 8 illustrates the successful outcomes of the same 5000 grasp hypotheses in the three different contexts of the corny object grasped with the SDH-2 using the preshape c_{par}. It is obvious that the added table constraint significantly reduces the number of successful grasps. However, it is not obvious if successes from the constrained environments (center and right image) will also be successes in the floating environment. In the following, we use the confusion matrices introduced

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

Table II. Success percentages of the simulated outcomes. In each major column, the left sub-column shows the percentages of success if collisions are not included, and the right column shows the success percentages if collisions are included.

		SDH-2				PG 70
		c_{par}	$c_{parsmall}$	c_{ball}	c_{cyl}	c_0
o_{corny}	s_{float}	43.1% 25.0%	45.5% 5.2%	70.3% 38.0%	52.9% 28.0%	12.3% 0.7%
	s_{side}	5.8% 0.4%	0.0% 0.0%	5.8% 0.2%	5.3% 0.3%	0.1% 0.0%
	$s_{upright}$	42.1% 7.2%	41.5% 1.9%	70.9% 9.3%	51.9% 8.9%	14.4% 0.4%
o_{cup}	s_{float}	54.5% 42.6%	47.7% 6.0%	79.8% 61.1%	61.7% 45.7%	41.5% 3.7%
	s_{side}	47.0% 5.8%	48.4% 1.7%	72.5% 5.4%	47.3% 6.0%	44.0% 1.6%
	$s_{upright}$	45.0% 4.7%	65.4% 2.3%	81.9% 4.9%	47.7% 4.9%	70.4% 3.5%
o_{tomato}	s_{float}	48.5% 24.9%	30.4% 5.5%	84.9% 72.0%	72.3% 59.9%	3.8% 0.3%
	s_{side}	22.7% 2.3%	10.9% 0.3%	34.4% 1.9%	32.8% 3.4%	2.4% 0.1%
	$s_{upright}$	43.4% 6.6%	23.1% 1.4%	75.3% 7.5%	63.1% 10.4%	17.2% 0.6%

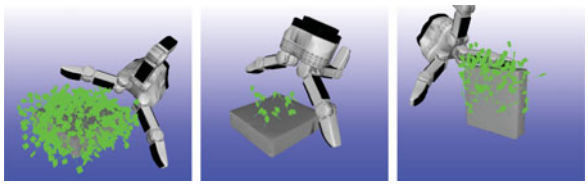


Fig. 8. The successes of grasping the Corny object in the same 5000 grasp simulations but performed in different contexts. From left to right: free-float, on table (side pose) and on table (upright pose).

in Section 3.1 to evaluate how well successes and failures in the floating simulations transfer to successes and failures in the table environments, and vice versa.

Multiple datasets were generated for each gripper. For the SDH-2 the datasets are characterized by a triple (o_i, s_j, c_k), where o_i is the object, s_j is the specific scene (free floating or on table with different poses), and c_k is the grasp strategy, which includes the number of fingers and the preshape used. The parallel gripper is simpler, and only one grasp strategy is used. Hence we describe datasets generated with the PG 70 by a pair (o_i, s_j).

For each floating environment experiment (o_i, s_{float}, c_k) 100.000 grasp simulations[2] were generated using the sampling approach presented in Section 5.2. The overall success probabilities of the simulations are available in Table II. These indicate the size of the success space of the grasp datasets in the individual contexts. Two success probabilities are given. The first shows the percentage of grasp successes from all grasps that were not initially in collision, the second shows the percentage of grasp successes from all grasps including the colliding ones. The latter is only interesting for evaluating how many grasps out of the database can actually be applied. It cannot be used to state anything on transferability since collisions are easily predicted online using collision detection between gripper geometry and a virtual model of the environment or sensor data.

In the same table the success probabilities of the simulations of the table environments (s_{side}, s_{uprigt}) are also shown. These simulations use the same grasp hypotheses as in the floating environment, but performed in the specific contexts (e.g., the object was placed upright on the table).

That the exact same grasps have been executed in the different contexts enables us to calculate the transferability measure presented in Section 3.1. Table III show the confusion matrices from which the transferability is calculated and the transferability quality of each confusion matrix is presented in Table IV.

6.2. Analysis of simulation results

The MCC metric is used directly on the confusion matrices of Table III (see Table I for explanation). Note that in Table III collisions with the table are filtered away. The results are shown in the rows of Table IV.

[2] For the tomato object the actual number of samples is slightly lower.

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

Table III. Confusion matrices of successes and failures from floating environment and the specific table environment (s_{side}, $s_{upright}$). See Table I for a detailed explanation of a single cell. The cells are colored in four gray scales to indicate the quality in terms of MCC. The lighter the gray color the higher the MCC value and the better the transfer.

		SDH-2				PG 70
		c_{par}	$c_{parsmall}$	c_{ball}	c_{cyl}	c_0
o_{corny}	s_{side}	159 221 933 5298	0 0 1 1495	74 113 539 2507	97 222 721 4991	0 1 0 1499
	$s_{upright}$	6271 861 1660 8183	1777 100 284 2379	8586 761 1127 2712	6296 830 1587 5017	327 46 34 2182
o_{cup}	s_{side}	4876 916 1903 4621	1548 126 241 1546	1459 407 303 406	893 314 474 871	1568 81 175 1924
	$s_{upright}$	3483 1219 1683 4059	2093 207 155 1061	1574 657 195 297	917 540 449 1146	3119 396 78 1399
o_{tomato}	s_{side}	1019 228 1310 2923	246 31 214 2045	445 112 570 494	1063 278 1212 1531	30 35 43 2568
	$s_{upright}$	1518 119 538 1591	547 82 219 1874	2167 91 261 477	2019 112 607 602	113 434 53 2584

Table IV. Quality estimates based on the MCC correlation coefficient of the confusion matrices in Table III. The cells have been colored in four different gray scales to visually highlight the values. The lighter the color the better the quality.

		SDH-2				PG 70
		c_{par}	$c_{parsmall}$	c_{ball}	c_{cyl}	c_0
o_{corny}	$s_{side,MCC}$	0.17	-	0.13	0.12	-
	$s_{upright,MCC}$	0.70	0.83	0.64	0.65	0.87
o_{cup}	$s_{side,MCC}$	0.55	0.79	0.34	0.39	0.86
	$s_{upright,MCC}$	0.45	0.78	0.25	0.35	0.79
o_{tomato}	$s_{side,MCC}$	0.43	0.64	0.26	0.33	0.42
	$s_{upright,MCC}$	0.67	0.72	0.67	0.52	0.32

The results in general indicate that the quality of the prediction depends highly on both the gripper and the object/context that is used. In Table IV the cells have been colored in four gray scales to indicate the value of the respective qualities.

In certain contexts (C_{par}, $O_{tomato,side}$), (C_{par}, $O_{Corny,side}$) and (C_{ball}, $O_{Corny,side}$) the prediction is so bad that the chances of choosing a successful grasp from s_{float} which will also be successful in s_1 is less than 20 %.

The first row $o_{corny,side}$ clearly have the worst transfer rate across all grippers. Inspecting the same row of confusion matrices in Table III we find that the most significant cause of this lies in grasps being successful in the float and not in the table context combined with the fact that only few grasps are successes at all (this is also indicated in center image in Fig. 8 and in success percentages in second row Table II). The low number of successes suggests that the particular grasping problem is difficult in which case only small context changes may influence the outcome.

The second row $o_{corny,side}$ has the best transfer rate across all grippers. Considering that the upright box on a table is an easy grasping problem and also that it is very similar to grasping the box in float context this outcome is expected.

The second column on the other hand, indicates that a specific gripper pre-shape has a higher transfer rate across all contexts than the others. The interesting part being that the gripper configuration $c_{parsmall}$ is the same as c_{par} only with a slightly more closed pre-shape. This suggests that pre-shapes that require large movements of the fingers and therefore potentially of the object are causing worse transfer rates, which is not counter intuitive since context differences impose different constraints on the movement of an object.

It is important to mention that in certain contexts quite a number of grasps that are not successful in the free-floating environment are successful in the constrained environment. This implies that in

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

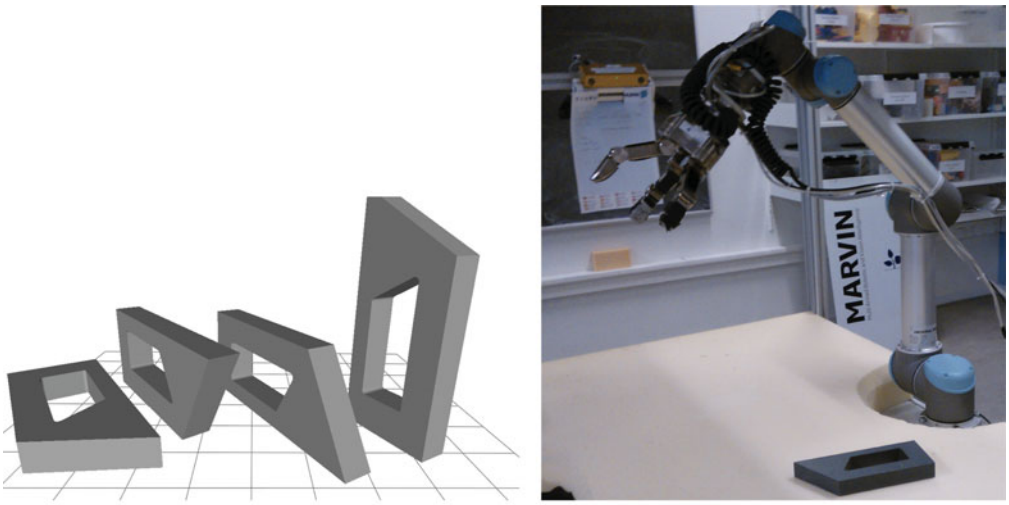


Fig. 9. To the left the Rump object used in the real world experiments. To the right the UR-6-85-5-A robot arm mounted with a Schunk Dexterous Hand (SDH). The object in the lower right corner is the Rump object.

certain scenarios, context-specific control strategies need to be taken into account in order to actually generate relevant grasp targets. For example, the movement of a flat object on a table when touched by one of the fingers might be utilized in the grasping process (see Fig. 1). These context-specific constraints cannot be accounted for in a free-floating scenario and need to be learned in the specific context.

In summary, the results show that difficult grasping problems seem to transfer poorly between contexts and that higher transfer quality for grasping problems in general may be obtained by avoiding large movements of the fingers eg. generating pre-shapes that are close to the surface of the object. Furthermore, it seems that some problems like grasping the corny object in an upright pose ($o_{corny,upright}$) are quite similar even across contexts, which suggests that one should use the expected grasping problem in his choice of context for generating grasp databases whenever possible.

6.3. Real world experiments

The real world experiments were performed using the *Rump* object illustrated in Fig. (9a). The *Rump* object weighs approximately 600 grams and was designed as a typical industrial object with a size to match the SDH gripper.

The grasping setup is illustrated in Fig. 9(b) where the SDH gripper is mounted on a universal robot arm UR-6-85-5-A that has a reach of approximately 1 m and a payload of 5 kg. In-between gripper and robot end-effector a Schunk 6D force torque sensor is mounted that is used for collision detection.

A single grasp database consisting of 434 grasps of the rump object using the SDH was used for the real world experiments. These grasps were generated using a state of the art filtering method[15] which not only keeps high quality grasps but also keeps grasps that increase the number of directions that the object can be reached from. For each grasp in the database, three different simulated quality metrics were computed. One used the radius of the largest sphere placed at the origin of the Grasp Wrench Space (GWS) and fully contained in the GWS.[13] The other two were based on directly computing the grasp success probability in two different contexts, namely a floating context and the full context.

The grasp database was used in a repetitive process that first located the Rump object on the table using a CAD based object localization algorithm. The localization used RGB-D information captured from a Microsoft Kinect camera which was converted into a point cloud and thereafter all points below and all points above 50 cm the largest plane was filtered away. The remaining points were clustered and the Iterative Closest Point (ICP) algorithm was used to fit the CAD model of the Rump to each cluster. The best fit was assumed to be the correct match.

After detecting the pose of the object, a suitable grasp from the database was chosen to be executed. If grasping was successful, then a new random drop pose above the table was calculated and the robot

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

Table V. Correlation between simulated grasp quality in two different contexts and the actual success probability achieved in real world use of the grasps.

	Float context	Full context
Pearson	0.30	0.67
Spearman	0.18	0.41

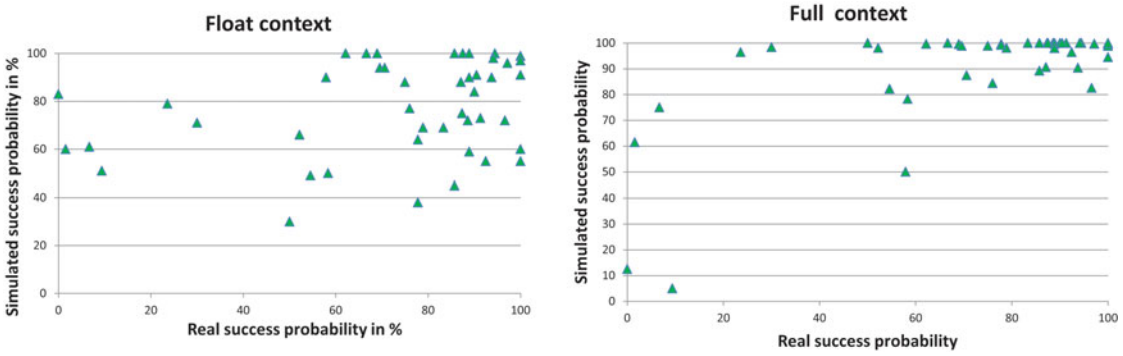


Fig. 10. (left) Correlation between float based robustness quality and success probability in the real world. (right) Correlation between success probability achieved in a simulation of the full context and the success probability in the real world.

executed motions to reach the drop pose. Finally, the object was released. If the grasp was unsuccessful then another grasp was selected for execution. If all grasp attempts for a specific object pose failed, a user intervened. It should be noted that only grasps that were kinematically feasible was selected. Meaning that for a grasp hypothesis to be selected it should have valid inverse kinematics (reachable), it should not collide with object or environment and there should exist a collision free path to reach the grasping configuration and to lift the object up from the table.

For each grasp execution attempt, the outcome (success/failed) was saved and used to later compute a success probability of each grasp. These success probabilities were then compared and correlated with the three different quality metrics.

A total of 2100 grasp experiments were performed in the real setup using 164 grasps out of the 434 grasps in the database. From these 164 grasps only 44 grasps have been used in more than 13 experiments. The grasp used most often had been used 106 times. Our correlations are based on these 44 grasps and the data is visualized in the graphs in Fig. 10.

The transferability measure based on the Spearman and Pearson correlation coefficients have also been computed for these two datasets. These are available in the Table V. The table shows that the full context generated success probabilities correlate significantly better than the float based method, measured with both Pearson and Spearman correlation coefficients.

6.4. Analysis of real world results

The correlation results in Table V show that the "full context" simulation has the best transferability. This was expected since the "full context" should model the real setup to a larger detail compared to the "float context".

Figure 10(right) provides further insight into the transferability for the "full context" where there is a tendency that samples are located in the upper left area of the correlation space, which suggest a systematic overestimation of success probability. The result of the float context is shown in Fig. 10(left), where there is a tendency in regard to over estimation for grasps with a very low success probability.

There are several factors that may cause the overestimation. Primarily the overestimation seem to stem from model parameter inaccuracies. In our case, the choice of friction had a significant impact on the results. High friction coefficients resulted in overestimation of success probability whereas low friction coefficients resulted in underestimated of success probability. This is not surprising

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

considering that a higher friction will allow a grasp to withstand a larger force. However, the friction coefficients of the real setup changed over time due to wear, heat and dirt, which we were not able to model in the simulation.

Secondly, the uncertainty model used in the evaluation of the simulated success probability may not accurately capture the accumulated uncertainties of the real system.

Both of these issues effects the simulated results. However, the "full context" simulation still performs better in terms of correlation with the real world, which clearly indicates the importance of context aware grasp databases.

7. Conclusion

Data-driven grasp planning has become increasingly popular and some grasp databases are available for public download.[3,27] When using a grasp target g_s from such a database, in systems with uncertainties, then it is most probable that the executed grasp will lie in the neighborhood of g_s. Hence, it is important to incorporate the transferability of the neighborhood of a grasp when using grasp database.

Moreover, in this paper we showed that it is quite likely that grasps performed on objects in different contexts have different outcomes. Hence, one should take care when using a grasp database in a context other than the context in which it was generated.

We specifically investigated the transfer between an unconstrained, free-floating environment and more constrained environments, which is important for applying learned grasp knowledge in novel contexts. For several different simulated contexts, we showed that the grasp success likelihood depends highly on the context that the object is embedded in and we used the Mathews correlation metric to evaluate how good this transfer is.

We also discussed how to evaluate transferability between simulated and real contexts and computed the transferability between two different simulated contexts to a real world context using a state of the art grasp database. These results should motivate the use of the detailed context in the simulation of grasps.

Further investigations on grasp transfer could include more real world experiments and a deeper analysis on the discrepancies between simulation and real world. Also evaluating the transferability of state of the art grasp quality metrics for different contexts, grippers and objects could provide further insights in how to choose a suitable grasp quality metric.

Acknowledgment

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. 270273, Xperience.

References

1. A. Morales, T. Asfour, P. Azad, S. Knoop and R. Dillmann, "Integrated Grasp Planning and Visual Object Localization for a Humanoid Robot with Five-Fingered Hands," *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems* (Oct. 2006) pp. 5663–5668.
2. D. Berenson, R. Diankov, K. Nishiwaki, S. Kagami and J. Kuffner, "Grasp Planning in Complex Scenes," *2007 7th IEEE-RAS International Conference on Humanoid Robots* (29 2007–dec. 1 2007) pp. 42–48.
3. C. Goldfeder, M. Ciocarlie, J. Peretzman, H. Dang and P. K. Allen, "Data-Driven Grasping with Partial Sensor Data," *IROS 2009. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009* (Oct. 2009) pp. 1278–1283.
4. C. Goldfeder, M. T. Ciocarlie, H. Dang and P. K. Allen, "The Columbia Grasp Database," *ICRA*, IEEE (2009) pp. 1710–1716.
5. C. Goldfeder and P. K. Allen, "Data-driven grasping," *Auton. Robots* **31**(1), 1–20 (2011).
6. J. Kim, K. Iwamoto, J. J. Kuffner, Y. Ota, and N. S. Pollard, "Physically-Based Grasp Quality Evaluation Under Uncertainty," *2012 IEEE International Conference on Robotics and Automation (ICRA)* (2012) pp. 3258–3263.
7. J. Weisz and P. K. Allen, "Pose error robust grasping from contact wrench space metrics," *ICRA*, IEEE (2012) pp. 557–562.

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press

8. H. Kruger, E. Rimon and A. F. van der Stappen, "Local Force Closure," *2012 IEEE International Conference on Robotics and Automation (ICRA)* (May 2012) pp. 4176 –4182.
9. R. Detry, D. Kraft, O. Kroemer, L. Bodenhagen, J. Peters, N. Krüger and J. Piater, "Learning grasp affordance densities," *Paladyn* **2**(1), 1–17 (June 2011).
10. D. Wren and R. B. Fisher, "Planning dextrous hand precision grasps from range data, using preshaping and finger trajectories," *Proceedings of the Symposium Intelligent Robotic Systems (SIRS95)* (1995).
11. K. Huebner, S. Ruthotto and D. Kragic, "Minimum Volume Bounding Box Decomposition for Shape Approximation in Robot Grasping," *IEEE International Conference on Robotics and Automation, 2008. ICRA 2008* (2008) pp. 1628–1633.
12. A. T. Miller, S. Knoop, H. I. Christensen and P. K. Allen, "Automatic Grasp Planning Using Shape Primitives," *Proceedings of the ICRA'03. IEEE International Conference on Robotics and Automation, 2003*, vol. 2 (2003), pp. 1824–1829.
13. C. Ferrari and J. Canny, "Planning Optimal Grasps," *Proceedings of the 1992 IEEE International Conference on Robotics and Automation, 1992*, vol. 3 (May 1992) pp. 2290–2295.
14. G. Kootstra, M. Popovi, J. A. Jørgensen, D. Kragic, H. G. Petersen and N. Krüger, "VisGraB: A benchmark for vision-based grasping," *Paladyn* **3**(2), 54–62 (2012).
15. D. Kraft, L.-P. Ellekilde and J. A. Jørgensen, "Automatic Grasp Generation and Improvement for Industrial Bin-Picking," **In**: *Gearing Up and Accelerating Crossfertilization between Academic and Industrial Robotics Research in Europe:*, volume 94 of *Springer Tracts in Advanced Robotics* (F. Röhrbein, G. Veiga and C. Natale, eds.) (Springer International Publishing, 2014) pp. 155–176.
16. J. A. Jørgensen and H. G. Petersen, "Grasp Synthesis for Dextrous Hands Optimised for Tactile Manipulation," *Proceedings for the Joint Conference of ISR 2010 (41st Internationel Symposium on Robotics) und ROBOTIK 2010 (6th German Conference on Robotics)*, VDE-Verlag (June 2010).
17. R. Suárez, M. Roa and J. Cornella, Grasp Quality Measures, *Technical Report* (Technical University of Catalonia, 2006).
18. R. Pelossof, A. Miller, P. Allen and T. Jebara, "An Svm Learning Approach to Robotic Grasping," *Proceedings of the ICRA '04. 2004 IEEE International Conference on Robotics and Automation, 2004*, vol. 4 (26-May 1, 2004), pp. 3512–3518.
19. N. Curtis and J. Xiao, "Efficient and Effective Grasping of Novel Objects through Learning and Adapting a Knowledge Base," *IROS 2008. IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008* (Sep. 2008) pp. 2252–2257.
20. C. Goldfeder, M. Ciocarlie, H. Dang and P. K. Allen, "The Columbia Grasp Database," *ICRA '09. IEEE International Conference on Robotics and Automation, 2009* (2009) pp. 1710–1716.
21. J. A. Jørgensen and H. Petersen, "Usage of Simulations to Plan Stable Grasping of Unknown Objects with a 3-Fingered Schunk Hand," *Workshop on Robot Simulators (IROS08)* (2008).
22. A. Morales, M. Prats, P. Sanz and A. P. Pobil, "An Experiment in the use of Manipulation Primitives and Tactile Perception for Reactive Grasping," *Robotics: Science and Systems, Workshop on Robot Manipulation: Sensing and Adapting to the Real World, Atlanta, USA* (2007).
23. K. Hsiao, S. Chitta, M. Ciocarlie and E. G. Jones, "Contact-Reactive Grasping of Objects with Partial Shape Information," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, vol. 10 (2010) pp. 2010.
24. J. A. Jørgensen, L. P. Ellekilde and H. G. Petersen, "Robworksim - An Open Simulator for Sensor Based Grasping," *Proceedings of Joint 41st International Symposium on Robotics (ISR 2010) and the 6th German Conference on Robotics*, Munich (2010).
25. J. A. Jørgensen, L. P. Ellekilde, D. Kraft, H. G. Petersen and N. Krüger, "On Transferability of Grasp-Affordances in Data-Driven Grasping," *Proceedings of the RAAD 2013 22nd International Workshop on Robotics in Alpe-Adria-Danube Region* (2013).
26. B. W. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA) - Protein Structure* **405**(2), 442–451 (1975).
27. A. Kasper, Z. Xue and R. Dillmann, "The KIT Object Models Database: An Object Model Database for Object Recognition, Localization and Manipulation in Service Robotics," *Int. J. Robot. Res.* **31**(8), 927–934 (July 2012).

https://doi.org/10.1017/S0263574714001192 Published online by Cambridge University Press