Reaching and Grasping in Peripersonal Space: Extended Abstract

Jonathan Juett[1] and Benjamin Kuipers[1]

We present a model by which an autonomous agent may construct reach and grasp actions, and learn features to make these actions more reliable. The model is evaluated with a physical Baxter Research Robot (pictured in Fig. 1e). The resulting behaviors may be compared qualitatively with early human infant manipulation actions, thanks to existing studies on infant learning and behavior (e.g. [2] for reaching).

Our model is driven by intrinsic motivation [12], [1], where the reward signal is based on the rate of learning or improvement in prediction or actions. New and unusual events perceived by the agent present opportunities to learn at a much higher rate than typical events that are already well explained by the current model. The agent will receive the highest intrinsic reward over time by repeating this procedure:

1) Practice the most recently learned action on self-identified target objects for several trials.
2) Identify a small group of observed results with a qualitative difference from the larger group. These cases will define a new unusual event.
3) Generate and test features of the executed trajectories and target position that best classify the results as unusual or usual.
4) Identify values of the selected features that make future attempts to repeat the unusual event most reliable. These will define the prerequisites and policy for a new action to cause the event purposefully.
5) Continue evaluation to refine each feature's ideal range of values. As the agent becomes more skilled and the action is performed reliably, the reward for observing the event it causes decreases. This makes way for a new unusual event, often a subset of, or related to, the previous event.

Our agent begins by constructing a graph representation of its environment, the *Peripersonal Space (PPS) Graph* by recording states visited during motor babbling. Nodes of this graph store the proprioceptive and visual states of the hand at the time of recording, with a vector of arm joint angles and segments of of RGB-D image, respectively. The edges of the PPS Graph connect nodes that have a feasible move between

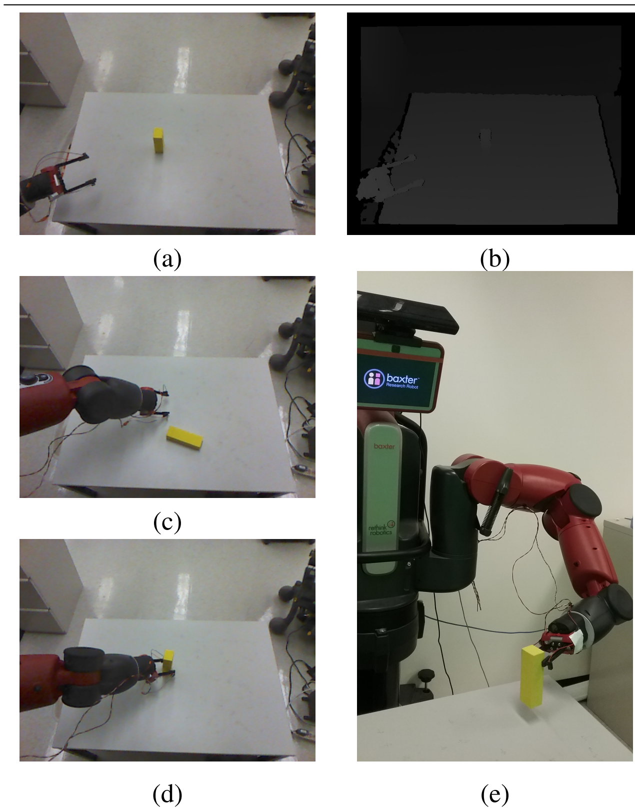

Fig. 1. **(a)** The RGB image portion of the visual percept at the beginning of a reaching or grasping trial. In this work all target objects are upright rectangular blocks, and can be identified by their distinct color. **(b)** The depth image portion of the visual percept taken at the same time as (a). The combination of disparity values in this image and the RGB image approximate human stereo vision. The agent stores prior percepts of the hand in various poses, but does not receive any other visual or geometric information for this task. **(c)** The agent's RGB percept after executing a successful reach. The object has been bumped to a new quasi-static location where it will remain as the agent returns the arm to the home position. **(d)** By revising the reach trajectory to satisfy learned prerequisites, the agent can grasp the object instead. This is the agent's RGB percept after initiating a grasp. **(e)** The Baxter Research Robot used to evaluate our learning model. In this image, the agent has returned the arm to the home position after initiating the grasp in (d) and observed the target object following the hand along this trajectory. It classifies this grasp attempt as successful.

This work has taken place in the Intelligent Robotics Lab in the Computer Science and Engineering Division of the University of Michigan. Research of the Intelligent Robotics lab is supported in part by a grant from the National Science Foundation (IIS-1421168).

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

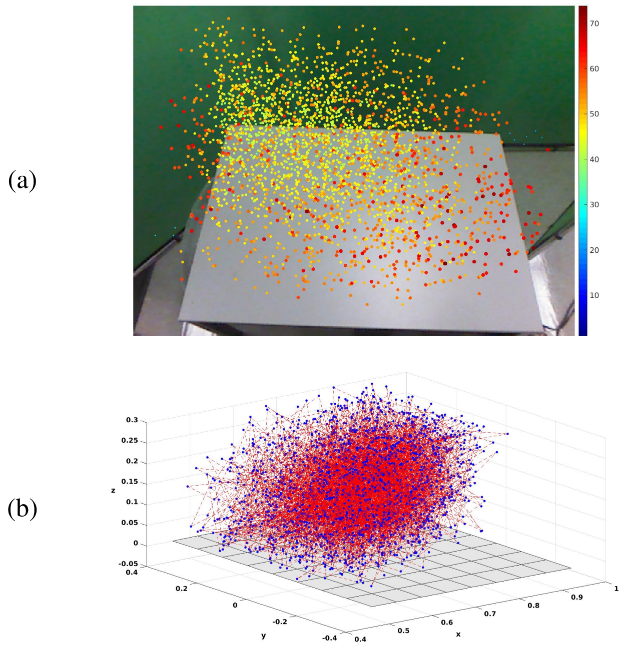

Fig. 2. **(a)** The Peripersonal Space (PPS) Graph shown in terms of the agent's visual space. Each marker superimposed with this base RGB percept corresponds to the (u, v) center of mass location of one of the 3000 nodes, with color and size scaled by the mean disparity value d of the node. **(b)** The PPS Graph in the default frame of reference for the robot, which is unavailable to the agent. Each blue point shows the true world (x, y, z) coordinates of a node, and each dotted red line corresponds to an edge of the graph. The 2999 edges shown were all proven by construction to be feasible moves during motor babbling, while additional edges added according to a length threshold for feasibility are not displayed here. For comparison, the surface of the table in the workspace is plotted as a gray plane. It can be observed that the surface and space above it are generally well-covered by the graph. As might be expected, one of the regions where nodes are more sparse is the corner opposite the natural position of the agent's left hand, and acting here is more difficult for the agent.

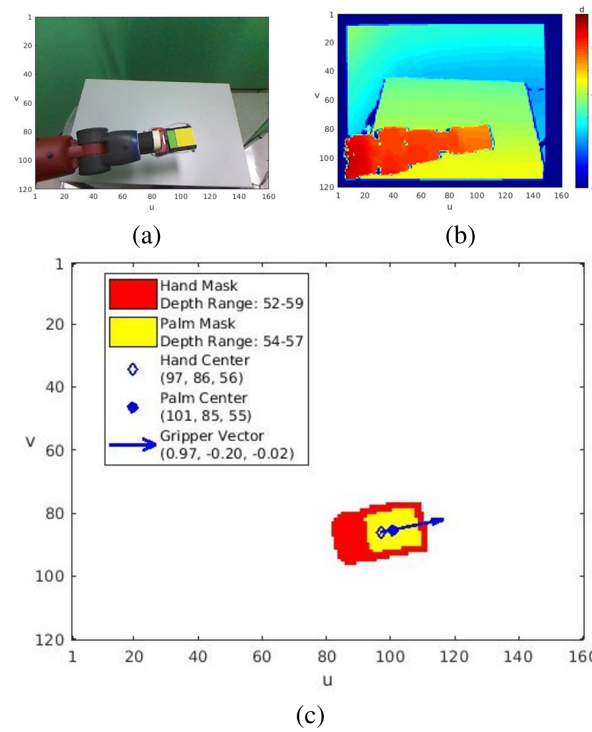

Fig. 3. **(a)-(b)** The agent's vision system provides an RGB image and a concurrent registered depth image in each percept. During motor babbling, these images are recorded along with the joint angles of a node. **(c)** The full representation the agent will store for this node, derived from the recorded images. The representation consists of binary masks for the full hand (red and yellow regions) and its grasping region (yellow), as well as the center of mass for each region and an estimated direction of the grippers. When planning a move, reach, or grasp, the PPS Graph serves as a mapping from a desired visual state to the configuration necessary to visit that state. The agent searches through the representations of all nodes for the one best fit to the task, and can look up the configuration recorded for that node.

them, and may be traversed by linear interpolation of the stored configurations of the endpoint nodes. Trajectories for moving the hand may be composed from these edges, and provide the agent with an initial *move(hand)* action. Fig. 2 provides two visualizations of the PPS Graph, and Fig. 3 gives an example of the state representation stored for a single node.

As the agent practices this move action, the typical result is for the appearance of the hand to change without a significant effect on the rest of the visual field. However, this static background model is violated in the rare event that the hand collides with another foreground object. This *bump* event can be detected by visually observing the object in a new quasi-static location. The agent learns that bumps are highly correlated with motions to a node where the stored image of the hand intersects the image of the target at the beginning of a trial. This is especially true when the intersection is in both the RGB image segments and the range of values for those pixels in the depth image. The *reach* action is defined as motion along a graph trajectory ending with a node with this property. A successful reach causes a bump, as in Fig. 1a-c. Our agent has nearly mastered reaching, with a success rate of 97%. We discuss an earlier process of learning to reach with an alternative visual system in [8].

Once the agent has a reliable reach action, it interacts with foreground objects much more often. In a few cases, this interaction places the object between the gripper fingers. Human infants have a *Palmar reflex* where the hand will automatically close around an object that touches the palm. Our robot does not have tactile sensors, but we have simulated the Palmar reflex with a break beam sensor. When an object passes between the gripper tips and breaks the beam, the grippers close. Both the real and simulated Palmar reflexes allow these rare cases with an object in the hand to cause an accidental grasp without an additional decision to close the hand at that precise moment. Removing one event from the necessary sequence does not make grasps common, but prevents them from being so rare that they cannot be observed and learned from. Our agent identifies *grasps* (Fig. 1d-e) by the new property of the object to move along with the hand until it is *ungrasped*, returning the object to a quasi-static state.

Learning a grasp action requires a much larger number of considerations than learning a reach action. At this time, our agent has improved the reliability of its grasp action in four ways. First, comparing hand images for a neighborhood of nodes in the PPS Graph allows estimation of adjustments to the final node configuration that will improve its intersection with the target object. This and continued practice with reaching have improved the accuracy of the agent's approach. Second, the agent has learned that the grippers are closed by the Palmar reflex more often if they are initially fully open to better allow the object to be surrounded. Collision with the object may still prevent a successful grasp, but increasing the frequency of Palmar reflex activations tends to increase the number of grasps. Third, the agent draws and evaluates specific vectors on the stored hand images to select a well-aligned penultimate node for the grasp trajectory. This is important so that the gripper opening leads the motion to the final node, avoiding a bump with the exterior of the hand that would prevent a grasp. Finally, the agent orients the wrist perpendicular to the major axis of the target object, attempting to ensure the gripper opening is wider than the cross section of the object it attempts to grasp. We discuss these prerequisites for grasping in more detail in [9]. With all of these features considered, our agent now grasps the object from previously unseen positions approximately 50% of the time. That grasping is much more difficult than reaching is understandable, and a focus of future work will be to improve this reliability. After grasping is suitably reliable, the agent may repeat the learning process to discover additional actions, such as specific ungrasps to set the object in a stable position, and a *move(target)* action to combine all of these capabilities.

Related Work

Some robotics researchers (e.g., [4], [13]) focus on learning the kind of precise model of the robot that is used for traditional forward and inverse kinematics-based motion planning. A goal of our model is to make much weaker assumptions about the variables and constraints in the model, and the information available from visual perception.

Other researchers (e.g., [10], [11]) structure their models according to hypotheses about the neural control of reaching and grasping, with constraints represented by neural networks that are trained from experience. Oztop, Bradley & Arbib [10] use a simulated robot arm and hand, focusing on learning an open-loop controller that is likely to terminate with a successful grasp, but assuming that reaching is already programmed in. Savastano and Nolfi [11] describe an embodied computational model with three developmental phases, implemented as a recurrent neural network, and evaluated on a simulation of the iCub robot. However, the transitions from one phase to the next are represented by manually adding certain links and changing certain parameters in the network, begging the question about how and why those changes take place.

Several recent research results are closer to our approach, in the sense of focusing on sensorimotor learning without explicit skill programming, exploration guidance, or labeled training examples. Each of these (including ours) makes simplifying assumptions to support progress at the current state of the art, but each contributes a "piece of the puzzle" for learning to reach and grasp.

Hülse, et al [6] present a very nice review of the (2010) neuroscience of visual search and reaching, and of the development of saccades, visual search, and reaching in human infants. They then present and evaluate two architectures for combining these capabilities, and discuss their significance for complete theories of reach learning.

Jamone, et al [7] define a Reachable Space Map, describing the learned reachability of fixated objects over gaze coordinates (head yaw and pitch, and eye vergence (to encode depth)). Within our framework, the PPS Graph [8] is learned during non-goal-directed motor babbling, at a developmentally earlier stage of knowledge, before goal-directed reaching has a meaningful chance of success.

Ugur, et al [14] demonstrate autonomous learning of behavioral primitives and object affordances, leading up to imitation learning of complex actions. However, they start with the assumption that peripersonal space can be described by a 3D Euclidean space model. Our agent starts with only the raw proprioceptively sensed joint angles in the arm, and the 2D images provided by vision sensors. The PPS graph represents a learned mapping between those spaces.

M. Hoffmann, et al [5] integrate empirical data from infant experiments with computational modeling on the physical iCub robot. Their model includes haptic and proprioceptive sensing, but not vision. They model the processes by which infants learn to reach to different parts of the body, prompted by buzzers on the skin. The model is implemented and evaluated on an iCub robot with artificial tactile-sensing skin. However, the authors themselves describe their success as partial, observing that the empirical data, conceptual framework, and robotic modeling are not well integrated. They aspire to implement a version of the sensorimotor account, but they describe their actual model as much closer to traditional robot programming.

Conclusion

Our agent has built a graph representation for peripersonal space from autonomous exploration which allows simple trajectory planning. These trajectories share qualities with early human movements, such as jerky submotions. In our model, these appear due to movements to discrete, memorized configurations stored in nodes along the shortest graph path. It has then applied an intrinsically motivated process twice, first to learn a reach action to replicate the unusual bump event, and then to learn to select reach trajectories that are more likely to also grasp a target object. These actions also have infant-like qualities, including the lack of dependence on a current percept of the hand for successful reaching (observed in [3], where infants reached for lit targets in a dark room), and the necessity of preshaping before a grasp. In our most recent results, the learned reach action is almost perfectly reliable, and the grasp action succeeds approximately 50% of the time.

REFERENCES

[1] G. Baldassarre and M. Mirolli, editors. *Intrinsically Motivated Learning in Natural and Artificial Systems*. Springer, 2013.

[2] N. E. Berthier. The syntax of human infant reaching. In *Unifying Themes in Complex Systems: Proc. 8th Int. Conf. on Complex Systems*, volume VIII of *New England Complex Systems Institute Series on Complexity*, pages 1477–1487. NECSI Knowledge Press, 2011.

[3] 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.

[4] M. Hersch, E. Sauser, and A. Billard. Online learning of the body schema. *Int. J. Humanoid Robotics*, 5(2):161–181, 2008.

[5] M. Hoffmann, L. K. Chinn, E. Somogyi, T. Heed, J. Fagard, J. J. Lockman, and J. K. O'Regan. Development of reaching to the body in early infancy: From experiments to robotic models. In *IEEE Int. Conf. Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2017.

[6] M. Hülse, S. McBride, J. Law, and M. Lee. Integration of active vision and reaching from a developmental robotics perspective. *IEEE Trans. Autonomous Mental Development*, 2(4):355–367, 2010.

[7] L. Jamone, M. Bradao, L. Natale, K. Hashimoto, G. Sandini, and A. Takanishi. Autonomous online generation of a motor representation of the workspace for intelligent whole-body reaching. *Robotics and Autonomous Systems*, 62(4):556–567, 2014.

[8] J. Juett and B. Kuipers. Learning to reach by building a representation of peri-personal space. In *IEEE/RSJ Int. Conf. Humanoid Robots*, 2016.

[9] J. Juett and B. Kuipers. Learning to grasp by extending the peri-personal space graph. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2018.

[10] E. Oztop, N. S. Bradley, and M. A. Arbib. Infant grasp learning: a computational model. *Experimental Brain Research*, 158(4):480–503, 2004.

[11] P. Savastano and S. Nolfi. A robotic model of reaching and grasping development. *IEEE Trans. Autonomous Mental Development*, 5(4):326–336, 2013.

[12] J. Schmidhuber. Curious model-building control systems. In *Proc. Int. J. Conf. Neural Networks*, pages 1458–1463. IEEE Press, 1991.

[13] J. Sturm, C. Plagemann, and W. Burgard. Unsupervised body scheme learning through self-perception. In *IEEE Int. Conf. Robotics & Automation (ICRA)*, 2008.

[14] E. Ugur, Y. Nagai, E. Sahin, and E. Oztop. Staged development of robot skills: behavior formation, affordance learning and imitation with motionese. *IEEE Trans. Autonomous Mental Development*, 7(2):119–139, 2015.