Eye Gaze for Assistive Manipulation

Reuben M. Aronson
rmaronson@cmu.edu
The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

Henny Admoni
henny@cmu.edu
The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania, USA

ABSTRACT

A key challenge of human-robot collaboration is to build systems that balance the usefulness of autonomous robot behaviors with the benefits of direct human control. This balance is especially relevant for assistive manipulation systems, which promise to help people with disabilities more easily control wheelchair-mounted robot arms to accomplish activities of daily living. To provide useful assistance, robots must understand the user's goals and preferences for the task. Our insight is that systems can enhance this understanding by monitoring the user's natural eye gaze behavior, as psychology research has shown that eye gaze is responsive and relevant to the task. In this work, we show how using gaze enhances assistance algorithms. First, we analyze eye gaze behavior during teleoperated robot manipulation and compare it to literature results on by-hand manipulation. Then, we develop a pipeline for combining the raw eye gaze signal with the task context to build a rich signal for learning algorithms. Finally, we propose a novel use of eye gaze in which the robot avoids risky behavior by detecting when the user believes that the robot's behavior has a problem.

CCS CONCEPTS

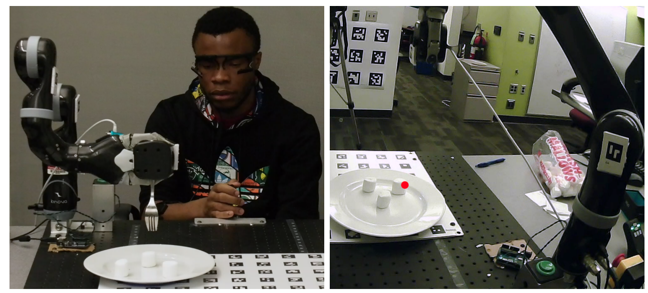
• **Human-centered computing → Collaborative interaction**; **Empirical studies in HCI**; • **Computer systems organization → Robotics**.

KEYWORDS

human-robot interaction, human-robot collaboration, physical assistance, shared control, eye gaze

ACM Reference Format:

Reuben M. Aronson and Henny Admoni. 2020. Eye Gaze for Assistive Manipulation. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3371382.3377434



(a) (b)

Figure 1: (a) A user performs a shared control task with a robot while wearing an eye tracker. (b) The user looks at the rightmost marshmallow; the dot indicates the user's gaze target.

1 INTRODUCTION

Assistive robotics is among the most promising practical applications for human-robot interaction to improve people's lives. Robot arms that mount on wheelchairs are used by people to perform activities of daily living today. However, these robots are typically operated by direct control. Teleoperating a robotic manipulator to perform a complex task is hard, especially when using low-dimensional input devices like joysticks.

HRI can help. Shared control methods [8, 11, 12, 18, 22] infer the operator's intended goal from their control input and combine this input with autonomous action towards that goal. We can enhance these systems by supplementing their inference from direct control with inference from natural, intuitive, indirect signals. These signals not only provide specific information about the user's intended goals, but can also give more information about the user's *mental state* during the task, which enables additional types of assistance.

One strategy for learning people's mental states during shared control is to monitor their eye gaze behavior. Gaze follows consistent patterns when individuals are performing specific tasks like walking [17] or manipulating objects [13, 16], and standard learning techniques including support vector machines [10], hidden Markov models [5], scanpath linguistic matches [15], and template matching [6] have been successful in understanding intent from an eye gaze signal. Moreover, these observations have been used to build collaborative systems that monitor gaze behavior to determine people's intentions during food serving [10] or handover [9] tasks. In addition, gaze is highly responsive to the situation; people can move their eyes to observe new data much faster than they can either move their hands or control a robot. Gaze behavior is a rich signal for understanding people's mental states.

In this project, we investigate how to use eye gaze signals for shared manipulator control. We begin by exploring where people look during teleoperated manipulation and compare their behavior with existing findings on eye gaze during by-hand manipulation. Next, we develop a pipeline for combining raw eye gaze data with the task and environment context to transform it into a signal

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-7057-8/20/03.
https://doi.org/10.1145/3371382.3377434

suitable for algorithmic assessment. Finally, we propose a novel application of eye gaze for assistance: a system that infers user discomfort by monitoring their eye gaze patterns and modifies its behavior appropriately. This research project shows how monitoring people's eye gaze improves shared control.

2 EYE GAZE DURING TELEOPERATED MANIPULATION

We begin by exploring patterns of human gaze during teleoperated manipulation in a user study [3]. Participants operated a Kinova Mico [14] robot arm to spear one of three marshmallows on a plate, and their gaze data was collected with a mobile eye tracker (Fig. 1). In the study, we found that people perform *anticipatory glances* towards their goals at specific times. While people mostly look at the robot's end-effector, people explicitly glance at their goal object before initiating different types of motion, especially before translation. In addition, they often alternate between looking at the end-effector and the goal during translational alignment. These patterns indicate that while gaze patterns during teleoperated manipulation differ from those during by-hand manipulation, they nevertheless remain informative about the operator's mental state.

People also displayed revealing eye gaze patterns when something went wrong. In another data set we collected and made available to researchers [19], we identified two categories of eye gaze behaviors that occurred when something went wrong in a task [1]. When the robot blocks the user's view of the goal, people move their heads for a better view. When the robot falls into a problematic kinematic configuration, such as a joint limit or collision preventing the robot from proceeding, users look at the problematic joint in question.This behavior shows that eye gaze can reveal not just people's intentions but also their view of the state of the task.

3 EYE GAZE PROCESSING PIPELINE

To use eye gaze for assistance, we must process the raw eye gaze signal into a usable form. Eye trackers typically report only a pixel location in an egocentric camera corresponding to where the user is looking. However, this contextless information is difficult to use directly. We observe that eye gaze during manipulation (both by-hand and teleoperated) is almost always directed at a specific, relevant scene object. Therefore, we augment the raw signal by matching each gaze fixation with a known object in the scene to determine at which object the user is looking (*semantic gaze labeling*). Then, we use timed, labeled fixations as inputs to a learning system. Typical approaches to semantic labeling use ray tracing to match the gaze location with scene objects [4, 20, 21]. To improve performance, we introduced velocity-based feature matching [2], which compares the relative motion of the gaze target over time with the motion of the objects in the scene. We show that this additional feature improves the robustness of the semantic gaze labeling procedure to constant errors such as those due to initial calibration. In addition, preliminary work shows that learning methods using this semantic gaze signal can predict task goal and failure conditions.

4 GAZE-BASED ASSISTANCE

Finally, we will apply the eye gaze analysis above to enhance robot assistance during shared control. First, we will build a model to predict the user's goal from their eye gaze behaviors. We can then compare eye gaze and joystick signals for intention recognition during assistance. In particular, eye gaze gives an absolute signal directed towards the ultimate goal, while joystick input only gives relative information about the direction towards the goal from the current state. Therefore, we expect that adding eye gaze will enhance the performance of shared control algorithms, and we will evaluate this claim in a user study.

In addition, we propose a new way to use people's eye gaze behavior for assistance. Existing psychology research [13] and our results both suggest that people look at aspects of a task that cause problems, such as obstacles that the robot must avoid. We hypothesize that when the user believes that the robot must be especially careful around a particular obstacle (due to sensing uncertainty, undetected obstacle properties like fragility, or user unfamiliarity with the system), the user will look more at that object. The robot can detect this eye gaze behavior and adjust its accordingly.

To validate this assistance behavior, we propose to develop a variant of the morsel spearing task, in which the user must also maneuver the robot around an obstacle. The robot assistance helps in obstacle avoidance, but it must trade off between optimal performance (passing as close to the obstacle as possible) and user confidence (giving the obstacle a wide berth to comfort the user) [7]. When the user glances more than usual towards the obstacle, the robot will give it a wider berth. We will evaluate this behavior in a user study to determine how this responsiveness enhances the robot's performance as well as users' confidence in its behavior.

5 CONCLUSIONS

In this work, we show how eye gaze can improve robot performance during shared control. First, we demonstrate features of users' eye gaze behavior during a teleoperation task. Then, we develop a pipeline for processing the raw eye gaze signal to include context. Finally, we propose an assistive application for eye gaze: adapting the safety margin of a robot navigating around an obstacle by measuring people's comfort with its behavior from their eye gaze. This project shows how eye gaze helps assistive robotics.

One important future direction of this project is to evaluate the systems we have developed with people with upper mobility impairments. While the systems described here have been entirely evaluated with non-disabled participants, making gaze-based assistance useful for people who use a wheelchair-mounted assistive robot arm requires additional studies and verification.

Beyond assistance, however, this work illustrates the usefulness of nonverbal natural signals like eye gaze behavior for understanding complex human mental states. Eye gaze research has shown that gaze patterns can reveal many different aspects of a person's mental state, from their expertise on a task to their cognitive load to which areas they are focusing more on. Incorporating natural signals enables human-robot collaboration paradigms to move beyond goal-only models of humans to encompass a wide variety of information about a collaborator's mental state.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation (IIS-1755823) and the Paralyzed Veterans of America.

REFERENCES

[1] Reuben M. Aronson and Henny Admoni. 2018. Gaze for Error Detection During Human-Robot Shared Manipulation. In *Fundamentals of Joint Action workshop, Robotics: Science and Systems*.

[2] Reuben M. Aronson and Henny Admoni. 2019. Semantic gaze labeling for human-robot shared manipulation. In *Eye Tracking Research and Applications Symposium (ETRA)*. Association for Computing Machinery. https://doi.org/10.1145/3314111.3319840

[3] Reuben M. Aronson, Thiago Santini, Thomas. C. Kübler, Enkelejda Kasneci, Siddhartha Srinivasa, and Henny Admoni. 2018. Eye-Hand Behavior in Human-Robot Shared Manipulation. In *ACM/IEEE International Conference on Human-Robot Interaction*.

[4] Matthias Bernhard, Efstathios Stavrakis, Michael Hecher, and Michael Wimmer. 2014. Gaze-to-Object Mapping During Visual Search in 3D Virtual Environments. *ACM Trans. Appl. Percept.* 11, 3, Article 14 (Aug. 2014), 17 pages. https://doi.org/10.1145/2644812

[5] Yu Chen and D. H. Ballard. 2002. Learning to recognize human action sequences. In *Proceedings - 2nd International Conference on Development and Learning, ICDL 2002*.

[6] Çağla Çığ and Tevfik Metin Sezgin. 2015. Gaze-based prediction of pen-based virtual interaction tasks. *International Journal of Human Computer Studies* 73 (2015), 91–106.

[7] Anca D. Dragan, Kenton C.T. Lee, and Siddhartha S. Srinivasa. 2013. Legibility and Predictability of Robot Motion. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction (HRI '13)*. IEEE Press, Piscataway, NJ, USA, 301–308. http://dl.acm.org/citation.cfm?id=2447556.2447672

[8] Anca D Dragan and Siddhartha S Srinivasa. 2013. A policy-blending formalism for shared control. *The International Journal of Robotics Research* 32, 7 (2013), 790–805. https://doi.org/10.1177/0278364913490324 arXiv:https://doi.org/10.1177/0278364913490324

[9] Elena Corina Grigore, Kerstin Eder, Anthony G. Pipe, Chris Melhuish, and Ute Leonards. 2013. Joint action understanding improves robot-to-human object handover. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*.

[10] Chien-Ming Huang and Bilge Mutlu. 2016. Anticipatory robot control for efficient human-robot collaboration. In *ACM/IEEE International Conference on Human-Robot Interaction*. 83–90.

[11] Siddarth Jain and Brenna Argall. 2018. Recursive Bayesian Human Intent Recognition in Shared-Control Robotics. *IROS* (2018).

[12] Shervin Javdani, Henny Admoni, Stefania Pellegrinelli, Siddhartha S. Srinivasa, and J. Andrew Bagnell. 2018. Shared autonomy via hindsight optimization for teleoperation and teaming. *The International Journal of Robotics Research* 37, 7 (2018), 717–742. https://doi.org/10.1177/0278364918776060 arXiv:https://doi.org/10.1177/0278364918776060

[13] Roland S Johansson, Gö Ran Westling, Anders Bä Ckströ, and J Randall Flanagan. 2001. Eye–Hand Coordination in Object Manipulation. *The Journal of Neuroscience* 21, 17 (2001), 6917–6932.

[14] Kinova Robotics, Inc. 2020. Robot arms. Retrieved Jan 13, 2018 from http://www.kinovarobotics.com/assistive-robotics/products/robot-arms/

[15] Thomas C. Kübler, Colleen Rothe, Ulrich Schiefer, Wolfgang Rosenstiel, and Enkelejda Kasneci. 2017. SubsMatch 2.0: Scanpath comparison and classification based on subsequence frequencies. *Behavior Research Methods* 49, 3 (2017), 1048–1064.

[16] Michael F. Land and Mary Hayhoe. 2001. In what ways do eye movements contribute to everyday activities? *Vision Research* 41, 25 (2001), 3559–3565.

[17] Jonathan Samir Matthis, Jacob L. Yates, and Mary M. Hayhoe. 2008. Gaze and the Control of Foot Placement When Walking in Natural Terrain. *Current Biology* (2008).

[18] Katharina Muelling, Arun Venkatraman, Jean-Sebastien Valois, John E. Downey, Jeffrey Weiss, Shervin Javdani, Martial Hebert, Andrew B. Schwartz, Jennifer L. Collinger, and J. Andrew Bagnell. 2017. Autonomy infused teleoperation with application to brain computer interface controlled manipulation. *Autonomous Robots* 41, 6 (01 Aug 2017), 1401–1422. https://doi.org/10.1007/s10514-017-9622-4

[19] Benjamin A. Newman, Reuben M. Aronson, Siddhartha S. Srinivasa, Kris Kitani, and Henny Admoni. 2018. HARMONIC: A Multimodal Dataset of Assistive Human-Robot Collaboration. *ArXiv e-prints* (July 2018). arXiv:cs.RO/1807.11154

[20] Thies Pfeiffer and Patrick Renner. 2014. EyeSee3D: A Low-cost Approach for Analyzing Mobile 3D Eye Tracking Data Using Computer Vision and Augmented Reality Technology. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '14)*. ACM, New York, NY, USA, 369–376. https://doi.org/10.1145/2578153.2628814

[21] Thies Pfeiffer, Patrick Renner, and Nadine Pfeiffer-Leßmann. 2016. EyeSee3D 2.0: Model-based Real-time Analysis of Mobile Eye-tracking in Static and Dynamic Three-dimensional Scenes. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications (ETRA '16)*. ACM, New York, NY, USA, 189–196. https://doi.org/10.1145/2857491.2857532

[22] Siddharth Reddy, Anca D. Dragan, and Sergey Levine. 2018. Shared Autonomy via Deep Reinforcement Learning. *Robotics: Science and Systems* (2018).