Learning to Grasp by Extending the Peri-Personal Space Graph

Jonathan Juett[1] and Benjamin Kuipers[1]

Abstract— We present a robot model of early reach and grasp learning, inspired by infant learning without prior knowledge of the geometry, kinematics, or dynamics of the arm.

Human infants at reach onset are capable of using a sequence of jerky submotions to bring the hand to the position of a nearby object. A robotic learning agent can produce qualitatively similar behavior by using a graph representation to encode a set of safe, potentially useful arm states and feasible moves between them. These observations show that the Peri-Personal Space (PPS) Graph model is sufficient for early reaching and suggest that infants may use analogous models during this phase.

In this paper, we show that the PPS Graph, with a simulated Palmar reflex (a reflex in infants that closes the fingers when the palm is touched), allows accidental grasps to occur during continued reaching practice. Given these occasional events, the agent can bootstrap to a simple deliberate grasp action. In particular, the agent must learn three new necessary conditions for a grasp: the hand should be open as the grasp begins, the final motion of the hand should be led by the gripper opening so that it reaches the target first, and the wrist must be oriented such that the gripper fingers may close around the target object, often requiring the opening to be perpendicular to the object's major axis. Combined with the existing capability to reach and interact with target objects, knowledge of these conditions allows the agent to learn increasingly reliable purposeful grasps. The first two conditions are addressed in this paper, and allow 45% of grasps to succeed.

This work contributes toward the larger goal of foundational robot learning after the model of infant learning, with minimal prior knowledge of its own anatomy or its environment. The ability to grasp will allow the agent to control the motion and position of objects, providing a richer representation for its environment and new experiences to learn from.

I. Introduction

How can an embodied intelligent agent acquire knowledge of the space within which it is embedded?

To answer this question, it is not adequate to assume that a human programmer, already knowledgeable about the nature and structure of space, designs and specifies the necessary data structures. This assumption accepts an infinite regress: How did the human programmer acquire his/her knowledge of space? The agent must construct its own representation for space from its own sensorimotor experience, acting with its body within its own spatial environment and observing how its sensory experience is affected.

Most work in robot manipulation assumes precisely known structure for the robot arm, so that forward and inverse kinematics can be inferred using well-understood geometric methods. However, a few researchers have considered the problem of learning a suitable model of the robot. Sturm,

[1] Division of Computer Science and Engineering, University of Michigan, 2260 Hayward St, Ann Arbor, MI 48109. {jonjuett, kuipers} at umich.edu.

Plagemann & Burgard [11] learn a Bayesian generative model of the robot arm from a sequence of visual images of the arm while motor babbling. They identify the network topology by starting with a fully connected model and eliminating unnecessary links. However, they also label the parts of the robot with markers that allow the 6D pose to be extracted from each visual image, whereas we make much weaker assumptions about the nature of visual perception. Hersch, Sauser & Billard [4] learn a body schema for a humanoid robot, modeled as a tree-structured hierarchy of frames of reference. They assume that the robot can perceive and track the 3D position of each end-effector, and that the robot is given the topology of the network of joints and segments. Jamone, Natale, Nori, Metta & Sandini [5] focus on autonomous learning of goal-directed reaching, where the robot fixates at the target object, and then attempts to bring its hand to the fixation point. Their robot maintains and adapts its model of manipulator kinematics, and the properties of the visual system. They also build and update a numerical model of reachable space. Their reaching controller uses both open-loop and closed-loop controllers. Our method relies on weaker assumptions about the motor and perceptual system, and on qualitative models of peri-personal space.

In previous work, we have described how a mobile robot, with uninterpreted sensors and end effectors, can construct a useful model of its own sensorimotor system [10], and how such a model can be used to learn to distinguish objects from the static environment, and to learn simple models of actions to transform the states of objects [8]. We also took initial steps toward understanding how an autonomously exploring agent could construct its own hierarchical models of actions for a manipulating robot [9].

In this paper (building on [6]), we describe how a physical robot (Baxter), without given knowledge of the spatial structure of its workspace or the configuration space of its manipulator arm, can learn through autonomous exploration how to reach and grasp nearby objects with increasing reliability.

II. Peri-Personal Space

Peri-personal space (PPS) is the space immediately around the agent, accessible to one or more manipulator arms, and generally (when not occluded) by visual perception. Traditional robotics makes use of detailed and precise knowledge of manipulator and perceptual geometry to build and use a precise model of the three-dimensional workspace. In the developmental robotics context, we assume that the agent

(human baby or baby robot) cannot do that, because manipulator and perceptual geometry are unknown and changing.

A. Our Representation for Peri-Personal Space

We have previously introduced a graph-based model for autonomous learning in peri-personal space. In this hypothesized model, a physically-embodied robot learning agent starts with minimal knowledge of its sensors, its effectors, and its environment. It learns a graph (the PPS Graph) which represents a correspondence between visual and proprioceptive perception of the environment, including its own body. The PPS graph is a kind of *probabilistic roadmap* [7] approximating the configuration space for the manipulator arm, constructed from unguided motor babbling experience.

Each node n_i in the PPS Graph represents a state of the arm in the environment, and is associated with the visual sense p_i and the proprioceptive sense vector q_i observed at that state. In our previous work, the visual sense p_i was a vector of 2-D images the robot received from three independently-placed fixed cameras. Here, p_i is a single low-resolution (160×120) RGB-D image instead. This change makes the vision system more similar to human vision, with the depth image closely related to the disparity measure from stereo vision. The proprioceptive sense vector q_i is the vector of joint angles of the arm. An edge $e(n_i, n_j)$ connects n_i and n_j if the distance between them, defined as $||q_i - q_j||$, is sufficiently small to assume interpolated motion is safe.

We assume that the agent is able to use simple image processing techniques to generate binary images corresponding to important segments of images. Let g_i denote the binary image for the gripper segment in p_i. The two-dimensional center of mass of g_i can be expressed by coordinates (u_i, v_i). The agent can also determine the range of depth channel values within an arbitrary region of an RGB-D image I specified by a pixel mask M, denoted $D(I, M)$. The range of depths occupied by the gripper at n_i is represented as $D(p_i, g_i)$. The center of the hand as seen by the RGB-D camera is $c_i = (u_i, v_i, d_i)$, where d_i is the mean of all values in the segment of the depth image specified by g_i.

For both reaching and grasping, the agent's goal is to interact with a target object in its peri-personal space. This requires a trajectory through the PPS Graph, ending at a *final node* n_f that will accomplish the desired interaction. Successful reaching depends only on a good selection of n_f, but successful grasping depends on the final segment $e(n_p, n_f)$ from the *penultimate node* n_p to the final node.

III. MAKING RARE EVENTS RELIABLE

The behavior and learning by the agent is driven by a kind of *intrinsic motivation* [1]. At a given stage in the learning process, the agent practices some kind of behavior, and learns the expected results of those actions. Occasionally, it experiences an unusual and interesting type of result. It then attempts to learn additional preconditions or modified actions that allow it to achieve that rare and unusual result reliably.

Prior to learning to reach, motor babbling actions would reliably move the hand, usually leaving all other aspects of the environment unchanged. The unusual event would be to change the position of another object in the workspace, for example by bumping and knocking over a block. The agent then takes on the goal of deliberately and reliably changing the position of a selected object.

Once the agent has learned to reach, many actions leave other objects unchanged, but a reaching action reliably moves the target object, quasi-statically. That is, the object would change from one static position to another. The unusual event would be for the target object to become (temporarily) linked with the manipulator hand, via a grasp action, so it continues to move as the hand moves. (This link is eliminated when the object is ungrasped.)

IV. EARLY REACH PLANNING WITH THE PPS GRAPH

We have demonstrated [6] that an agent using a learned PPS Graph model of about 3000 nodes is capable of learning a reach action by which a physical robot may interact with an object in its environment. A successful reach is one that causes a persistent change in the visual percept of the target object. Most often, the interaction in a successful reach is simply a bump, changing the position of the object or knocking it over. To plan a reach in the PPS Graph, the agent selects a final node, and then finds a least cost path along graph edges, which may be traversed by linear interpolation between the stored configuration for each node. Our agent learned that its reaches were most reliable when the final node was chosen such that the region defined by stored images of the gripper's final move to this node met certain conditions. In particular, the agent had a much greater chance of success when all three cameras perceived a significant intersection between this region and the target object in the visual percept at planning time. Selecting final nodes for manipulator trajectories according to this criterion yielded successful reaches on 90% of trials.

With our change to a new vision system, this condition has been replaced. The new reach criteria are applied to the RGB-D images stored in the PPS Graph, $\{p_i\}$, and the current visual percept, an image I_0 in which the target is identified by pixel mask t and center of mass c_t. Candidate final nodes n_i must satisfy two criteria. First, there is a non-empty spatial intersection in the RGB image, $g_i \cap t \neq \emptyset$. Second, within that 2D intersection, there should be an overlap in depth ranges, that is, $D(p_i, g_i \cap t) \cap D(I_0, g_i \cap t) \neq \emptyset$. If multiple nodes meet these criteria, n_f is chosen to maximize

$$\frac{|D(p_f, g_f \cap t) \cap D(I_0, g_f \cap t)|}{|D(p_f, g_f \cap t) \cup D(I_0, g_f \cap t)|}. \tag{1}$$

Fig. 1 gives an example of reach planning by selecting n_f and searching for the shortest graph path to it.

The learned reach has several important qualitative properties in common with the early stages of human infant reach learning. First, reaching trajectories are notably jerky [12]. In our model, this is a consequence of motion taking place along paths in the PPS graph, which (even with 3000 nodes)
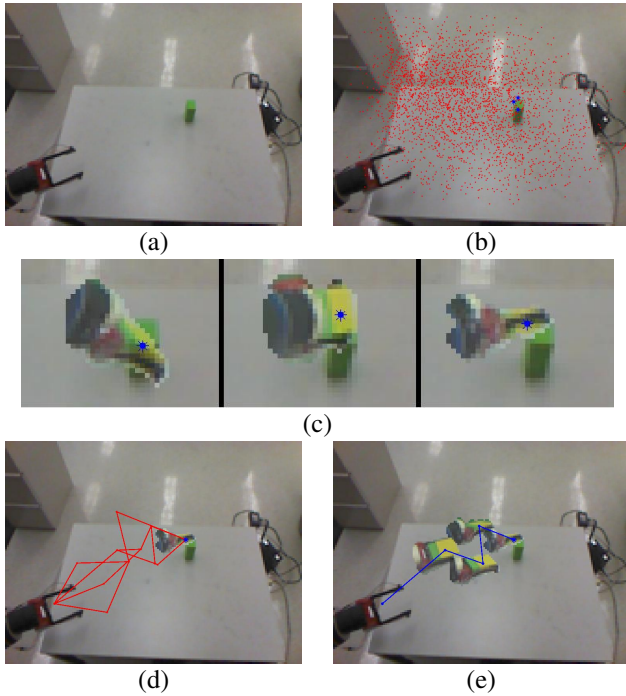
Fig. 1. **(a)** For a reaching task, the input to our agent is a single RGB-D image of the workspace. The agent compares this current percept to a stored image to segment the added target object. The agent can derive properties of the object in image space, but is not given the target's geometry, position in 3D space, or the arm configuration required to reach the target with the hand. **(b)** The agent has previously explored 3000 arm configurations, and stores each as a node in the PPS Graph. Each node is defined by the configuration of joint angles, and is annotated with the image of the hand in that configuration. Each node is represented by a point (usually red) overlayed on the input image (a), indicating the center of mass of the grippers in the image. The three larger blue points correspond to the candidate target nodes shown in (c). **(c)** The agent tests every node by finding the intersection of its stored image of the hand (shown grasping colored blocks, which simplified tracking) and the current image of the target object. The three nodes shown have nonempty intersections in the image plane and overlapping depth ranges within that intersection. All other nodes shown red in (b) fail to intersect in this way or are missing valid depth data. **(d)** The third candidate node from (c) has the largest intersection over union of depth ranges, the node feature found to most reliably predict reaching success. This node is selected as the target for the reach action. Several possible trajectories from the home node to the final node along PPS Graph edges are shown. **(e)** A graph search determines the shortest path in configuration space, which is not necessarily the shortest path in image space. With all joints weighted equally, this path minimizes the effort the agent expends during the reach. The agent chooses this trajectory, shown here with the stored images and gripper centers of mass. The similar gripper poses shown explain the low cost to traverse this path.

is a relatively sparse approximation of the seven-dimensional configuration space. Second, reaching is no less reliable under conditions when the agent cannot see its own hand during motion [2]. In our model, selection of the final node and the trajectory to reach it depends on current perception of the target object, but *not* on current perception of the hand. The hand images are collected from experience during motor babbling, and stored on the nodes of the PPS graph.

V. LEARNING TO GRASP

Once the agent has learned to reach reliably, events where an object is moved from one static position to another are no longer unusual. However, young children have the *Palmar*

reflex, where an object touching the palm of the hand causes the fingers to close, automatically grasping the object [3]. We have implemented a similar reflex on our Baxter robot, where a break-beam sensor between the fingers of the gripper causes the gripper to close.

The unusual event, then, occurs when an attempt to reach a block happens to trigger the Palmar reflex, and the hand closes so as to grasp the object, rather than simply moving it or knocking it over. This unusual situation is detected perceptually because when the hand moves, the grasped object moves with the hand, rather than remaining in a new static position. Grasping places an object that is not part of the agent, under the control of the agent, so it can be moved as naturally as the agent can move its own hand. That control is relinquished with an (easily learned) "ungrasp" action, leaving the object in a static pose where it will remain until moved again. Just as before, the agent identifies an unusual action, and seeks to learn how to make that action reliable.

A. Select Among Nodes in the Target Neighborhood

While learning to reach, the attention of the agent was focused on the final node of the trajectory. For learning to grasp, the important consideration is the final edge, and whether three conditions are met: (1) the gripper must be open; (2) along the final edge, the hand must approach the object with the open gripper facing the direction of motion; and (3) the wrist must be rotated so that the open gripper is transverse to the main axis of the block to be grasped. These conditions on the pose of the hand ensure that the hand is "pre-shaped" for the intended grasp. The agent must learn to concretely apply these abstract conditions in its motion planning, from unguided experience, just as the conditions on the target node were learned when learning to reach.

To select the penultimate node n_p and the final edge $e(n_p, n_f)$ along which the trajectory should reach a final node n_f, the agent must consider the immediate neighborhood of the final node: the (reasonably small) set of nodes $n_i \in N(n_f)$ that are linked to the final node n_f by a single edge $e(n_i, n_f)$. The learner seeks a feature that determines which final edge $e(n_i, n_f)$ is most likely to result in a successful grasp when the hand reaches n_f.

B. Predicting Favorable Alignments with Cosine Similarity

For a successful grasp, it is important that the opening between the gripper fingers is the first part of the hand to reach the object. If any other part of the hand leads the approach, the fingers or wrist will bump the target object out of the way before it can be grasped. Using subscripts p and f for features of n_p and n_f, we define vectors in image (u, v, d)-space to describe the alignment between these nodes and the target object, as applied in Fig. 2:

$$
\begin{aligned}
v_p &= \text{vector parallel to the gripper fingers in } p_p \\
v_f &= \text{vector parallel to the gripper fingers in } p_f \\
v_{pf} &= c_f - c_p \\
v_{pt} &= c_t - c_p \\
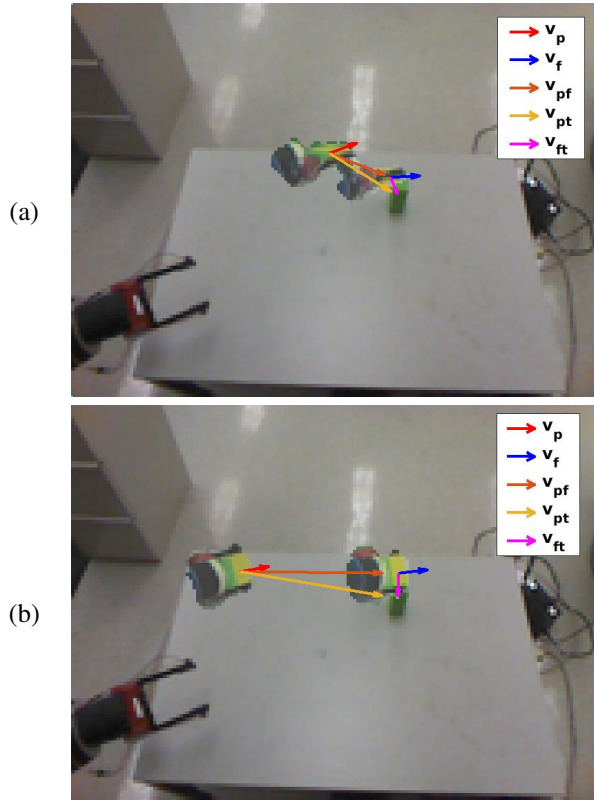v_{ft} &= c_t - c_f
\end{aligned}
\tag{2}
$$

Fig. 2. **(a)** A baseline grasping method uses the reach trajectory as determined in Fig. 1, approaching the target with the same move $e(n_p, n_f)$, shown by the endpoints' stored hand images. Also shown are the two-dimensional components of the four predictive vectors and v_{ft}, which is not useful for predicting the success of a grasp. In this case, the hand orientations are similar at both nodes, and the direction of motion from the penultimate node to the final node and target are similar, but there is little agreement between the orientation vectors and direction of motion vectors. As a result, a gripper is unlikely to face the target object as it approaches and will bump it instead. **(b)** The agent can improve its success rate by selecting a different trajectory such that the alignment of the final move is better. The final node of this trajectory may be any of the candidate nodes n_j predicted to interact with the target (as in Fig. 1(c)). Each n_j has a set of candidate penultimate nodes $n_i \in N(n_j)$ such that n_i is not also a final node candidate. The agent calculates (3) for each pair of candidate penultimate and final nodes. The nodes shown here have the maximal value of 0.94 for the given target object, compared to a value of 0.61 for those in (a). The high agreement between the four vectors suggests the opening between the gripper fingers is likely to lead the motion, which is a necessary condition for a grasp. The agent chooses to end the trajectory with a move between these maximal value nodes, and the rest of the trajectory is given by the shortest path in configuration space between the home node and n_p.

After a number of attempted grasp trajectories, we find the pairwise cosine similarities between these vectors for each. For pairs between any of the first four vectors, the cosine similarities from trials that produced successful grasps cluster near 1, with the cosine similarities from unsuccessful trials significantly lower and more spread out. The result that high cosine similarities correlate with successful grasps can be explained intuitively. Similar v_p and v_f imply minimal rotation of the gripper opening during the motion, and the gripper opening will lead the motion when these also align with v_{pf}. If v_{pf} also agrees with v_{pt}, motion along the edge will move directly toward the target. So with all four vectors aligned, a grasp attempt should have a higher chance of

success. However, for all pairs involving the fifth vector v_{ft}, the distribution of cosine similarities was not significantly different for successful and unsuccessful trials. We hypothesize that these cosine similarities are not informative due to the inconsistent relationship between c_f and the point of contact where the grasp occurs. If the grasp occurs past c_f at the very end of the motion, v_{ft} often agrees with the other vectors, but if it occurs earlier, v_{ft} points back toward c_t and may have a dissimilar or even opposite direction.

The results from the initial set of grasps indicate that a grasp trajectory with maximal cosine similarity between each pair in $\{v_p, v_f, v_{pf}, v_{pt}\}$ will have the highest chance of success. It is further observed that approaches with a single low cosine similarity often fail, even if the others are near 1. By contrast, if all cosine similarities are only moderately high, the grasp may still succeed. Therefore, the agent will choose from the set of reach trajectories the trajectory with the final edge that maximizes the minimum cosine similarity between any of the six informative pairs of vectors. That is, n_p and n_f are chosen to maximize

$$\min(\text{cosine\_similarity}(v_1, v_2)) \qquad (3)$$

for all $v_1, v_2 \in \{v_p, v_f, v_{pf}, v_{pt}\}$ and $v_1 \neq v_2$.

C. Estimating the Local Jacobian at a PPS Graph Node

The Jacobian describes the relationship between changes in the arm's configuration and changes in the observed center of mass of the hand in image space (u, v, d)-coordinates, that is, between Δq and Δc. This relationship is nonlinear and dependent on the current configuration, making the global model difficult to learn and use. However, the relationship at any particular configuration can be modeled by a linear approximation, and within the neighborhood of that configuration the error of the model is generally small enough that the predicted change and actual change have the same qualitative result, such as a bump or miss. A method for making a local Jacobian estimate at a configuration stored in a PPS Graph node follows. In practice, this method provides an estimate at any configuration, as the estimate for the nearest configuration in a node is sufficient.

Local Jacobian estimates can be calculated from information stored in the PPS Graph. For a node n_i and each $n_j \in N(n_i)$, the edge $e(n_i, n_j)$ provides a training example of a change in joint configuration labeled with its change in image coordinates. The information for all m edges can be grouped into an $m \times 7$ matrix ΔQ of joint angle changes between neighboring nodes and an $m \times 3$ matrix ΔC of image coordinate changes between the same pairs of nodes. We represent the local Jacobian with a 7×3 Matrix $J(n_i)$ and can find its least squares solution $\hat{J}(n_i)$ in $\Delta Q \hat{J}(n_i) = \Delta C$. In order for a solution to be defined, n_i must have $m \geq 7$ neighbors and valid depth information. We denote the set of all such nodes N_J, with $|N_J| = 2720$.

D. Inverse Jacobian Estimate for the Target Configuration

With an adequate model of the local structure of the configuration space, the agent can estimate how to perturb

the joint angle vector to more accurately reach to a pose that is near, but not exactly on, a node in the PPS graph. A particular application of this model allows the agent to estimate the configuration that corresponds to the center of mass of the target object and reach there from the penultimate node. Once a trajectory passing through $n_p \in N_J$ is chosen, the agent finds v_{pt}, an indication of the necessary image-space change to move the hand's center onto the target's center. $\hat{J}(n_p)$ is estimated by the above method, and its right psuedo-inverse $\hat{J}(n_p)^{-1}$ is computed. Multiplying this inverse local Jacobian by a desired change in image space predicts the change in joint angles needed to produce that change. Therefore, the agent can more precisely reach for the target's center by moving from q_p to a target configuration

$$q_p + \Delta q = q_p + v_{pt}\hat{J}(n_p)^{-1} \qquad (4)$$

instead of the original final configuration q_f given by the final node.

VI. EXPERIMENTS

A. Experiment 1 - Learning the Gripper Open Requirement

While the alignment and orientation requirements are interdependent and also influenced by the full configuration of the arm, holding the gripper open at the beginning of the grasp is independent, and involves only one degree of freedom. For this reason, the openness requirement is a good first learning goal. Note that it is not necessary to study this requirement in isolation, and infants will often make use of their longer available training time to observe multiple phenomena and task components at once. In this work we motivate our agent to consider only this requirement first, in order to greatly reduce the number of trials necessary to observe a number of successful grasps and draw conclusions.

Despite the low probability of an accidental grasp, a small number occurred while our agent learned to reach with the Palmar reflex present. Returning to these reach trials provides the agent with suitable trajectories for grasps, including sufficiently good values for the alignment and orientation conditions on the final move. To study the effect of the gripper degree of freedom on the probability of a successful grasp, these trajectories are repeated, varying the percent openness while holding all other conditions constant. The results of this experiment are shown in Fig. 3.

We can observe that all settings less than 70% open prevent any grasps from being repeated. On the low end of this range, the opening is smaller than the width of the object, and closer to 70%, fitting the object between the gripper fingers instead of knocking it over requires implausible precision for movements within the PPS Graph. Setting the gripper 100% open allows all grasps experienced thus far to be consistently repeated, and is significantly more reliable than all other settings. The unexpected spike in performance at 80% suggests that there may be some special case grasps where a smaller initial setting is at least equally viable, but at this time the agent will attempt all new grasps with the gripper fully open during the approach.
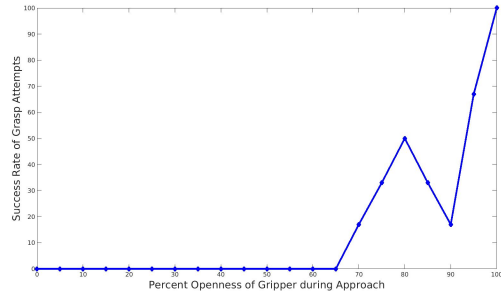

Fig. 3. The agent investigates the effect of the value of the gripper degree of freedom on the reliability of a grasp. Trials that have previously produced a grasp are repeated with gripper settings between 0% open and 100% open at 5% intervals, keeping the trajectory and target position constant over all repetitions. While the success rate does not monotonically increase with gripper openness, it is clear that a grasp is not possible if the gripper is too far closed during the approach. Further, only the fully open setting was capable of repeating every successful grasp in this experiment. The agent chooses the 100% open setting as the most reliable for grasping, and will use it in all trials in the remainder of this paper.

B. Experiment 2 - Selecting Well-Aligned Final Moves

Our agent uses two new methods for improved grasp reliability over the baseline of reusing the reaching trajectory. In the Cosine Similarity method (Fig. 2), the trajectory with the most aligned penultimate and final nodes is chosen according to a maximal value for (3). This method will not necessarily maximize (1) with its choice of n_f, so the chance of an interaction with the target is predicted to decrease, but when interactions occur, they are more likely to be grasps and not merely bumps. In the Cosine Similarity and Local Jacobian method, the agent starts with the trajectory chosen by the Cosine Similarity method and then estimates the target configuration by (4). This trajectory moves normally along PPS Graph edges from the home node to n_p, but the final motion sets the arm to this target configuration instead of moving to n_f. Moving the center of the hand to the center of the object increases the chance the gripper will surround the object for a grasp, and reduces the chance of a miss.

The results for each method are presented in Fig. 4. We can observe the expected small decrease in reaching performance when using only the Cosine Similarity method. The grasp success rate is not higher for this method than the baseline, but the agent does achieve a number of Palmar Bumps, where the Palmar reflex is triggered as the object goes between the grippers, but the object is bumped away. These near miss grasps suggest attention to the alignment criteria has allowed the agent to perform better, but not well enough for full success. The Cosine Similarity and Local Jacobian method combines this better alignment with precise positioning and produces the most successful grasps, 4.5 times the number achieved by the baseline. Fig. 5 shows the results of individual trials. We hypothesize that the central region where grasping is currently most reliable exists primarily due to the higher density of the PPS Graph near the center, providing more options for n_p and additional information for estimating the target configuration.

Method	Reach	Grasp	Palmar Bump
Reaching Baseline	90%	10%	0%
Cosine Similarity	82.5%	10%	12.5%
Cosine Similarity & Local Jacobian	97.5%	45%	22.5%

Fig. 4. Success rates at reaching and grasping using either the reaching trajectory or one of two methods for selecting a trajectory more likely to meet the prerequisites for a successful grasp. The first column gives the percent of attempted grasp trials that successfully reach to the target object and interact with it in some way (including bumps, Palmar bumps, and grasps). The second and third columns give the percent of all trials where a reach is completed and the interaction is a successful grasp or a near miss Palmar Bump, respectively. Using both the local Jacobian and cosine similarity features yields the best performance in all three categories.
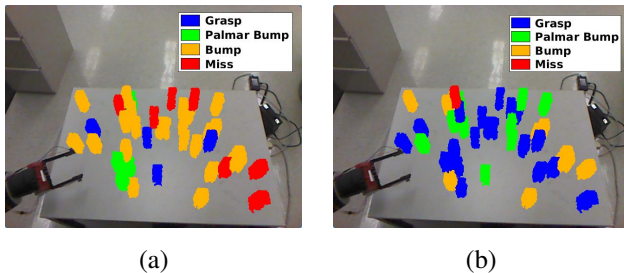

(a)   (b)

Fig. 5. **(a)** The 40 random positions of the target object as seen by the agent, color coded according to the result of the grasp attempt using the Cosine Similarity method to choose a trajectory with well-aligned penultimate and final nodes. Prioritizing alignment over the best position for the final node causes an increased number of failed reaches, with most of these misses near the edges of the workspace. **(b)** The results for the same 40 positions when using the Cosine Similarity and Local Jacobian method. This method allows the agent to choose these nodes and then refine the final joint configuration of the trajectory, which improves the rate of successful reaches and especially improves the rate of successful grasps.

VII. Future Work

The 40 positions and results shown in Fig. 5(b) allow the construction of a learning problem. Can the agent identify new features common to the successful or unsuccessful attempts that allow it to correct some of its mistakes? We will investigate possible transfers of the appropriate wrist joint angle for gripper orientation from a past success to new attempts or retries of past failures, especially if the target positions are nearby. Once the grasp action is reliable, it may be used in symbolic planning for larger tasks.

We have also done preliminary work to allow the agent to reach for targets in a smooth motion using inverse local Jacobians, rather than in jerky submotions along PPS Graph edges. Initial tests show that this method causes the hand to smoothly converge to the target. We will investigate the increase in efficiency and methods to incorporate the grasp prerequisites in smooth motion.

VIII. Conclusions

We present an updated Peri-Personal Space Graph that can be built through autonomous motor babbling and only requires input from a single RGB-D camera. This model supports the selection of desirable goal arm states for reaches and grasps, and planning trajectories to reach those goals without geometric or kinematic models. An agent with access to the PPS Graph and the Palmar reflex learns to apply two of three prerequisites to make its grasps more reliable.

To learn to keep the gripper open, our agent repeats previously successful grasps with its full range of gripper settings. Successful grasps begin to appear with the gripper 70% open, but this boundary value allows only 17% of the grasps to be repeated. Only when the gripper is set 100% open do all of the grasp repetitions succeed. Satisfying the requirement that the gripper faces the target during the approach depends on the entire configuration of the arm, a prohibitively large space for search. The set of motions defined by edges to candidate final nodes is much smaller. The edge with the highest similarity of image-space vectors representing alignment can be selected, and provides the best aligned penultimate and final nodes. Perturbing the final node to an estimate of the target's location in configuration space improves accuracy of the approach. Trajectories using this final motion and an open gripper successfully grasped the target at 45% of previously unseen locations, up to a total of 67.5% reliability once near misses can be corrected for.

References

[1] G. Baldassarre and M. Mirolli, editors. *Intrinsically Motivated Learning in Natural and Artificial Systems*. Springer, 2013.
[2] R. K. Clifton, D. W. Muir, D. H. Ashmead, and M. G. Clarkson. Is visually guided reaching in early infancy a myth? *Child Development*, 64(4):1099–1110, 1993.
[3] Y. Futagi, Y. Toribe, and Y. Suzuki. The grasp reflex and Moro reflex in infants: Hierarchy of primitive reflex responses. *International Journal of Pediatrics*, 2012(191562), 2012. doi:10.1155/2012/191562.
[4] M. Hersch, E. Sauser, and A. Billard. Online learning of the body schema. *Int. J. Humanoid Robotics*, 5(2):161–181, 2008.
[5] L. Jamone, L. Natale, F. Nori, G. Metta, and G. Sandini. Autonomous online learning of reaching behavior in a humanoid robot. *Int. J. Humanoid Robotics*, 9(3):1250017, 2012.
[6] J. Juett and B. Kuipers. Learning to reach by building a representation of peri-personal space. In *IEEE/RSJ Int. Conf. Humanoid Robots*, 2016.
[7] L. E. Kavraki, P. Svestka, J.-C. Latombe, and M. H. Overmars. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Trans. Robotics and Automation*, 12(4):566–580, 1996.
[8] J. Modayil and B. Kuipers. The initial development of object knowledge by a learning robot. *Robotics and Autonomous Systems*, 56:879–890, 2008.
[9] J. Mugan and B. Kuipers. Autonomous learning of high-level states and actions in continuous environments. *IEEE Trans. Autonomous Mental Development*, 4(1):70–86, 2012.
[10] D. M. Pierce and B. J. Kuipers. Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92, 1997.
[11] J. Sturm, C. Plagemann, and W. Burgard. Unsupervised body scheme learning through self-perception. In *Int. Conf. Robotics and Automation (ICRA)*, 2008.
[12] E. Thelen, D. Corbetta, K. Kamm, J. P. Spencer, K. Schneider, and R. F. Zernicke. The transition to reaching: mapping intention and intrinsic dynamics. *Child Development*, 64(4):1058–1098, 1993.