Mode Switch Assistance to Maximize Human Intent Disambiguation

Deepak Gopinath, Department of Mechanical Engineering, Northwestern University

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*. It has been established that mode switching adds to the cognitive burden and results in a 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* [3] to these systems helps to address these issues by letting the system take over control 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 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 $n_g = |\mathcal{G}|$. Every goal $g \in \mathcal{G}$ has an associated confidence c_g which is a measure of the robot's estimate that g is the user's intended goal. The set of all confidences is 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

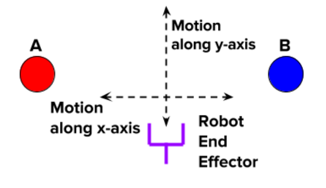

Fig. 1. Illustration of goal disambiguation along various control dimensions. A and B indicate two point goal locations.

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 $\frac{\partial c}{\partial x_k}$, where $c \in \mathcal{C}$ and x_k is the component of robot's position along control dimension k[1]. Lastly, we define a disambiguation metric $\boldsymbol{D}_m \in \mathbb{R}$ for each control mode $m \in \mathcal{M}$. $\boldsymbol{D}_m$ is a measure of how legible the user commands would be if the user were to control the robot in mode m.

We identify four important considerations to inform the design of D_k.

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

$$\Lambda_k = \sum_{p=1}^{n_g}\sum_{q=p}^{n_g}|c^p_{\delta_x} - c^q_{\delta_x}|$$

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 confidences (Γ_k) is a good measure of the system's overall certainty in accurately estimating human intent. This is computed as

$$\Gamma_k = \max_{1 \leq i \leq n_g} c^i_{\delta_x}$$

3) Difference between largest confidences: Accurate disambiguation also benefits from a large separation between the 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 $\frac{\partial c}{\partial x}$ $\forall\, c \in \mathcal{C}$. The spread of the gradients (Υ_k) is computed as

$$\Upsilon_k = \sum_{p=1}^{n_g}\sum_{q=p}^{n_g}\Big|\frac{\partial c^p}{\partial x} - \frac{\partial c^q}{\partial x}\Big|$$

$$\boldsymbol{u} = \alpha \cdot \boldsymbol{u}_r + (1-\alpha)\cdot \boldsymbol{u}_h$$

[1]The subscript k will be dropped from x_k for brevity in notation

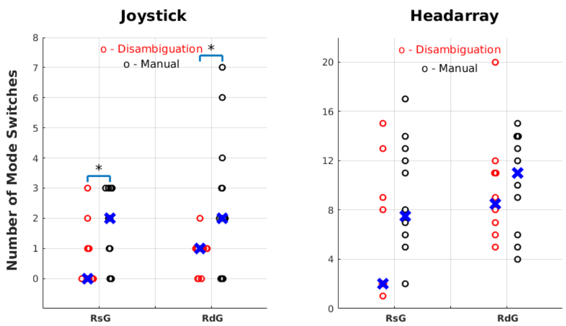

Fig. 2. Comparison of *Disambiguation* and *Manual* paradigms.

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 task-specific weight set by the system-designer. The disambiguation metric $\boldsymbol{D}_m$ for control mode m is calculated as $\boldsymbol{D}_m = \sum_j D_j \ \forall \ j \in m$. The control dimension with highest disambiguation capability $\boldsymbol{k}^*$ is given by $\text{argmax}_k\, D_k$ and the mode with highest disambiguation capability $\boldsymbol{m}^*$ is given by $\text{argmax}_m\, \boldsymbol{D}_m$. 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 pilot experiment was 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 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. 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 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 order to quantify the impact of different confidence functions on the computation of $\boldsymbol{k}^*$ and $\boldsymbol{m}^*$, simulations were performed in which $\boldsymbol{k}^*$ was computed at 2000 uniformly sampled points in the workspace of the robot. Two different confidence functions were used: Proximity-based (**C1**) and Directedness-based (**C2**). The goal configuration was same as that in *RsG* in which the goal positions were spread out maximally along x and z axes. Intuitively the robot will be able to infer the human's intent if the human control command is along the x or z axes. Table I reports the number of times the algorithm picked each of the three control dimensions, for each confidence function. **C1** often was unable to capture the human intent properly. Furthermore, **C1** had "null" spaces where all confidences were identically equal to zero and therefore disambiguation was not possible. By contrast, with **C2** the algorithm identified x as the preferred dimension 1711 out of 2000 samples, and z in 196 of the remaining 289 samples, which indicates that the confidence function along with our algorithm was able to select the disambiguating dimensions over 95% of the time.

Best control dimension distribution				
Confidence Function	**X**	**Y**	**Z**	**NULL**
C1	579	615	446	360
C2	1711	93	196	0

TABLE I
BEST CONTROL DIMENSION DISTRIBUTION FOR TWO DIFFERENT CONFIDENCE FUNCTIONS.

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] 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.