Mode Switch Assistance to Maximize Human Intent Disambiguation

Deepak Gopinath and Brenna D. Argall

I. INTRODUCTION

Assistive and rehabilitation devices such as powered wheelchairs, robotic arms and myoelectric prostheses play an important role in the lives of people with motor impairments. These devices help to increase their ability to perform activities of daily lives and reduce their dependence on caretakers.

The common paradigm for control of such high-dimensional devices has the human directly control the motion via a control interface. These interfaces (for example, a switch-based head array and Sip-N-Puff) are lower in dimensionality and bandwidth and due to this dimensionality mismatch, these interfaces operate in *modes* which correspond to different partitions of the control space. In order to have full control of the robot the user will have to switch between modes and this is known as *mode switching* [3]. It has been established that mode switching adds to the cognitive burden and results in degradation of task performance [2]. Furthermore, the more severe a person's impairment, the more limited are the control interfaces available for use.

The introduction of *shared autonomy* [4] to these systems helps to address these issues by letting the system take responsibility to some extent. For any assistive autonomy, the system typically needs an idea of what it is the human is trying to do—either by explicit indication from the user of the task or goal, or by inferring the human's intent from their control signals and/or sensor data. In our work, we develop an assistance paradigm that helps with intent inference, by selecting the control mode in which robot motion will *maximally disambiguate* human intent. In many shared-control systems, the assistance provided by the robot is regulated by a confidence metric that is the system's confidence in its own estimate of human intent.

In the example illustrated in Figure 1, a human control command issued along 'x' carries more information than other dimensions and is more intent-expressive. With the disambiguation assistance scheme developed in this work, we hope to elicit more legible human control commands by placing the user control in those modes with *maximum disambiguation* between the various goals in the scene. A *legible* motion in this context is one which will help the observer decipher the intent more *quickly* and *confidently*. The legibility and predictability of robot motion *to the human* have already been thoroughly investigated [1]. We apply similar concepts of legibility to the *human control* commands and propose a paradigm of *inverse legibility* in which the roles are switched and the human-generated actions *help the robot*.

II. ALGORITHM DESIGN

Let $\mathcal{G}$ be the set of all candidate goals in the scene with $n_g = |\mathcal{G}|$. The set of goals results in an associated set of confidences denoted as $\mathcal{C}$. Let $\mathcal{K}$ be the control space in which the robot operates and let $\mathcal{M}$ refer to the set of all modes that $\mathcal{K}$ is partitioned into.

We define a *disambiguation* metric, $D_k \in \mathbb{R}$ for each control dimension $k \in \mathcal{K}$, which is a function of c and $\frac{\partial c}{\partial k}$, where $c \in \mathcal{C}$. Lastly, we define a disambiguation metric $D_m \in \mathbb{R}$ for each control mode $m \in \mathcal{M}$. D_m is a measure of how legible the user commands would be if the user were to control the robot in mode m. The confidence function used in this work tries to capture the "directedness" of the human control commands and is given by

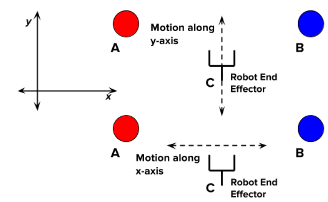

Fig. 1. Illustration of goal disambiguation along various control dimensions. A and B indicate two point goal locations.

$$c_g(\boldsymbol{x}, \boldsymbol{u}_h, \boldsymbol{u}_{r,g}) = \boldsymbol{u}_h^{trans} \cdot (\boldsymbol{x}_g - \boldsymbol{x})^{trans} + \boldsymbol{u}_h^{rot} \cdot \boldsymbol{u}_{r,g}^{(rot)}$$

where $\boldsymbol{u}_h$ is the human control command, $\boldsymbol{x}$ is the current position of the robot, $\boldsymbol{x}_g$ is the location of goal g, $trans$ refers to the translational and rot refers to the rotational parts of the entire control space.

Let the component of $\boldsymbol{x}$ along control dimension k be denoted as x. We identify four important considerations to inform the design of D_k.

1) Separation in confidences: A good measure for evaluating the confidence disambiguation potential of a control dimension is to compute the *separation*, Λ_k, in goal confidences.

$$\Lambda_k = \sum_{p=1}^{n_g} \sum_{q=p}^{n_g} |c_{\delta_x}^p - c_{\delta_x}^q|$$

where δ_x indicates $x + \Delta x$ or $x - \Delta x$ depending on the direction of perturbation and $|\cdot|$ denotes the absolute value.

2) Max of confidences: The maximum of the goal confidences is a good measure of the system's overall certainty in accurately estimating human intent. The max (Γ_k) is computed as

$$\Gamma_k = \max_{1 \le i \le n_g} c_{\delta_x}^i$$

3) Difference between largest confidences: Accurate disambiguation also benefits from a large separation between the

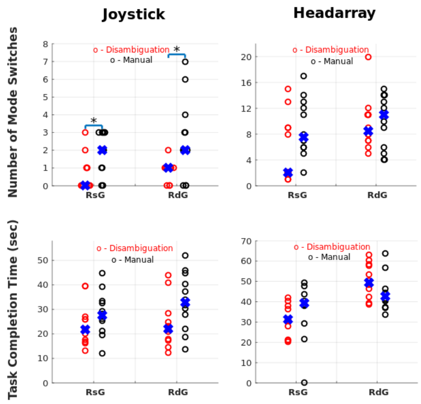

Fig. 2. Comparison of Disambiguation and Manual paradigms.

first and second most confident goals. This difference (Δ_k) is computed as

$$\Delta_k = \max(\mathcal{C}) - \max(\mathcal{C} \setminus \max(\mathcal{C}))$$

4) Gradients: The propensity for change and information gain upon the continuation of motion along control dimension k is encoded in the spread of gradients, Υ_k which is computed as $\frac{\partial c}{\partial k}\ \forall\ c \in \mathcal{C}$.

$$\Upsilon_k = \sum_{p=1}^{n_g} \sum_{q=p}^{n_g} \left| \frac{\partial c^p}{\partial k} - \frac{\partial c^q}{\partial k} \right|$$

Putting it all together: Λ_k, Γ_k, Δ_k and Υ_k are then combined to compute D_k as

$$D_k = \boldsymbol{w} \cdot (\Lambda_k \cdot \Gamma_k \cdot \Delta_k) + (1 - \boldsymbol{w}) \cdot \Upsilon_k$$

where $\boldsymbol{w}$ is a user-defined weight. The disambiguation metric D_m for control mode m is calculated as $D_m = \sum_j D_k \forall\ k \in m$. The control dimension with highest disambiguation capability $\boldsymbol{k}^*$ is given by $\mathrm{argmax}_k\, D_k$ and the mode with highest disambiguation capability $\boldsymbol{m}^*$ is given by $\mathrm{argmax}_m\, D_m$. Disambiguation mode $\boldsymbol{m}^*$ is the mode that the algorithm chooses *for* the human to better their intent. Any subsequent control command issued by the user in $\boldsymbol{m}^*$ is likely to be more legible due to maximal goal confidence disambiguation.

III. EXPERIMENTS AND RESULTS

The experiments were performed using the MICO robotic arm (Kinova Robotics, Canada). Three kinds of mode switching paradigms were evaluated: *Manual* – User manually performs all mode switches, *Disambiguation* – The disambiguation system is activated right at the beginning of a trial and the trial starts in $\boldsymbol{m}^*$, *On Demand* – User can request a mode switch assistance at any time during task execution.

Three different reaching tasks, namely 1) Simple Reaching (*R*) 2) Reaching with Same Grasp (*RsG*) and 3) Reaching with Different Grasp (*RdG*) with different goal configurations were developed for our pilot study. Analysis was performed only on data collected from *RsG* and *RdG*. The human control command $\boldsymbol{u}_h$ was captured using two different control interfaces: 2-axis joystick and a head array.

We investigated the use and utility of our disambiguation approach in a pilot study. Four subjects participated in the pilot study, (3 male, 1 female), and all were lab members. Task completion times and the number of mode switches were evaluated during the pilot study.

An improvement in task performance in terms of a decrease in the number of mode switches was observed across all task–interface combinations. Figure 2 (top row) reveals the general trend of a decrease in the number of mode switches. The difference in the number of mode switches was statistically significant only when using the joystick. In Figure 2 (bottom row), a decrease in task completion times during *Disambiguation* paradigm was observed in all but one case: *RdG*–headarray combination. However, these differences were not statistically significant for any of the task-interface combinations. Although the subjects demonstrated a wide range of disambiguation request behaviors, we were able to observe a general trend of an increase in disambiguation requests with an increase in task difficulty.

IV. DISCUSSION AND CONCLUSION

In future versions of the algorithm, it might be useful to bias the computation of the "best control mode" by looking for cues in the past history of the robot trajectory and control commands such that the control mode chosen will always be useful for the human in reaching the goal s/he has in mind. This will also likely improve the robustness and result in higher user acceptance. A more extensive user study with motor-impaired subjects will be conducted in the future to evaluate the utility value of the disambiguation assistance system and further explore and understand the disambiguation request patterns of users. The goal of the algorithm developed in this paper is to seek legible control commands from the human by placing the control in those modes able to *maximally disambiguate* between the various goals in the scene. Moreover, preliminary pilot study results indicated that the disambiguation paradigm proposed was a promising approach in decreasing task effort (number of mode switches) across interfaces and tasks.

REFERENCES

[1] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 301–308. IEEE, 2013.

[2] Stephen Monsell. Task switching. *Trends in cognitive sciences*, 7(3):134–140, 2003.

[3] Katherine Tsui, Holly Yanco, David Kontak, and Linda Beliveau. Development and evaluation of a flexible interface for a wheelchair mounted robotic arm. In *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pages 105–112. ACM, 2008.

[4] Ivan Volosyak, Oleg Ivlev, and Axel Graser. Rehabilitation robot friend ii-the general concept and current implementation. In *9th International Conference on Rehabilitation Robotics (ICORR), 2005*, pages 540–544. IEEE, 2005.