Understanding Human Trust in Assistive Teleopertion

Maunika Atmakuri, Annie Black, Angela Li, Vinitha Ranganeni, Daphne Tan
{matmakur, ablack, qingyanl, vrangane, hdtan}@andrew.cmu.edu

1 Introduction

Assistive healthcare robotic arms are promising in that they increase the freedom and abilities of people who lack dexterity. These high-dimensional robot arms are often controlled by a low-dimensional joystick input and requires switching modes. Each mode controls a subset of the degrees of freedom of the robot. However, changing modes is a harmful distraction that impedes efficient control (Herlant, Holladay, and Srinivasa 2016). *The goal of the project is to improve the performance of completing tasks, with robot assistance, while maintaining user trust.* Herlant et. al proposed a simple model to automatically switch modes that increases user satisfaction while maintaining performance (Herlant, Holladay, and Srinivasa 2016).

This project is an extension of Herlant et. al's work that incorporate two areas of Human-Robot Interaction: **algorithmic design** and **user studies**. Our work analyzes user trust, in assistive teleoperation that uses time-optimal mode switching, as the user's visibility of the environment varies.

The remainder of the report will be structured as follows: we provide a brief overview of current work in assistive teleoperation and maintaining user trust (Sec. 2). In Sec. 3 we present our extension of Herlant et. al's time-optiaml mode switching algorithm. Then we discuss our user study design (Sec. 5) and the design of the interface used in the study (Sec. 4). Finally, we provide a description of our experiments and results (Sec. 6) and a discussion of the limitations and potential future directions (Sec. 7).

Figure 1: On the left is a figure of a Kinova JACO mounted to a wheelchair (Robotics 2016) that must be controlled by changing between a series of modes (Admoni 2018).

2 Related Work

Herlant et. al identified that mode switching consumes about 17.4% of execution time even for able bodied users (Herlant, Holladay, and Srinivasa 2016). Thus, they proposed an algorithm to automatically switch the modes in order to minimize the time taken to complete the task. This algorithm was tested on three different levels (Herlant, Holladay, and Srinivasa 2016):

1. **Manual**: The user has full control over mode switching; the robot provides no assistance
2. **Automatic**: The robot automatically switches the mode whenever it enters a new region based determined by an optimality map. This change happens the first time the robot enters the zone. The user can change the mode as they please.
3. **Forced**: The robot automatically switches the mode to the time-optimal mode. However, after every action the user took, the robot would switch back to the time-optimal mode.

Herlant et. al examined the assistance types through task performance and user satisfaction. However, trust is also an essential component in scenarios where a robot and human must work together to reach a common goal (Atkinson, Clancey, and Clark 2014). Thus, we are interested in exploring how human trust varies as a robot provides different levels guidance through assistive teleoperation.

In order to test the effectiveness of the assistance the robot provides, we must be able to measure trust. Yagoda et. al grouped trust into three categories: *performance*, *function* and *semantics* (Yagoda and Gillan 2012). There are two basis of trust in the performance category that are relevant to our work. The first is **timely** which is task completion occurring at a favorable time. The second is **dependability** which is the degree to which behavior is consistent and expected.

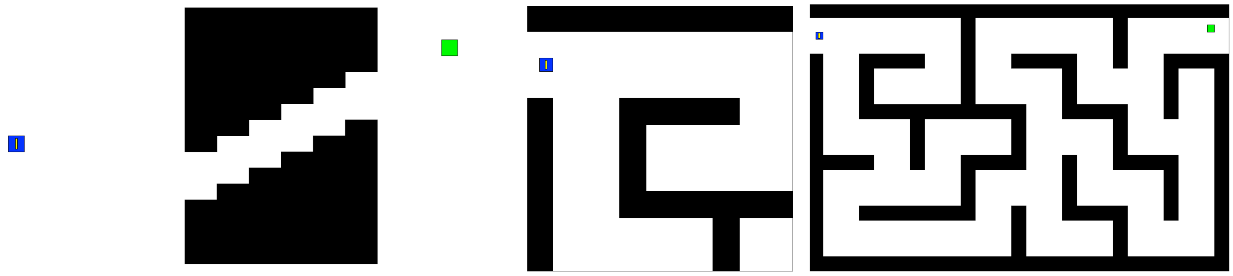


Figure 2: (a) Practice map for user study. (b) Partially visible map where the field of view changes as the robot moves. (c) Fully visible map where the position of the robot and goal are always visible.

3 Algorithmic Contribution

Herlant et. al determined the optimal mode by running Dijkstra's algorithm on a 2D workspace $\mathcal{W} \subset \mathbb{R}^2$ augmented by mode $m \in \mathcal{M}, \mathcal{M} = \{m_h, m_v\}$. m_h is the mode that allows the robot to move horizontally and m_v is the mode that allowed the robot to move vertically. Everytime the user moved the robot to a new position (x_{new}, y_{new}) they determined the optimal mode by running Dijkstra's algorithm to find the cost $c((x_{new}, y_{new}, m), (x_g, y_g))\ \forall m \in \mathcal{M}$. Note, that the mode at the goal position is irrelevant. The optimal mode $m* = \text{argmin}_m\, c((x_{new}, y_{new}, m), (x_g, y_g))$.

Continuously running Dijkstra's algorithm is inefficient as one state can be re-expanded multiple times. Our variation of the time-optimal mode switching algorithm, utilizes Dijkstra's algorithm but only runs it once as a pre-processing phase. $c((x_{new}, y_{new}, m), (x_g, y_g))$ gives the cost of the shortest path from (x_{new}, y_{new}, m) to (x_g, y_g). However, if we run Dijkstra's algorithm once backward from (x_g, y_g) till every state has been explored, we can get a cost from every (x, y, m), $\forall m \in \mathcal{M}$ to the goal. This method allows use to generate an *optimality map* that we can simple query at every timestep to get the optimal mode.

This approach does not have a significant impact on the efficiency of the simulation. However, if this algorithm was extended to a higher-dimensional robot, such as the Kinova JACO, there would be a significant improvement in the time taken to find the optimal mode.

4 User Interface Design

To emulates having a controller that can control only a subset of the degrees of freedom of the robot in each mode, we had users navigate a two-dimensional space with a one dimensional controller. For each time step, the robot had a specific mode, denoted by the orientation of the yellow line at the center of the robot (blue square) (Fig. 2). When the line is horizontal, the robot can only move in the horizontal direction. In this mode, the up and down arrow keys would map to left and right movement respectively. Likewise, when the line is vertical the robot can move vertically. In this mode, the up and down arrow keys would map to up and down movement respectively. The user could switch between the modes using the spacebar.

Our interface also implements all the three assistance types (see Sec. 2). With the automatic assistance type, we introduce the concept of *optimal zones* which are sections of the map where one mode is more optimal than the other. These optimal zones are determined by the optimality map (see Sec. 3). An intuitive example is when navigating a vertical hallway, being in vertical mode would lead to fewer number of mode switches compared to being in horizontal mode. This is because the user would eventually switch to vertical to continue moving down the hallway, making there be one more mode switch compared to beginning with vertical mode.

5 Study Design

In our user study, we asked each participant to teleoperate a robot (blue box) and navigate it to a goal location (green box)in a 2-dimensional map using the mode-switching interface we provided.

Our hypotheses are as follows:

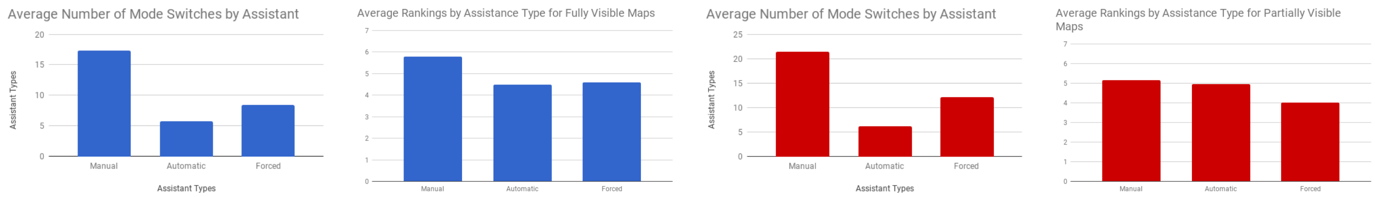
H.1 Users will only trust robots automatic assistance types on fully-visible environments.

H.2 Users will trust robots automatic and forced assistance types on partially-visible environments.

H.3 There will be an overall decrease in time taken to reach the goal when the robot provides forced or automatic assistance versus manual.

We designed our user study to manipulate two independent variables: visibility and assistance types (see Sec. 2). Visibility had two levels: partial and full. The partially visibile only allowed the user to see a portion of the map as they moved the robot (Fig. 2b). In contrast, the fully visible map allowed the user to see the entire map as the moved the robot (Fig. 2c).

We developed 3×2 factorial design to test our independent variables. There were 6 total conditions, with 21 users in each condition. We used a within-subjects design and gave users a practice map (Fig. 2a) before running our study to avoid the expertise effect.



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

Figure 3: (a) The average number of mode switches made the user in the fully-visible map for each assistance type. (b) The average rankings for how much the user trusted the robot for each assistance type in the fully-visible map. (c) The average number of mode switches made the user in the partially-visible map for each assistance type, (d) The average rankings for how much the user trusted the robot for each assistance type in the partially-visible map.

6 Experiments and Results

To assess trust, we primarily considered the number of mode switches made by the user, and secondarily considered the comfort rankings submitted by the user; we concluded that the number of mode switches would be an accurate measure for trust because this would show if the user had trusted the robot enough that there were significantly fewer mode switches, whereas user rankings was used as a secondary metric because comfort doesnt necessarily equate to trust, and also user rankings are subjective and very dependent on the order of the trials. We chose to perform one-tail t-tests to determine significance in our data, because our hypotheses look specifically for either a significant increase or a significant decrease in the metric, rather than any significant difference. We compare our p-values against $\alpha = 0.05$.

For **H.1**, we did not expect the user to be satisfied with the forced assistance type because if they could see the entire map, we believed they would be less likely to trust the robot's assistance. Performing a one-tail t-test to compare the number of mode switches, we saw a $p = 1.88 \times 10^{10}$, which is very significant. This shows that the users manually switched modes significantly less on the automatic assistance type trials in comparison with the manual assistance type trials (Fig. 3a).

In contrast, while looking at the survey data for fully-visible environment (Fig. 3b), the lesser rankings for forced/automatic versus manual is significant $p = 0.009$ and $p = 0.003$ respectively). This demonstrates that, while the users may have trusted the robot more, in terms of mode-switching, they were not comfortable with the assistance. As we are prioritizing the mode switching as our metric for trust, we do not reject the first hypothesis, but we note that the trust was prevelant but the comfort was not. Additionally, we accept that our hypothesis was wrongly limited, given that the number of mode switches for forced assistance types was also highly significant, with a $p = 2.29 * 10^5$. To conclude, for **H.1**, users trusted the robot's automatic and forced assistance but were not comfortable with it.

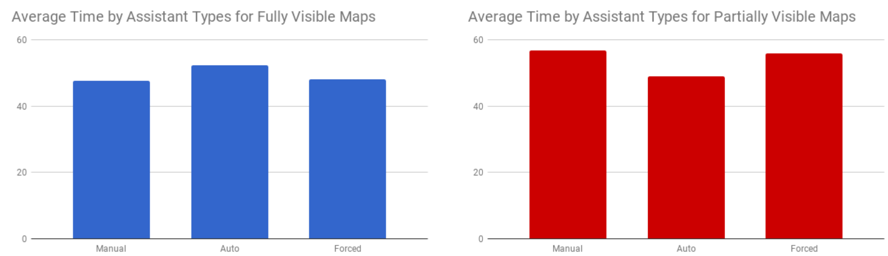
For **H.2** we expected that, because the user had a limited understanding of the environment in comparison with the robot, the user would be more inclined to trust the robot's assistance. There obvious decrease in the number of mode switched for both forced and automatic versus manual (Fig. 3c). Because the $p = 0.01$ for forced versus manual and $p = 4.58 * 10^{-10}$ for automatic versus forced, we conclude that both results are significant. Additionally, users trust rankings for each assistance type are very similar; $p = 0.03$ for forced and $p = 0.35$ for automatic (Fig. 3d). Therefore, users do not mind the assistance given by the automatic in partially-visible environments but are still not comfortable with forced. To conclude, for **H.2**, we accept it for both the forced and automatic assistance types.

Lastly, for **H.3** we hypothesized that there would be a decrease in time taken to reach the goal when the users are given forced/automatic assistance rather than manual control. [1] For each map type, the assistance types demonstrate very similar times (Fig. 4a, Fig. 4b). All the p-values are greater than $\alpha = 0.05$ (Fig. 4c), and so we have no significant differences in time. Thus, we reject **H.3**. Intuitively, we believe that while mode switching may have been significantly decreased, users still required time to reassess their situation after the robot switched modes, and so the times balanced out regardless of the robots assistance.

7 Discussion

Limitations While our work has shown promising results, there are many limitations to this work. First, users may have potentially learned the map layout even though we changed the goal position. Second, the optimal modes always guided the robot towards the walls users often found this uncomfortable. They did not want to crawl along the wall and often changed the mode to move the robot to the middle of the maze. Third, users may have responded differently depending on the randomized order of assistance types. Though all users got the partially visible map first, the assistance type was randomized. Users who got forced assistance early on may have reacted differently to the rest of the maps compared to users who received the manual or automatic assistance type first. Finally, our sample population was not diverse or large enough. Most of our sample size were young females majoring in a STEM field. For more insight, wed ideally test with a variety of people across different age

[1] Note that this hypothesis is regardless of map type; however, we expect that the full and partial maps would have differing times by design and so we still compare this data separately.



(a)

(b)

P-Values for Time by Assistance and Map Types

	Fully Visible	Partially Visible
Manual vs Auto	0.25	0.19
Manual vs Forced	0.47	0.45

(c)

Figure 4: (a) The average time taken to complete the task on the fully-visible map for each assistance type. (b) The average time taken to complete the task on the partially-visible map for each assistance type. (c) The p-values for the time taken to complete the task comparing assistance types and map visibility.

groups and backgrounds. The current data as it stands is insufficient for us to draw realistic conclusions on how this can be applied outside of this specific user study.

Future Work Even with limitations, there are many possible future directions for this work to explore new concepts and address these limitations. One interesting concept to explore is user trust when the robot provides suboptimal (maximizes mode switches) vs optimal guidance (minimizes mode switches). Our user study only examined user trust when the robot provided optimal assistance.

Second, wed also like to extend the algorithm to address the user discomfort with the mode switches that caused the robot to crawl along the walls. Ideally, the algorithm would not only optimize for the minimum number of mode switches but would take the user perceptions into account. This extension would allow us to truly understand user trust since the current algorithms lack of optimality from the users perspective may have potentially resulted in a lower comfort rating as reflected in our survey data.

Third, the ordering of different assistance types can be experimented with. We could potentially address the ordering of assistance types by perhaps fixing the map and assistance type for the first round (after the training map) so that we can establish consistency and better measure user trust. Since we believe this ordering could have impacted user comfort, having a fixed order for part of the study may help us counteract the potential inconsistency and mixed reports on comfort levels.

Lastly, this work could be extended to run user studies with the Kinova JACO arm. While the 2D interface does mimic the user interface of controlling the JACO arm, to a certain degree, we acknowledge potentially significant impact of leveraging a real robot for user studies. The physical operation of a robot could have impact on emotional comfort or discomfort and attitudes toward robot assistance.

References

[Admoni 2018] Admoni, H. 2018. Kinova jaco modes.

[Atkinson, Clancey, and Clark 2014] Atkinson, D. J.; Clancey, W. J.; and Clark, M. H. 2014. Shared awareness, autonomy and trust in human-robot teamwork. In *Artificial Intelligence and Human-Computer Interaction: Papers from the 2014 AAAI Spring Symposium*.

[Herlant, Holladay, and Srinivasa 2016] Herlant, L. V.; Holladay, R. M.; and Srinivasa, S. S. 2016. Assistive teleoperation of robot arms via automatic time-optimal mode switching. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press.

[Robotics 2016] Robotics, K. 2016. Jaco assistive robotic device: Empowering people with disabilities through innovative algorithms.

[Yagoda and Gillan 2012] Yagoda, R. E., and Gillan, D. J. 2012. You want me to trust a robot? the development of a human–robot interaction trust scale. *International Journal of Social Robotics*.