Empowering Robots for Object Handovers

Alap Kshirsagar
Cornell University
Ithaca, NY, USA
ak2458@cornell.edu

Guy Hoffman
Cornell University
Ithaca, NY, USA
hoffman@cornell.edu

***Abstract*—This work investigates the collaborative task of object handovers between a human and a robot, a central aspect of human-robot collaboration. Our research contributes along three directions: first, designing robot controllers for previously unexplored human-robot handover scenarios; second, investigating gaze behaviors of a receiver in human-to-human and human-to-robot handovers; and third, investigating human behavior in bimanual and multiple sequential human-to-human handovers. Our contributions could help enable robots perform the complex but essential tasks of handing over objects to and receiving objects from humans.**

***Index Terms*—physical human-robot interaction, social human-robot interaction, formal methods, reinforcement learning**

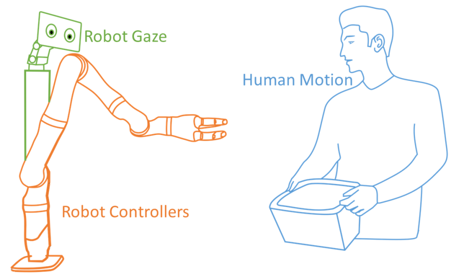

Fig. 1. Our work focuses on developing robot motion controllers and robot gaze behaviors for human-robot handovers, and investigating human motion in handovers.

I. Introduction and Research Questions

This project studies the collaborative task of object handovers between a human and a robot. People frequently perform object handovers with each other, and robots will be expected to perform such handovers with people, both at work and at home. Examples include a collaborative factory robot exchanging parts with a human co-worker, a surgical assistant robot transferring instruments to or from a surgeon, a warehouse robot helping a human shelve items, and a caregiver robot providing food or medicine to a patient. These tasks include both human-to-robot and robot-to-human handovers.

In this work, we seek to address three research questions related to object handovers as depicted in Figure 1:

- How to control a robot to perform handovers with humans? We focus on three previously unexplored handover scenarios: automated synthesis from high-level specifications, unknown robot dynamics, and human-like bimanual handovers.
- Which gaze behavior should a robot utilize while receiving an object from a human?
- How do humans perform bimanual and multiple sequential handovers with each other?

II. How to control robots to perform handovers?

The first research direction is the study of robot controllers in three previously unexplored scenarios:

Scenario 1: Synthesis from High-level Specifications

Existing robot controllers for object handovers [1]–[7] do not allow users to specify and dynamically change robot behaviors, or require users to tune non-intuitive controller parameters. Also, there is no unified framework to easily switch between the multitude of handover strategies that have been proposed in the literature. To overcome these shortcomings, we developed a controller for human-robot handovers that is automatically synthesized from high-level specifications in Signal Temporal Logic (STL) [8]. We developed specification templates both for human-to-robot and for robot-to-human handovers, and illustrated the flexibility of our approach by reproducing existing human-robot handover strategies found in the literature. Our controller allows end-users to specify and dynamically change the robot's behaviors using high-level requirements of goals and constraints that are intuitive. For example, they can specify the time taken by the robot to reach the handover location. This specification is then converted into a Model Predictive Control (MPC) problem to generate the robot's reactive motion in a handover.

Scenario 2: Unknown Robot Dynamics

Another drawback of existing handover controllers is that these methods rely on accurate models of the robot's dynamics and/or of the human kinematics, which may be difficult to obtain for custom built robots and for commercial robots with proprietary claims. Robot dynamics may also change due to different, and possibly unknown, masses of the objects to be handed over. Recently researchers have suggested a reinforcement learning (RL) algorithm called "Guided Policy Search" (GPS) [9] that does not require prior knowledge of the robot or environment dynamics. While GPS has been demonstrated on a number of autonomous manipulation and navigation tasks [9]–[11], it was not tested in a human-robot collaborative task such as a handover. We evaluated the potential of GPS to train a robot controller for human-robot object handovers [12]. Our results in a simulation environment indicate that GPS is limited in the spatial generalizability over variations in the target location, but that this issue can be mitigated with the addition of local controllers trained over target locations in the high error regions. Moreover, learned policies generalize well over a large range of end-effector masses. Moving targets can be reached with comparable errors using a global policy trained on static targets, but this results in inefficient, high-torque trajectories. Training on moving targets improves trajectories, but results in worse worst-case performance.

Scenario 3: Human-like Bimanual Handovers

There are numerous applications where bimanual handovers are useful or necessary. Bimanual handovers are necessary when handing over large rigid objects such as crates, deformable objects such as folded clothes, spherical objects such as a basketball, and delicate objects like crockery. Also in some cultures it is a rule of etiquette to hand over objects with two hands. However, there is very little work on robot controllers for bimanual human-robot handovers [13]–[16]. Further, these existing approaches do not take into account the human's preferences for the handover configurations and the robot's handover motion is not human-like. For single handed or uni-manual handovers, several researchers have proposed robot controllers for human-like reaching motions [17]–[21]. But, to the best of our knowledge, there is no prior work on human-like robot motion in bimanual handovers. We seek to address this gap by developing robot controllers for human-like bimanual handovers. We will explore the application of movement primitives such as Pro-MPs [18] for this purpose.

III. Which gaze behavior should a robot utilize while receiving an object from a human?

Our second research direction concerns robot gaze behaviors when receiving an object from a human. Past research has shown that the robot's head gaze behaviors affect the subjective experience and timing of robot-to-human object handovers [2], [22]–[24]. However, all of these works only studied robot-to-human handovers, where the robot was the giver. Human-to-robot handovers, where robot is the receiver, are equally important with many applications in various domains. Therefore, we investigated robot gaze behaviors in human-to-robot handovers [25], [26].

To find candidates for robot gaze behaviors, we analyzed gaze behaviors of human receivers in human-to-human handovers by annotating gaze locations in over 14000 frames of a public data-set of handovers [27]. We then implemented and compared four human-inspired robot head gaze behaviors during the reach phase of human-to-robot handovers. Results revealed that observers of a handover perceived a *Face-Hand transition* gaze, in which the robot initially looks at the giver's face and then at the giver's hand, as more anthropomorphic, likable and communicative of timing compared to continuously looking at the giver's face (*Face* gaze) or hand (*Hand* gaze). Participants in a handover perceived *Face* gaze or *Face-Hand transition* gaze as more anthropomorphic and likable compared to *Hand* gaze. However, these results were limited to a specific scenario of object handover, a common limitation of other prior studies as well. To expand and generalize the findings from our work [25], we studied human preferences towards robot gaze behaviors in human-to-robot handovers for four different object types (large, small, fragile, non-fragile) and two human postures (standing, sitting) [26]. Our results revealed that, for both observers and participants in a handover, when the robot exhibited a *Face-Hand-Face* gaze (gazing at the giver's face and then at the giver's hand during the reach phase and back at the giver's face during the retreat phase), participants considered the handover to be more likable, anthropomorphic, and communicative of timing. We did not find evidence of any effect of the object's size or fragility or the giver's posture on the gaze preference.

IV. How do humans perform bimanual and multiple sequential handovers with each other?

The final research direction studies bimnanual and sequential handovers. In the past, researchers have studied various aspects of human-to-human handovers to understand how people perform this complex maneuver. Some studies investigated different phases of handovers, such as approach [28], reach [29]–[31], and transfer [32], [33]. Others have built data-sets of human-to-human handovers [27], [34], [35]. However, all of these works have studied handovers of objects with single handed or uni-manual grasps. To the best of our knowledge there is no prior work studying human-to-human handovers of objects requiring bimanual grasps, even though such handovers are equally important and even more challenging. Further, several collaborative tasks require multiple sequential handovers. For example, shelving in a warehouse, automotive assembly, unloading dishes or groceries etc. In these tasks some objects need to be handed over with two hands, owing to their shape, size or fragility, while others can be handed over with one hand. Sometimes the giver and/or the receiver need/s to perform a self-handover i.e. transferring the object from one hand to another. When people perform such tasks, they effortlessly switch between these different types of handovers.

We propose to create a multi-sensor data-set of sequential uni-manual and bimanual object handovers between humans. We will ask participants to perform a shelving task involving object handovers with different types of objects, requiring uni-manual and/or bimanual grasps. We will record the trajectories followed by the giver and the receiver with a motion tracking system. We will also record the videos of handovers to obtain grasp configurations. Our data-set will be made public to conduct further research. It could help in designing controllers for robots to perform bimanual and multiple sequential handovers with humans.

V. Conclusion

We presented three directions of research related to human-robot object handovers. First, we developed robot controllers to address three previously unexplored scenarios of human-robot handovers. Second, we investigated a robot receiver's gaze in human-to-robot handovers. Third, we proposed to create a data-set of multiple sequential unimanual/bimanual human-human handovers that will provide new insights on human behavior in handovers.

REFERENCES

[1] K. Strabala, M. K. Lee, A. Dragan, J. Forlizzi, S. S. Srinivasa, M. Cakmak, and V. Micelli, "Toward seamless human-robot handovers," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 112–132, 2013.

[2] A. Moon, D. M. Troniak, B. Gleeson, M. K. Pan, M. Zheng, B. A. Blumer, K. MacLean, and E. A. Croft, "Meet me where i'm gazing: How shared attention gaze affects human-robot handover timing," in *ACM/IEEE International Conference on Human-robot Interaction*, 2014, pp. 334–341.

[3] C.-M. Huang, M. Cakmak, and B. Mutlu, "Adaptive coordination strategies for human-robot handovers," in *Robotics: Science and Systems*, Rome, Italy, 2015.

[4] S. M. zu Borgsen, J. Bernotat, and S. Wachsmuth, "Hand in hand with robots: differences between experienced and naive users in human-robot handover scenarios," in *International Conference on Social Robotics*. Springer, 2017, pp. 587–596.

[5] M. Cakmak, S. S. Srinivasa, M. K. Lee, J. Forlizzi, and S. Kiesler, "Human preferences for robot-human hand-over configurations," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 1986–1993.

[6] J. R. Medina, F. Duvallet, M. Karnam, and A. Billard, "A human-inspired controller for fluid human-robot handovers," in *IEEE-RAS 16th International Conference on Humanoid Robots (Humanoids)*, 2016, pp. 324–331.

[7] E. A. Sisbot and R. Alami, "A human-aware manipulation planner," *IEEE Transactions on Robotics*, vol. 28, no. 5, pp. 1045–1057, 2012.

[8] A. Kshirsagar, H. Kress-Gazit, and G. Hoffman, "Specifying and synthesizing human-robot handovers," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 5930–5936.

[9] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 1334–1373, 2016.

[10] S. Levine and P. Abbeel, "Learning neural network policies with guided policy search under unknown dynamics," in *Advances in Neural Information Processing Systems*, 2014, pp. 1071–1079.

[11] S. Levine, N. Wagener, and P. Abbeel, "Learning contact-rich manipulation skills with guided policy search," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2015. [Online]. Available: https://doi.org/10.1109/icra.2015.7138994

[12] A. Kshirsagar, G. Hoffman, and A. Biess, "Evaluating guided policy search for human-robot handovers," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3933–3940, 2021.

[13] A. M. Bestick, S. A. Burden, G. Willits, N. Naikal, S. S. Sastry, and R. Bajcsy, "Personalized kinematics for human-robot collaborative manipulation," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Sep. 2015. [Online]. Available: https://doi.org/10.1109/iros.2015.7353498

[14] S. Sina Mirrazavi Salehian, N. Figueroa, and A. Billard, "Coordinated multi-arm motion planning: Reaching for moving objects in the face of uncertainty," in *Proceedings of Robotics: Science and Systems*, 2016.

[15] S. S. M. Salehian, N. Figueroa, and A. Billard, "A unified framework for coordinated multi-arm motion planning," *The International Journal of Robotics Research*, vol. 37, no. 10, pp. 1205–1232, Apr. 2018. [Online]. Available: https://doi.org/10.1177/0278364918765952

[16] W. He, J. Li, Z. Yan, and F. Chen, "Bidirectional human-robot bimanual handover of big planar object with vertical posture," *IEEE Transactions on Automation Science and Engineering*, 2021.

[17] M. Prada, A. Remazeilles, A. Koene, and S. Endo, "Dynamic movement primitives for human-robot interaction: comparison with human behavioral observation," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2013, pp. 1168–1175.

[18] G. Maeda, M. Ewerton, R. Lioutikov, H. B. Amor, J. Peters, and G. Neumann, "Learning interaction for collaborative tasks with probabilistic movement primitives," in *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE, Nov. 2014. [Online]. Available: https://doi.org/10.1109/humanoids.2014.7041413

[19] A. Kupcsik, D. Hsu, and W. S. Lee, "Learning dynamic robot-to-human object handover from human feedback," *Robotics Research*, vol. 1, p. 161, 2017.

[20] K. Yamane, M. Revfi, and T. Asfour, "Synthesizing object receiving motions of humanoid robots with human motion database," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2013, pp. 1629–1636.

[21] W. Yang, C. Paxton, M. Cakmak, and D. Fox, "Human grasp classification for reactive human-to-robot handovers," *arXiv preprint arXiv:2003.06000*, 2020.

[22] M. Zheng, A. Moon, E. A. Croft, and M. Q.-H. Meng, "Impacts of robot head gaze on robot-to-human handovers," *International Journal of Social Robotics*, vol. 7, no. 5, pp. 783–798, 2015.

[23] K. Fischer, L. C. Jensen, F. Kirstein, S. Stabinger, Ö. Erkent, D. Shukla, and J. Piater, "The effects of social gaze in human-robot collaborative assembly," in *International Conference on Social Robotics*. Cham: Springer, 2015, pp. 204–213.

[24] B. Kühnlenz, Z.-Q. Wang, and K. Kühnlenz, "Impact of continuous eye contact of a humanoid robot on user experience and interactions with professional user background," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. Lisbon, Portugal: IEEE, 2017, pp. 1037–1042.

[25] A. Kshirsagar, M. Lim, S. Christian, and G. Hoffman, "Robot gaze behaviors in human-to-robot handovers," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 6552–6558, 2020.

[26] T. Faibish*, A. Kshirsagar*, G. Hoffman, and Y. Edan, "Human preferences for robot eye gaze in human-to-robot handovers," *International Journal of Social Robotics (in press)*, 2021.

[27] A. Carfì, F. Foglino, B. Bruno, and F. Mastrogiovanni, "A multi-sensor dataset of human-human handover," *Data in Brief*, vol. 22, pp. 109 – 117, 2019.

[28] P. Basili, M. Huber, T. Brandt, S. Hirche, and S. Glasauer, "Investigating human-human approach and hand-over," in *Human Centered Robot Systems. Cognitive Systems Monographs*, 2009, vol. 6, pp. 151–160.

[29] M. Huber, M. Rickert, A. Knoll, T. Brandt, and S. Glasauer, "Human-robot interaction in handing-over tasks," in *17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2008, pp. 107–112.

[30] S. Parastegari, B. Abbasi, E. Noohi, and M. Zefran, "Modeling human reaching phase in human-human object handover with application in robot-human handover," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 3597–3602.

[31] A. H. Mason and C. L. MacKenzie, "Grip forces when passing an object to a partner," *Experimental brain research*, vol. 163, no. 2, pp. 173–187, 2005.

[32] W. P. Chan, C. A. Parker, H. M. Van der Loos, and E. A. Croft, "Grip forces and load forces in handovers: Implications for designing human-robot handover controllers," in *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*. New York, NY, USA: ACM, 2012, pp. 9–16.

[33] S. Endo, G. Pegman, M. Burgin, T. Toumi, and A. M. Wing, "Haptics in between-person object transfer," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 2012, pp. 103–111.

[34] W. P. Chan, M. K. X. J. Pan, E. A. Croft, and M. Inaba, "An affordance and distance minimization based method for computing object orientations for robot human handovers," *International Journal of Social Robotics*, vol. 12, no. 1, pp. 143–162, Apr. 2019. [Online]. Available: https://doi.org/10.1007/s12369-019-00546-7

[35] R. Ye, W. Xu, Z. Xue, T. Tang, Y. Wang, and C. Lu, "H2o: A benchmark for visual human-human object handover analysis," 2021.