Value of Assistance for Grasping

Mohammad Masarwy, Yuval Goshen, David Dovrat and Sarah Keren

***Abstract*— In multiple realistic settings, a robot is tasked with grasping an object without knowing its exact pose and relies on a probabilistic estimation of the pose to decide how to attempt the grasp. We support settings in which it is possible to provide the robot with an observation of the object before a grasp is attempted but this possibility is limited and there is a need to decide which sensing action would be most beneficial. We support this decision by offering a novel *Value of Assistance* (VOA) measure for assessing the expected effect a specific observation will have on the robot's ability to complete its task. We evaluate our suggested measure in simulated and real-world collaborative grasping settings.**

I. INTRODUCTION

Task-driven agents often need to decide how to act based on partial and noisy state estimations which may greatly compromise performance. We consider settings in which an agent is tasked with grasping an object based on a probabilistic estimation of its pose. Before attempting the grasp, another agent may assist by performing a sensing action and sending its observation to the grasping agent. In our settings of interest, sensing and communication may be costly or limited and there is a need to support the decision of which observation to perform by offering principled ways to assess the expected benefit.

To demonstrate, consider the simplified automated manufacturing setting depicted in Figure 1. One agent, denoted as the *actor*, is a robotic arm with a parallel-jaw gripper that is tasked with grasping an object (here, an adversarial object from [1]). After a successful grasp, the object drops unexpectedly. Since the actor does not have a functioning sensor it can attempt to grasp the object based only on its estimation of the current position of the object. Alternatively, it can attempt the grasp after receiving an observation from another agent, the *helper*, that is equipped with a functioning sensor (here, an OnRobot 2.5D Vision System). The question that we pose is whether the helper can provide valuable assistance and what is the best position for the helper from which to provide the sensor reading among the possible options. Of course, the same question arises in single-agent settings in which it is the agent itself that needs to decide whether to perform a costly sensing action.

Beyond this illustrative example, grasping is an essential task for a wide range of robotic applications, including industrial automation, household robotics, agriculture, and more [2], [3]. Accordingly, research on effective grasping capabilities has resulted in many solution approaches that can be generally divided into two main categories [4], [3]. In analytical approaches, a representation of the physical and dynamical models of the agent and the object are used when choosing a configuration from which to attemp a grasp [5], [6], [7], [8]. In contrast, data-driven approaches rank labeled samples to come up with grasping policies. The ranking is usually based on a heuristic or on experiences collected from simulated or real robots [9], [10], [11], [12], [13].

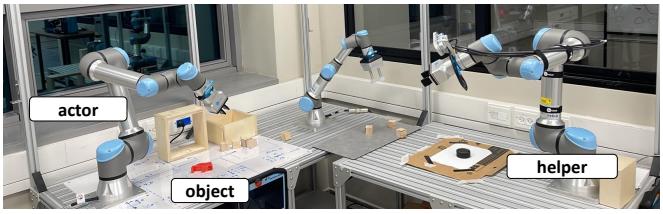


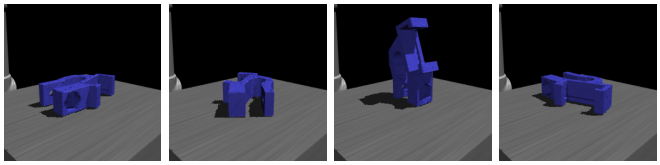
Fig. 1: Collaborative Grasping Example.

We assume the actor is associated with a procedure for choosing a grasp given its *belief* which represents its knowledge about the position of the object. To support the decision of which observation would be most beneficial, we formulate *value of assistance* (VOA) *for grasping* and offer ways to compute it for estimating the benefit a sensing action will have on the probability of a successful grasp. This involves accounting for how the actor's estimation will change based on the acquired observation and assessing how this change will affect the actor's decision of how to attempt the grasp.

VOA is used to assess the informative value an observation will have and is therefore closely related to the well-established notions of *value of information* (VOI) and *information gain* (IG) [14], [15], [16], [17], [18], which are widely used across multiple AI frameworks to assess the impact information will have on agents decisions and expected utility. We adapt these ideas to robotic settings. While in [19] we used VOA for assessing the effect localization information would have on a navigating robot's expected cost, here we use it in the context of a grasping task.

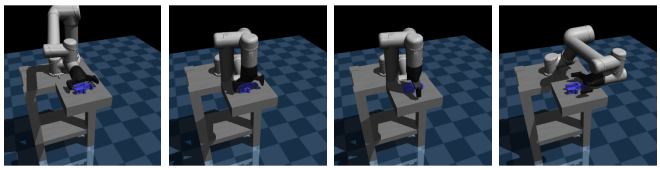
Perhaps closest to our work is *active perception* and *sensor planning* [20], [21], [22], [23], [24], [25], [26], [27], [28] which refer to the integration of sensing and decision-making processes within a robotic system. This involves actively acquiring and utilizing information from the environment and selecting viewpoints or trajectories likely to reveal relevant information or reduce uncertainty [29], [30]. While these include work on active perception in manipulation tasks they mostly focus on assessing the effect various perspectives will have on the ability to correctly locate and classify objects [20], [31]. We offer a general formulation of VOA for grasping and novel ways to estimate the effect an observation will have on the probability of accomplishing a grasp.

Since our focus is on settings in which information acquisition actions are limited and may be performed by



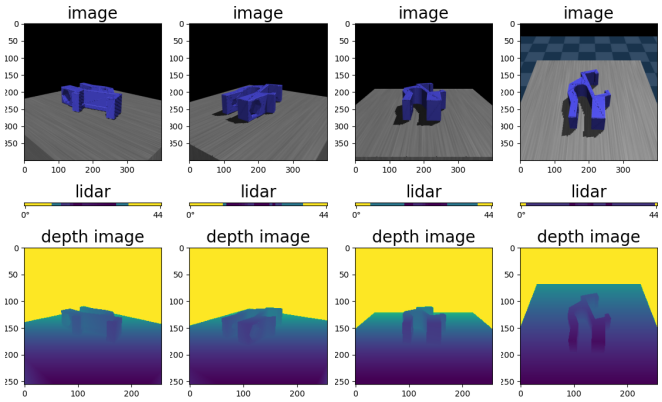
(a) (b) (c) (d)

Fig. 2: Example stable poses.



(a) (b) (c) (d)

Fig. 3: Example grasp configurations from which the actor can attempt to grasp the object - each configuration is associated with a score, i.e., probability of success.



(a) (b) (c) (d)

Fig. 4: Example sensor configurations. Each column represents the RGB image [top] lidar reading [middle] and depth image [bottom] for a sensor configuration-object pose pair.

another agent, our work is also highly related to *decision-theoretic communication* in particular, where agents communicate over a limited-bandwidth channel and messages are chosen to maximize the utility or effectiveness of the communication [32], [27], [26], [33], [34]. Our novelty is in offering measures that account for the manipulation and sensing capabilities of robotic agents when assessing the value of communicating an observation within a collaborative grasping setting. Our key contributions are the following:

1) We introduce and formulate *Value of Assistance* (VOA) for grasping.
2) We instantiate VOA for a collaborative grasping setting with a robotic arm equipped with a gripper and another agent equipped with either a lidar or a depth camera.
3) We empirically demonstrate in both simulated and real-world robotic settings how VOA predicts the effect an observation will have on performance and how it can be used to identify the best assistive action.

II. Preliminaries

To support a grasping task, where the object is assumed to be in a stable pose, we use a function that assigns a score to a grasping configuration - object stable pose pair.

Definition 1 (Grasp Score): *Given a set of object poses $\mathcal{P}$ and a set of grasp configurations $\mathcal{G}$, a* grasp score function $\gamma : \mathcal{G} \times \mathcal{P} \mapsto [0,1]$ *specifies the probability that an actor applying grasp configuration $g \in \mathcal{G}$ will successfully grasp an object at pose $p \in \mathcal{P}$, i.e. $\gamma(p,g) = P(s|p,g)$, where s is the event of a successful grasp.*

The grasp score function may be evaluated analytically, by considering diverse factors such as contact area, closure force, object shape, and friction coefficient [35], [3] or empirically, by using data-driven approaches such as [36] where a deep learning model is trained to predict the quality of grasps based on depth images of the objects.

The actor's choice of a grasp configuration relies on a *pose belief* which describes the perceived likelihood of each object pose within the set of possible stable poses $\mathcal{P}$.

Definition 2 (Pose Belief): *A* pose belief $\beta : \mathcal{P} \mapsto [0,1]$ *is a probability distribution over $\mathcal{P}$.*

The pose belief is affected by different factors, including the model of the object and its dynamics and the collected sensory information. We formulate the initial belief after the object is dropped (Figure 1) using a joint probability model that captures the prior probability of stable poses, a von Mises PDF for the angle [37], and a multivariate normal distribution for the position on the plane. See Section I of our online appendix for the formulation[1].

Definition 3 (Expected Grasp Score): *For grasp configuration g and pose belief β, the* expected grasp score $\bar{\gamma}(g,\beta) = \mathbb{E}_{p\sim\beta}[\gamma(g,p)]$ *is the weighted aggregated grasp score over the set of possible poses. A* maximal grasp *of β, denoted $g^{max}(\beta)$, maximizes the expected grasp score, i.e., $g^{max}(\beta) = \arg\max_{g\in\mathcal{G}_a} \bar{\gamma}(g,\beta)$.*

The actor receives an *observation* from the helper which corresponds to its readings from a specific sensor configuration $x \in \mathcal{X}$ and object pose $p \in \mathcal{P}$. Our formulation of the *observation space* O is general and represents the set of readings that can be made by the sensor that is available in the considered setting. In our evaluations, we used a planar lidar sensor for which a reading is an array of non-negative distances per angle $\mathbb{R}^{360}$ and a depth camera which emits a 2D array $\mathbb{R}^{w\times h}$, where $w \times h$ are the image dimensions.

As is common in the literature, (e.g., [38], [39]), we consider obtaining an observation as a stochastic process.

Definition 4 (Sensor Function): *Given object pose $p \in \mathcal{P}$, sensor configuration $x \in \mathcal{X}$, and observation $o \in O$, sensor function $\mathcal{O}(o,p,x) = P(o|p,x)$ provides the conditional probability of obtaining o from x for p.*

Notably, an agent may not be aware of the actual distribution and may instead only have a *predicted sensor function*

[1] https://github.com/CLAIR-LAB-TECHNION/GraspVOA

$\hat{\mathcal{O}}$ and a *predicted observation probability* $\hat{P}(o|p,x)$, based on a distribution which may be incorrect or inaccurate.

When receiving an observation o, the actor updates its belief using its *belief update function* which defines the effect an observation has on the pose belief.

Definition 5 (Belief Update): *A belief update function* $\tau : \mathcal{B} \times O \times \mathcal{X} \mapsto \mathcal{B}$ *maps belief* $\beta \in \mathcal{B}$*, observation* $o \in O$ *and sensor configuration* $x \in \mathcal{X}$ *to an updated belief* $\beta^{o,x}$.

The literature is rich with approaches for belief update (e.g., [38], [17], [23], [40]). We use a Bayesian filter such that for any observation $o \in O$ taken from sensor configuration $x \in \mathcal{X}$, the updated pose belief $\beta^{o,x}(p)$ for pose $p \in \mathcal{P}$ is given as

$$\beta^{o,x}(p) = \frac{\hat{P}(o|p,x)\,\beta(p)}{\int_{p' \in \mathcal{P}} \hat{P}(o|p',x)\,\beta(p')\,dp'} \tag{1}$$

where $\beta(p)$ is the estimated probability that p is the object pose prior to considering the new observation o.

When $\hat{\mathcal{O}}$ and $\hat{P}$ describe stochastic processes they can be directly used for belief update using Equation 1. Sometimes, it may be useful to consider deterministic sensor functions where the conditional distribution of an observation is replaced by a deterministic mapping $\hat{\mathcal{O}} : \mathcal{P} \times \mathcal{X} \mapsto O$. For example, a deterministic sensor model would assign the value of a cell in a lidar reading based on the predicted distance between the lidar and the object surface at a specific angle. A stochastic sensor model would sample from a Gaussian distribution with this value as the mean and the specified error margins as the standard deviation.

In some cases, such as when using a deterministic sensor model, a similarity score $\omega : O \times O \mapsto [0,1]$ is used to compare the predicted and received observations and to compose a valid distribution function for $\hat{P}$:

$$\hat{P}(o|p,x) = \frac{\omega(\hat{\mathcal{O}}(p,x),o)}{\int_{p' \in \mathcal{P}} \omega(\hat{\mathcal{O}}(p',x),o)dp'} \tag{2}$$

The literature is rich with ways to measure ω, which may vary between applications and sensor types. See Appendix III for a description of several approaches including using MSE for assessing the similarity between lidar sensor readings and an SSIM-based measure [41] for depth images.

III. Value of Assistance (VOA) for Grasping

We offer ways to assess the effect sensing actions will have on the probability of successfully grasping an object. We formulate this as a two-agent collaborative grasping setting with an *actor* that is tasked with grasping an object for which the exact pose is not known and a *helper* that can move to a specific configuration within its workspace to collect an observation from its sensor and share it with the actor.

We note that we use a two-agent model since it clearly distinguishes between the grasping and sensing capabilities. Depending on the application, this model can be used to support an active perception setting in which a single agent needs to choose whether to perform a manipulation or sensing action if it is capable of both.

The actor chooses a grasp configuration based on its *pose belief* (Definition 2). While the object's exact pose is unknown, its shape is given and it is assumed to be in a stable pose (see Figure 2). The actor needs to choose a feasible *grasp configuration* $g \in \mathcal{G}$ from which a grasp will be attempted (see Figure 3) based on its pose belief and grasp score function. The helper can choose among a set of *sensor configurations* $x \in \mathcal{X}$ that offer different points of view of the object (Figure 4) and a potentially different effect on the actor's pose belief.

We aim to assess *Value of Assistance* (VOA) for grasping as the expected benefit an observation performed from a sensor configuration will have on the actor's probability of a successful grasp. Our perspective is that of the helper and its decision of which sensing action to perform. Accordingly, we seek a way to estimate beforehand the effect an expected observation will have on the actor's belief and on its choice of configuration from which to attempt the grasp.

Importantly, the helper's belief may be different than that of the actor. We denote the helper and actor pose beliefs as β_h and β_a, respectively. We note that when considering a single agent with sensing and grasping capabilities the computation remains the same, with $\beta_a \equiv \beta_h$.

A key element in VOA computation is the expected difference between the utility of the actor with and without the intervention. Here, utility is the grasp score of the configuration chosen by the actor based on its belief.

Definition 6 (Value of Assistance (VOA) for Grasping): *Given actor belief* $\beta_a \in \mathcal{B}$*, helper belief* $\beta_h \in \mathcal{B}$*, helper perceived sensor function* $\hat{P}_h$*, sensor configuration* $x \in \mathcal{X}$ *and actor belief update function* τ_a*,*

$$U_x^{VOA}(\beta_h,\beta_a) \stackrel{\text{def}}{=\!=} \mathbb{E}_{p\sim\beta_h}\left[\mathbb{E}_{\hat{o}\sim\hat{P}_h(o|p,x)}\left[\gamma(g^{max}(\beta_a^{\hat{o},x}),p)\right] - \gamma(g^{max}(\beta_a),p))\right] \tag{3}$$

where $\hat{o}$ *is the predicted observation and* $\beta_a^{\hat{o},x}$ *is the predicted actor's belief after receiving* $\hat{o}$ *from sensor configuration* x *and updating its belief, i.e.,* $\beta_a^{\hat{o},x} = \tau_a(\beta_a,\hat{o},x)$.

In the definition above, the maximal grasp g^{max} is the one that maximizes the expected grasp score $\bar{\gamma}$ as in Definition 3. Since VOA estimation is performed by the helper, observation $\hat{o}$ is extracted from its predicted observation probability $\hat{P}_h$. In contrast, the actor's belief update is based on τ_a and uses Equation 1 with its own perceived observation probability $\hat{P}_a$.

For simplicity of presentation, we hereon assume that the actor and helper share an initial pose belief β before the grasp attempt. In addition, we assume the predicted observation $\hat{o}$ is generated using a deterministic sensor function such that $\hat{o} = \hat{\mathcal{O}}(p,x)$. The updated belief is then a function of p and x and is denoted as $\beta^{\hat{o}}$. We note that our evaluation and analysis can be adapted to the more general settings in which these assumptions are relaxed, but this allows us to use a simplified VOA formulation as follows

$$U_x^{VOA}(\beta) \stackrel{\text{def}}{=\!=} \mathbb{E}_{p\sim\beta}\left[\gamma(g^{max}(\beta^{\hat{o}}),p) - \gamma(g^{max}(\beta),p)\right] \tag{4}$$

Algorithm 1 in Appendix II describes how VOA can be used for supporting the helper's decision of which sensing action to perform. The algorithm includes the ComputeVOA function for computing VOA for a sensor configuration. We also provide a complexity analysis for the algorithm.

IV. Empirical Evaluation

The objective of our evaluation is to examine the ability of our proposed VOA measures to predict the effect sensing actions will have on the probability of a successful grasp and on finding one that maximizes this probability. With this objective in mind, our evaluation is comprised of three parts.

1) **Evaluating grasp score:** measuring the success ratio $\gamma(p, g)$ for grasping an object at pose p from grasp configuration g for different objects.
2) **Evaluating $\hat{\mathcal{O}}$:** examining the difference between the predicted sensor function $\hat{\mathcal{O}}$ and the readings of the actual sensor $\mathcal{O}$.
3) **Assessing VOA**: assessing how well VOA estimates the effect observations will have on grasp score and its ability to identify the best sensing action.

A. Experimental Setting

We performed our evaluation in a two-agent robotic setting, both at the lab and in simulation. In our lab setting, depicted in Figure 1, the actor is a UR5e robotic arm [42] with an OnRobot 2FG7 parallel jaw gripper [43]. We used two implementations for the helper with two different sensors that might be available, depending on the setting: a LDS-01 lidar [44] that could be moved on the x-y plane and a 2.5D Onrobot vision system [43] mounted on an adjacent UR5e arm. For simulation, we used a MuJoCo [45] environment (depicted in Figure 3) [46]. We simulated a lidar sensor using the MuJoCo depth camera, taking only one row of the camera's readings. The simulated gripper was a Robotiq 2F-85 parallel jaw [47]. We used five objects for the simulation and lab experiments (Figure 5). Objects meshes are based on the Dex-Net dataset [1].

We sampled a set $\tilde{\mathcal{P}} \subseteq \mathcal{P}$ of stable poses and considered four possible grasps g indexed $1-4$ (see Figure 7). We used both the lidar and depth camera. For each object pose $p \in \tilde{\mathcal{P}}$, we recorded the observation o and the predicted observation $\hat{o}$ received from each sensor configuration, indexed $I - IV$ for the lidar and $I - VI$ for the depth camera.

We empirically evaluated grasp score by examining a set of pose-grasp pairs for each object in both simulated and lab settings. Due to space constraints, the full details and results can be found in Section IV of our online appendix.

B. Evaluating the Predicted Sensor Function $\hat{\mathcal{O}}$

We evaluated our predicted sensor functions $\hat{\mathcal{O}}$ for both the lidar and depth camera by comparing their predicted observations to the readings collected from the sensor (Figure 6). For each sensor configuration-object pose pair, we recorded the mean error of the difference between the measured and predicted readings. For the lidar, the set $\tilde{\mathcal{X}}$ included four sensor configurations for each cardinal direction and

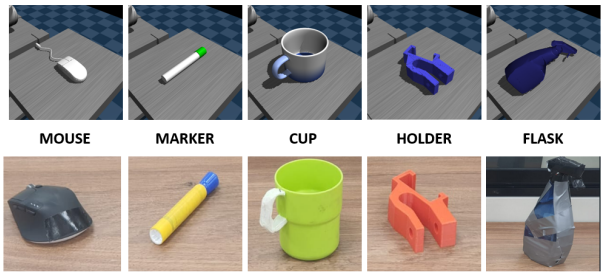
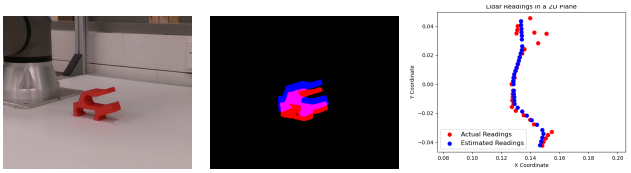


Fig. 5: Evaluation objects



(a) (b) (c)

Fig. 6: Observations of HOLDER (a) Actual scene. (b) Comparing the predicted depth image (blue) and the lab-recorded image (red). (c) Comparing the predicted (blue) and lab-recorded (red) 2D representation of the lidar reading.

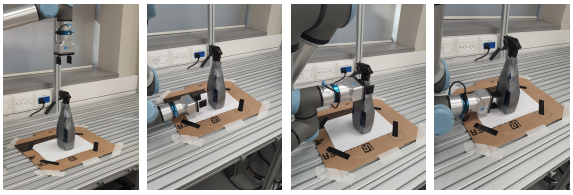
$\hat{\mathcal{O}}$ computed the closest intersection between the simulated lidar ray and the object meshes for the relevant FoV. The actual observations generated by $\mathcal{O}$ for this setup were collected in simulation and the lab. However, while in the lab the lidar was able to capture all objects, in simulation MOUSE and MARKER, could not be captured.

For the depth camera, $\hat{\mathcal{O}}$ rendered synthetic images using the pyrender library [48] which involved projecting the 3D model of an object (transformed into a specific pose), onto a 2D plane using the camera's intrinsic and extrinsic parameters. We evaluated our observation prediction accuracy compared to the actual image using the Intersection over Union (IoU) measure: we preprocessed the RBG images to extract the object masks and computed IoU between these and the corresponding synthetic images. The set $\tilde{\mathcal{X}}$ of sensor configurations was generated by randomly sampling robot configurations $q \in (-\pi, \pi]^6$ and translating them into camera poses using forward kinematics i.e. $x = FK(q)$. Each sensor configuration $x \in \tilde{\mathcal{X}}$ was ranked using a heuristic

$$H(x) = \left(1 - \frac{D(x)}{D_{max}}\right) + V(x) \tag{5}$$

where $D(x)$ is the Euclidean distance between the camera's center and the Point of Interest (PoI), representing the mean landing position after the object is dropped, D_{max} is a maximum acceptable distance used for normalization, and $V(x)$ is a visibility score, which assesses how centered the PoI is within the camera's FoV and is computed as the distance between the projection of the PoI onto the image plane and the center of the image divided by R_{ref}, a reference radius within the image plane that represents the boundary of acceptability.

Results: Table I presents results per sensor for HOLDER



(a) (b) (c) (d)

Fig. 7: Four grasp configurations for FLASK at the lab

Lidar [mm]				Depth Camera			
x	Avg. Err.	Min. Err.	Max. Err	x	Avg. IoU	Max IoU	Min. IoU
I	1.6	0.7	3.2	I	0.6	0.8	0.4
II	3.9	0.7	4.8	II	0.5	0.7	0.4
III	0.8	0.5	1.0	III	0.6	0.7	0.5
IV	3.0	0.6	4.7	IV	0.6	0.7	0.6
-	-	-	-	V	0.4	0.6	0.3
-	-	-	-	VI	0.3	0.5	0.1

TABLE I: Sensor prediction evaluation for the lab setting of HOLDER. For the lidar the lower values are better while for the depth camera higher values are better.

(results for all objects are in our online appendix). For each sensor configuration x of the lidar, the table shows the average, minimal and maximal error (Avg. Err., Min. Err. and Max. Err., respectively) in mm over object poses $\tilde{\mathcal{P}}$. Similarly, for the depth camera, we computed the average, maximal, and minimal IOU values.

Results for the lidar show that the prediction errors are negligible given the dimensions of the objects examined. In contrast, for the depth camera, errors are more substantial with a maximal average of 0.6. At the same time, results show varying performance across different configurations, depending on the object and its pose. For example, configuration I gives high accuracy for some objects, while configuration IV excels for others.

Inconsistencies between the observations are the result of several factors including the noise of the sensor itself, inconsistent scaling of the meshes with regard to the real objects, mismatches between objects and the meshes used for estimation, and inaccuracies in the placement of the objects in the lab (while the observation prediction is based on perfect object positioning). In addition, as the distance between the sensor and the object increases, the object occupies less of the sensor's FoV and the readings include fewer and less informative data points. Specific to the depth camera is the confusion caused by the reflection of bright light on shiny surfaces which may distort object shape.

Figure 6 demonstrates inconsistencies between the predicted and actual observation of HOLDER at the lab. Here, this is due to the misplacement of the object. As we show next, despite these inconsistencies, the estimated observations are still useful for VOA computation.

C. Assessing VOA

We estimate the benefit of using VOA as a decision-making tool by assessing the benefit of choosing which sensor configuration to apply based on VOA values. Our evaluation uses the setting depicted in Figure 1 and assumes the initial pose belief (after the object drops) is shared by the actor and helper. The model of the initial belief is described in Section 2.

We used three belief update functions per sensor, each based on a different similarity metric. For the lidar, τ_1 uses a deterministic update rule that considers two observations o_i, o_j as equivalent if for all angles the values are within a margin of 8 mm, τ_2 uses the similarity metric $\omega(o_i, o_j) = e^{-\|o_i - o_j\|}$ to update the belief based on Equation 2, while τ_3 uses a multidimensional Gaussian over one observation while the other observation is the mean vector and the covariance matrix is the identity matrix. For the depth camera, τ_4 is based on the structure element of SSIM [41], τ_5 is based on IoU between the two observations, and τ_6 employs the cv2 library [49] for contour matching to quantify the similarity between two masks by detecting their primary contours and comparing their shapes through a shape-matching algorithm.

Results: Table II presents our evaluation for the different objects at the lab[2]. For each setting, we consider three grasps: g^* is an optimal grasp, g_i is the grasp chosen by the actor based on its initial belief, and g_f is its post-intervention choice, i.e., after receiving an observation from a sensor configuration with the highest VOA value (see Figure 8). For each belief update function and object, the table reports the average values of:

- $\delta = \mathbb{E}_{p\sim\beta}[\gamma(g_f, p) - \gamma(g_i, p)]$: the weighted score difference between the chosen grasp after and before the intervention.
- $\delta^* = \frac{\mathbb{E}_{p\sim\beta}[\gamma(g_f,p)-\gamma(g_i,p)]}{\mathbb{E}_{p\sim\beta}[\gamma(g^*,p)-\gamma(g_i,p)]}$: the ratio between δ and the weighted score difference between the best configuration and the configuration chosen before the intervention.
- $\mathcal{A} = \frac{\delta^*(x)}{E_{x'\sim\tilde{\mathcal{X}}}[\delta^*(x')]}$: the advantage of choosing the maximal VOA sensor configuration defined as the ratio between δ^* and the average over all configurations. This represents the difference between choosing a sensor configuration using VOA to choosing randomly.

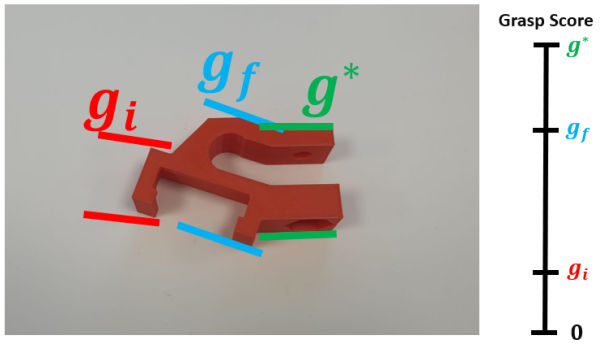


Fig. 8: Grasps for HOLDER: best grasp g^*, initial chosen grasp g_i and chosen grasp after the intervention g_f.

[2]due to space constraints, complete results as well as our implementation, are in our online appendix.

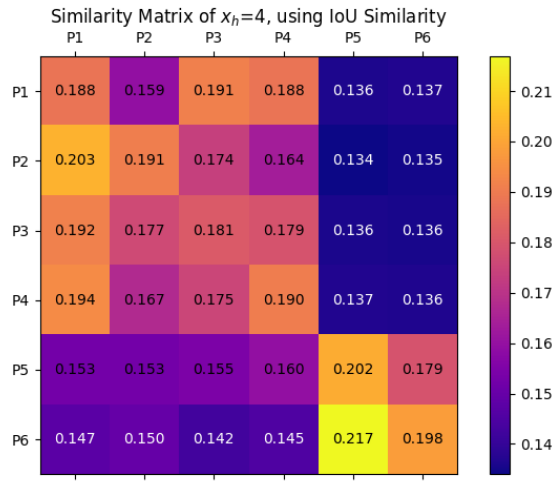


Fig. 9: Similarity score between actual observation (rows) for each pose and predicted observations (columns).

(a) P1 (b) P2 (c) P3 (d) P4 (e) P5 (f) P6

Fig. 10: Observations for different object poses of HOLDER

Results show that for all objects and for 4 of our examined belief update functions (τ_1-τ_4), selecting the sensor configuration with the highest VOA value is beneficial in terms of the three examined measures. The smallest benefit is for MOUSE for which the initial grasp is optimal for all poses except one for which the optimal grasp has only a slight advantage. This subtlety is captured only by τ_4.

Notably, belief update functions τ_5 and τ_6 which rely on IOU and contour matching, respectively, did not perform well on average for any of the objects. We associate this with the fact that the objects examined are small relative to their distance from the sensor, which is something these update functions are sensitive to. Figure 9 presents the similarity scores between the actual (rows) and predicted (columns) observations for τ_5 that relies solely on the IoU of the masks without considering depth values. This makes it hard for τ_5 to differentiate between object poses that occupy the same area across multiple poses, as depicted in Figure 10. The matrix shows a clear distinction between the standing positions 5 and 6 and the laying positions $1-4$, but the distinction within these two groups is challenging. Another critical issue is demonstrated by the values on the diagonal that show the low similarity scores between the predicted and actual observations. These indicate the sensitivity of τ_5 to noise in the actual image, where even slight distortions can dramatically impact prediction quality. Similar results were observed for τ_6 where noise dramatically distorts contours.

V. CONCLUSION

We introduced *Value of Assistance* (VOA) for grasping and suggested ways to estimate it for sensing actions. Our experiments in both simulation and real-world settings demonstrate how our VOA measures predict the effect an observation will have on performance and how it can be used to support the decision of which observation to perform.

Future work will account for optimization considerations of the helper and for integrating VOA in long-term and complex tasks. Another extension will consider multi-agent settings in which VOA can be used not only for choosing which assistive action to perform but also for choosing which agent to assist.

		δ	δ^*	$\mathcal{A}$
HOLDER	τ_1	**0.19**	0.23	0.23
	τ_2	0.15	0.23	**0.29**
	τ_3	0.2	**0.26**	0.28
	τ_4	0.03	0.15	0.13
	τ_5	0.0	0.0	0.0
	τ_6	0.0	0.0	0.0
EXPO	τ_1	**0.29**	0.31	**0.29**
	τ_2	**0.29**	0.31	**0.29**
	τ_3	**0.29**	0.31	**0.29**
	τ_4	0.04	**0.37**	**0.29**
	τ_5	0.00	0.00	0.00
	τ_6	-0.00	0.29	**0.29**
MOUSE	τ_1	0.00	0.06	0.04
	τ_2	0.00	0.05	0.04
	τ_3	0.00	0.07	0.00
	τ_4	**0.04**	**0.44**	**0.53**
	τ_5	0.00	0.00	0.00
	τ_6	0.00	0.00	0.00
CUP	τ_1	0.30	0.33	0.30
	τ_2	0.30	0.33	**0.41**
	τ_3	**0.37**	**0.63**	0.37
	τ_4	0.24	0.38	0.27
	τ_5	0.00	0.00	0.00
	τ_6	0.00	0.00	0.00
FLASK	τ_1	**0.25**	**0.25**	**0.25**
	τ_2	**0.25**	**0.25**	**0.25**
	τ_3	**0.25**	**0.25**	**0.25**
	τ_4	0.02	**0.25**	**0.25**
	τ_5	0.00	0.00	0.00
	τ_6	0.00	0.00	0.00
AVG	τ_1	0.19	0.23	0.23
	τ_2	0.19	0.23	0.29
	τ_3	**0.20**	**0.26**	0.28
	τ_4	0.08	0.22	**0.34**
	τ_5	0.00	0.00	0.00
	τ_6	0.00	0.00	0.00

TABLE II: Results per belief update function (best results per criteria are highlighted)

REFERENCES

[1] "dexnet," https://berkeleyautomation.github.io/dex-net.

[2] O. Kroemer, S. Niekum, and G. Konidaris, "A review of robot learning for manipulation: challenges, representations, and algorithms," *J. Mach. Learn. Res.*, vol. 22, no. 1, jan 2021.

[3] A. Bicchi and V. Kumar, "Robotic grasping and contact: A review," in *Proceedings of the 2000 IEEE International Conference on Robotics and Automation, (ICRA)*. IEEE, 2000.

[4] A. Sahbani, S. El-Khoury, and P. Bidaud, "An overview of 3d object grasp synthesis algorithms," *Robotics Auton. Syst.*, 2012.

[5] A. M. Hands, "Kinematic and force analysis of," *Journal of Mechanisms, Transmissions, and Automation in Design*, 1983.

[6] D. Prattichizzo and J. C. Trinkle, "Grasping," in *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Springer, 2008.

[7] F. T. Pokorny and D. Kragic, "Classical grasp quality evaluation: New algorithms and theory," in *Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, Year, p. Page Range.

[8] D. Berenson, S. Srinivasa, and J. Kuffner, "Task space regions: A framework for pose-constrained manipulation planning," *The International Journal of Robotics Research*, 2011.

[9] 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.

[10] R. Balasubramanian, L. Xu, P. D. Brook, J. R. Smith, and Y. Matsuoka, "Physical human interactive guidance: Identifying grasping principles from human-planned grasps," *IEEE Transactions on Robotics*, vol. 28, no. 4, pp. 899–910, 2012.

[11] L. Pinto and A. Gupta, "Supersizing self-supervision: Learning to grasp from 50k tries and 700 robot hours," in *Proc. IEEE Int. Conf. Robotics and Automation (ICRA)*, 2016.

[12] S. Nahavandi, R. Alizadehsani, D. Nahavandi, C. P. Lim, K. Kelly, and F. Bello, "Machine learning meets advanced robotic manipulation," *Information Fusion*, vol. 105, p. 102221, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1566253523005377

[13] M. Khansari, D. Kappler, J. Luo, J. Bingham, and M. Kalakrishnan, "Action image representation: Learning scalable deep grasping policies with zero real world data," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 3597–3603.

[14] R. A. Howard, "Information value theory," *IEEE Transactions on Systems Science and Cybernetics*, vol. 2, no. 1, pp. 22–26, 1966.

[15] S. Russell and E. Wefald, "Principles of metareasoning," *Artificial Intelligence*, vol. 49, no. 1, pp. 361–395, 1991.

[16] S. Zilberstein and V. Lesser, "Intelligent information gathering using decision models," Computer Science Department, University of Massachussetts Amherst, Tech. Rep. 96-35, 1996.

[17] C. Stachniss, G. Grisetti, and W. Burgard, "Information gain-based exploration using rao-blackwellized particle filters." in *Robotics: Science and systems*, vol. 2, 2005, pp. 65–72.

[18] C. Cai and S. Ferrari, "Information-driven sensor path planning by approximate cell decomposition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 3, pp. 672–689, 2009.

[19] A. Amuzig, D. Dovrat, and S. Keren, "Value of assistance for mobile agents," 2023.

[20] J. Shin, S. Chang, J. Weaver, J. C. Isaacs, B. Fu, and S. Ferrari, "Informative multiview planning for underwater sensors," *IEEE Journal of Oceanic Engineering*, vol. 47, no. 3, pp. 780–798, 2022.

[21] È. Pairet, J. D. Hernández, M. Lahijanian, and M. Carreras, "Uncertainty-based online mapping and motion planning for marine robotics guidance," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 2367–2374.

[22] V. Chandrasekhar, W. K. Seah, Y. S. Choo, and H. V. Ee, "Localization in underwater sensor networks: survey and challenges," in *Proceedings of the ACM international workshop on Underwater networks*, 2006, pp. 33–40.

[23] V. Indelman, L. Carlone, and F. Dellaert, "Planning in the continuous domain: A generalized belief space approach for autonomous navigation in unknown environments," *The International Journal of Robotics Research*, vol. 34, no. 7, pp. 849–882, 2015.

[24] V. Indelman, S. Williams, M. Kaess, and F. Dellaert, "Information fusion in navigation systems via factor graph based incremental smoothing," *Robotics and Autonomous Systems*, vol. 61, no. 8, pp. 721–738, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092188901300081X

[25] Y. Tian, Y. Chang, F. H. Arias, C. Nieto-Granda, J. P. How, and L. Carlone, "Kimera-multi: Robust, distributed, dense metric-semantic slam for multi-robot systems," *IEEE Transactions on Robotics*, vol. 38, no. 4, 2022.

[26] S. Koenig and Y. Liu, "Sensor planning with non-linear utility functions," in *Recent Advances in AI Planning*, S. Biundo and M. Fox, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 265–277.

[27] R. Becker, A. Carlin, V. Lesser, and S. Zilberstein, "Analyzing myopic approaches for multi-agent communication," *Computational Intelligence*, 2009.

[28] R. Mirsky, W. Macke, A. Wang, H. Yedidsion, and P. Stone, "A penny for your thoughts: The value of communication in ad hoc teamwork," *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence*, pp. 254–260, July 2020.

[29] A. T. Taylor, T. A. Berrueta, and T. D. Murphey, "Active learning in robotics: A review of control principles," *Mechatronics*, vol. 77, p. 102576, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957415821000659

[30] R. Bajcsy, Y. Aloimonos, and J. K. Tsotsos, "Revisiting active perception," *Autonomous Robots*, vol. 42, pp. 177–196, 2018.

[31] Q. V. Le, A. Saxena, and A. Y. Ng, "Active perception: Interactive manipulation for improving object detection," *Standford University Journal*, 2008.

[32] A. Fern, S. Natarajan, K. Judah, and P. Tadepalli, "A decision-theoretic model of assistance," *Journal of Artificial Intelligence Research*, vol. 50, pp. 71–104, 2014.

[33] C. V. Goldman and S. Zilberstein, "Optimizing information exchange in cooperative multi-agent systems," in *Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems*, ser. AAMAS '03. New York, NY, USA: Association for Computing Machinery, 2003, p. 137–144.

[34] R. J. Marcotte, X. Wang, D. Mehta, and E. Olson, "Optimizing multi-robot communication under bandwidth constraints," *Autonomous Robots*, vol. 44, no. 1, pp. 43–55, Jan. 2020.

[35] R. M. Murray, S. Sastry, and Z. Li, "A mathematical introduction to robotic manipulation," 1994.

[36] J. Mahler, J. Liang, S. Niyaz, M. Laskey, R. Doan, X. Liu, J. A. Ojea, and K. Goldberg, "Dex-net 2.0: Deep learning to plan robust grasps with synthetic point clouds and analytic grasp metrics," 2017.

[37] K. V. Mardia and P. E. Jupp, *Directional statistics*. John Wiley & Sons, 2009.

[38] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*, ser. Intelligent Robotics and Autonomous Agents series. MIT Press, 2005. [Online]. Available: https://dl.acm.org/doi/10.5555/1121596

[39] M. Lauri, D. Hsu, and J. Pajarinen, "Partially observable markov decision processes in robotics: A survey," *IEEE Transactions on Robotics*, vol. 39, no. 1, pp. 21–40, 2022.

[40] H. Kurniawati, "Partially observable markov decision processes and robotics," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, pp. 253–277, 2022.

[41] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600 – 612, 05 2004.

[42] "Universal Robotics-UR5," https://www.universal-robots.com/products/ur5-robot/.

[43] "Onrobot," https://onrobot.com/en/products/2fg7.

[44] "Robotis," www.robotis.us/360-laser-distance-sensor-lds-01-lidar.

[45] "mujoco," https://mujoco.org/.

[46] P. Daniel, "Mujoco_rl_ur5," https://github.com/PaulDanielML/MuJoCo_RL_UR5.

[47] "robotiq," https://robotiq.com/products/2f85-140-adaptive-robot-gripper.

[48] mmatl, "Pyrender," 2019, python 3D rendering toolkit. [Online]. Available: https://github.com/mmatl/pyrender

[49] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[50] K. Andrey, *Foundations of the Theory of Probability*. New York: Chelsea Publishing Company, 1933.

APPENDIX

I. POSE BELIEF

The pose belief is affected by different factors, including the model of the object and its dynamics and collected sensory information. We formulate the initial belief after the object is dropped (Figure 1) using a joint probability model that captures the prior probability of stable poses, a von Mises PDF for the angle[37], and a multivariate normal distribution for the position on the plane.

Formally,

$\beta(p = (C, \theta, X, Y)) = P(C) \cdot f_\theta(\theta|C; \mu_{\theta,C}, \kappa_C) \cdot f_{XY}(X, Y|C; \mu_C, \Sigma_C)$

In our model, the mean of the Gaussian distribution, μ_C, represents the point where the agent drops the object, directly linking the expected landing position to the drop location. The covariance matrix, Σ_C, reflects the uncertainty in the object's landing position, which is influenced by the height from which the object is dropped; a higher drop height introduces greater variability in the landing outcome. The orientation parameter, $\mu_{\theta,C}$, is affected by the object's pose just before the drop, incorporating the initial orientation's influence on the final resting orientation.

- $P(C)$ is the probability of the category C, where C is a discrete random variable representing the category of stable poses for a given object. The variable C takes values from a finite set of predefined categories, each corresponding to a distinct stable pose of the object. For example, the adversarial object in our example has five distinct categories, corresponding to the sides with a large enough surface.
- $f_\theta(\theta|C; \mu_{\theta,C}, \kappa_C)$ is a von Mises probability distribution function [37] that represents the conditional probability density function of the angle θ, given the category C, with $\mu_{\theta,C}$ as the mean direction and κ_C as the concentration parameter specific to category C.
- $f_{XY}(X, Y; \mu, \Sigma)$ denotes the conditional probability density function of the position (X, Y), given the category C, with μ_C as the mean position and Σ_C as the covariance matrix, again specific to category C.

The actor uses its pose belief β and grasp score function γ to select a grasping configuration $g \in \mathcal{G}_a$ from which to attempt to grasp the object. A reasonable choice is to select a configuration that maximizes the *expected grasp score*.

II. FINDING MAXIMAL VOA

We observe that $\hat{\mathcal{O}}(p, x)$ could be a bottleneck in the algorithm, depending on the sensor used. We will denote its complexity as $O(\hat{\mathcal{O}})$. Also it depends on the implementation, but ω can be a bottleneck.

Analysing the time complexity of `ComputeVOA`:

1) Outer Loop Over $\tilde{\mathcal{P}}$: The outer loop iterates over the sampled set of poses $\tilde{\mathcal{P}}$, so it runs $|\tilde{\mathcal{P}}|$ times.
2) item Observation Prediction Function and Belief Update: Within each iteration of the outer loop, the $\hat{\mathcal{O}}$ function is called, adding $O(\hat{\mathcal{O}})$. The belief update is also called, which contains its own internal loop over $\tilde{\mathcal{P}}$ and calls $\hat{\mathcal{O}}$ and ω in each iteration. Therefore, the belief update contributes $O(|\tilde{\mathcal{P}}|(\hat{\mathcal{O}} + \omega))$.
3) Grasp Calculation Over $\mathcal{G}_a$ and $\tilde{\mathcal{P}}$: The best grasp calculation, which occurs for each pose, has a complexity of $|\mathcal{G}_a||\tilde{\mathcal{P}}|$ since it runs for each grasp configuration and each pose.

Combining the complexities from these components, we get:

- For each pose in $\tilde{\mathcal{P}}$, the estimated sensor function and the belief update yield $O(\hat{\mathcal{O}}) + |\tilde{\mathcal{P}}| \times O(\hat{\mathcal{O}})$.

Algorithm 1: Choose helping action with VOA

Input: Grasp configuration set $\mathcal{G}_a$
Input: Sampled subset of help configurations $\tilde{\mathcal{X}} \subseteq \mathcal{X}$
Input: Actor belief β_a.
Input: Helper belief β_h.
Input: Grasp score function γ.
Input: Observation prediction function $\hat{\mathcal{O}}$
Input: Sampled subset of object pose set $\tilde{\mathcal{P}} \subseteq \mathcal{P}$.
Input: Belief update function τ
Output: Sensor configuration $x^* \in \tilde{\mathcal{X}}$

Compute $g^{max}(\beta_a)$; ▷ Chosen grasp without help. Definition 3
Compute $\bar{\gamma}(g^{max}(\beta_a), \beta_h)$; ▷ Mean score without help
$x^* \leftarrow Null$; ▷ Chosen helping action
$U^{VOA^*} \leftarrow 0$; ▷ best VOA
for $x \in \tilde{\mathcal{X}}$ **do**
  $U_x^{VOA} \leftarrow$ `ComputeVOA`$(x, \tilde{\mathcal{P}}, \bar{\gamma}(g^{max}(\beta_a), \beta_h), \beta_a, \beta_h, \tau)$;
  if $U_x^{VOA} > U^{VOA^*}$ **then**
    $U^{VOA^*} \leftarrow U_x^{VOA}$;
    $x^* \leftarrow x$;
  end
end
return x^*;
Function `ComputeVOA` $(x, \tilde{\mathcal{P}}, \bar{\gamma}(g^{max}(\beta_a), \beta_h), \beta_a, \beta_h, \tau)$**:**
  $\mathbf{\Gamma}_{\tilde{\mathcal{P}}} \leftarrow \emptyset$; ▷ Score per sampled pose after help
  for $p \in \tilde{\mathcal{P}}$ **do**
    $\hat{o} \leftarrow \hat{\mathcal{O}}(p, x)$; ▷ predicted observation for pose
    $\beta_a^{o,x} \leftarrow \tau(\beta_a, \hat{o}, x)$; ▷ Predicted belief. Eq. 1
    Compute $g^{max}(\beta_a^{o,x})$; ▷ Grasp with help. Def. 3
    Compute $\gamma(g^{max}(\beta_a^{o,x}), \beta_h)$; ▷ Predicted score
    $\mathbf{\Gamma}_{\tilde{\mathcal{P}}} \leftarrow \mathbf{\Gamma}_{\tilde{\mathcal{P}}} \cup \{\gamma(g^{max}(\beta_a^{o,x}), \beta_h)\}$;
  end
  return `Mean`$(\mathbf{\Gamma}_{\tilde{\mathcal{P}}}) - \bar{\gamma}(g^{max}(\beta_a), \beta_h)$;
end function

- The grasp calculation contributes $|\mathcal{G}_a||\tilde{\mathcal{P}}|$.

So the total complexity is given by:

$$|\tilde{\mathcal{P}}| \times \left(O(\hat{\mathcal{O}}) + |\tilde{\mathcal{P}}| \times O(\hat{\mathcal{O}}) + |\tilde{\mathcal{P}}| \times O(\omega) + |\mathcal{G}_a||\tilde{\mathcal{P}}|\right)$$

.

Simplifying this expression, we get:

$$O\left(|\tilde{\mathcal{P}}|^2 \left(|\mathcal{G}_a| + \hat{\mathcal{O}} + \omega\right)\right)$$

In 1 we iterate over $\tilde{\mathcal{X}}$, where in each iteration we run `ComputeVOA`, Thus, the total complexity of the algorithm is: $O\left(|\tilde{\mathcal{X}}||\tilde{\mathcal{P}}|^2 \left(|\mathcal{G}_a| + \hat{\mathcal{O}} + \omega\right)\right)$

One significant optimization we've introduced involves the computation of $\hat{\mathcal{O}}$, which are utilized at least $|P_o|$ times within the algorithm. Recognizing this, we precalculate a similarity matrix denoted as S, which consists of all pairwise similarities between the estimated observations. In other words, $S_{i,j} = \omega(o_i, o_j)$. It's important to note that the observations are exclusively used for the purpose of calculating $\omega(o_i, o_j)$ in the belief update. The resulting complexity of `ComputeVOA` is:

$$O\left(|\tilde{\mathcal{P}}|^2 \left(|\mathcal{G}_a| + \omega\right) + |\tilde{\mathcal{P}}|\hat{\mathcal{O}}\right)$$

III. Observation Similarity Scores

The following discussion presents three ways to implement ω but we note that our framework is not restricted to these methods. The first measure deterministically decides whether two observations o and o' are equivalent, i.e., $o \equiv o'$. We note that such a measure does not constrain the sensor to be deterministic but instead imposes a deterministic rule for deciding whether two observations are considered equivalent, e.g., if the difference between their values is within some bound.

A second option is to use any arbitrary similarity metric as long as it simulates a probability in the sense that the Kolmogorov axioms [50] are met when comparing one observation with the simulated probability distribution's mean. For example, to assess structural similarity between depth image pairs o and o', we applied the structure element of Structural Similarity Index (SSIM) [41] using the formula:

$$s(o, o') = \frac{\sigma_{oo'} + c}{\sigma_o \sigma_{o'} + c}$$

In this formula:

- σ_x and σ_y are the standard deviation of the observations o and o'.
- $\sigma_{oo'}$ is the covariance between observations o and o'.
- c is a small constant to stabilize the division.

As a third option, uncertainty is approximated using a compact parametric model, e.g., a Gaussian distribution, which is defined by two parameters: the mean μ and the variance σ^2. In the context of a similarity model, the mean could represent the point of maximum similarity (which could be zero distance for identical vectors), and the variance could control how quickly the similarity decreases as the distance increases. Here we assume that the similarity between two observations o and o' can be modeled as a function of the distance between them in some feature space. Using the Euclidean distance yields:

$$\omega(o, o') = \exp\left(-\frac{\|o - o'\|}{2\sigma^2}\right)$$

Small σ^2 makes the similarity function sharply peak around zero distance, making the model very sensitive to small changes in distance and large σ^2 makes the decay of similarity with distance more gradual.

Examples of these methods of approximating the observation probability are given in our empirical evaluation in Section IV.

Note that the observation prediction function is equivalent to a deterministic sensing function, as defined in Definition 4.

In the general case, the observation function can be stochastic, thus the grasp score after help is stochastic, and there is an expectation in the definition. In the case of a deterministic observation function, the expectation is not necessary.

Ideally, the helper agent would be able to compute the value difference for to support it's decision about a helping action, but in practice, the pose is unknown, thus the actual value difference is intractable. Instead, the helper agent can estimate the value difference using it's belief β_h.

IV. Evaluating Grasp Score

A. Evaluating grasp score

Setup: We empirically evaluated grasp score by examining a set of pose-grasp pairs for each object. In the lab setting, we examined 6 poses and 4 grasps. In simulation, we examined 3-6 poses and 3-4 grasps. For every pose-grasp pair, we ran 100 and 15 grasp attempts in simulation and at the lab, respectively. In simulation, for each grasp attempt a Gaussian noise was added to the object position (in meters) and orientation (in radians). We ran 4 experiment sets with different standard deviations of the noise: 0, 0.01, 0.03 and 0.05. At the lab, we used a set of 15 noisy variations of the object position and orientation.

Results: For each pose-grasp pair, we counted the number of successful grasps. Results demonstrate the diversity of our setup, with some grasps that have a high probability of success for several poses and others that are adequate for a limited set of poses.

Figure 11 shows the ratio of successful grasps for FLASK in the simulated and lab settings. Rows represent stable poses and the columns represent grasp configurations (results for all other objects are in the supplementary material). In simulation, with a uniform belief over the poses, the agent should attempt grasp 2 because it has the highest expected grasp score. If, however, the belief associates a high probability to pose 2 or 3, the agent should attempt grasp 0.

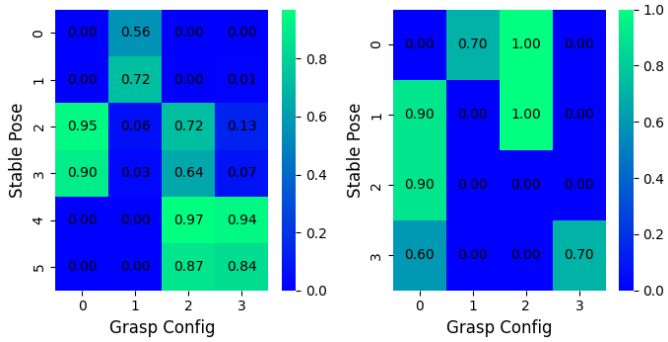


Fig. 11: FLASK grasp score for the simulated [left] and lab [right] settings.