Schedule-based Motion Prediction for Human-Centric Autonomous Observation

David Kent* and Sonia Chernova*

***Abstract*— As robot autonomy improves, robots are increasingly being considered in the role of autonomous observation systems — free-flying cameras capable of actively tracking human activity within some predefined area of interest. In this work, we formulate the autonomous observation problem through multi-objective optimization, presenting a novel Semi-MDP formulation of the autonomous observation problem that maximizes observation rewards while accounting for both human- and robot-centric costs. We demonstrate how the problem can be solved for a known human trajectory using Constrained MDPs, and extend the approach to incorporate human motion prediction based on noisy rationality models, defined over a set of goals extracted from a task schedule.**

I. Introduction

Human operations in extreme and remote environments, such as space and deep water domains, have the potential to benefit from robots with autonomous observation capabilities. Due to their high-cost and high-risk nature, human activities in such domains are often video recorded for documentation and later analysis. NASA, for example, collects video documentation of each experiment conducted on the International Space Station (ISS). As robot autonomy improves, robots are increasingly being considered in the role of *autonomous observation systems* — free-flying cameras capable of actively tracking human activity within a predefined area of interest. Example systems include the free-flying NASA Astrobee [1] and European Space Agency CIMON [2], as well as autonomous camera robots being considered for underwater exploration [3].

While existing robot hardware offers capable candidates for autonomous observation systems, the autonomous observation problem itself is complex and largely unsolved. Autonomous observation of humans moving in 3D space is challenging due to the proliferation of viewpoints required to cover unconstrained humans in 6-DOF environments. Further, the robot should act as a passive observer, causing minimal distraction to the human subject from both collisions and visual and auditory disturbance.

In this work, we formulate the autonomous observation problem through multi-objective optimization [4]. We present a novel Semi-MDP formulation of the autonomous human observation problem that maximizes observation rewards while accounting for both human- and robot-centric costs, solvable with Constrained Markov Decision Processes (CMDPs) [5]. We validate our approach on activity tracking using a simulated model of the Astrobee robot operating within a simulated ISS environment developed by NASA

*Georgia Institute of Technology, Atlanta, GA, USA {dekent, chernova}@gatech.edu

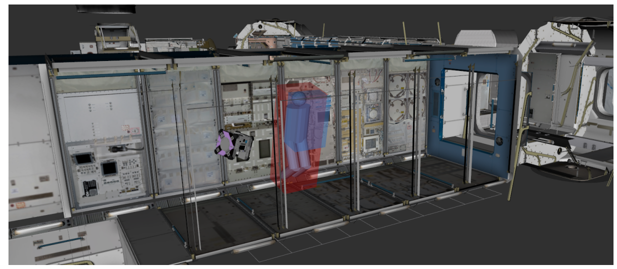

Fig. 1: The Astrobee platform in a module of the ISS, simulated in Gazebo and visualized in rviz.

(Figure 1). Additionally, we propose an extension to our approach by solving the problem formulation over a set of trajectories sampled from noisy rationality models designed to predict goal-based human motion [6], where our goal-based models are informed by a given task schedule.

II. Problem Formulation

We define the autonomous observation problem as a Semi-Markov Decision Process (SMDP) [7] with the components $(S, A(s), p(s'|s,a), p(\tau|s,a,s'), r(s,a,s',\tau))$, defined as follows:

- $s(t) \in S$ is a state consisting of the robot state and the human subject's pose $[x_r, x_h]$. We assume the robot is at one of a set of waypoints $x_r.pose \in [w_0, w_1, \ldots, w_n]$, and that the human's trajectory $x_h(t)$ is known (we relax this assumption in Section IV).
- $A(s)$ is the set of actions available to the robot at the current state, which must include the subset $\{hold\_pos()\} \cup \{move(w_i) | w_i \in [w_0, w_1, \ldots, w_n]\}$.
- $p(s'|s,a)$, the state transition function, is the probability that executing action a in state s will result in state s'.
- $p(\tau|s,a,s')$, the time duration distribution function, is the probability that transitioning from state s to state s' with action a will take the duration τ.
- $r(s,a,s',\tau)$ is the reward function. We model this as an observation reward rate received over the period τ, i.e. $r(s,a,s',\tau) = r(s,a,s')\tau$.

We define the observation reward based on subject coverage and resolution (a function of distance). We calculate the reward as the percentage of a region-of-interest (ROI) covered by the robot's field of view V_r, scaled by the distance from the robot to the ROI center.

$$r(s,a,s') = \frac{1}{||ROI_c - x_r.pose||} \frac{||V_r \bigcap ROI||}{||ROI||} \tag{1}$$

The ROI can be defined as a human-centric task workspace (e.g. the blue region in Figure 1), the subject's full bounding box, or whatever area the robot's camera should capture.

Additionally, we introduce a set of cost functions $c_i(s, a, s', \tau)$ to model human- and robot-centric costs. As with the reward function, we accumulate costs over a time duration. The costs are as follows:

- $c_0(s, a, s')$ represents collision between the robot and the human, calculated based on the distance from the robot to the human's workspace[1], shown in red in Figure 1. The platform-specific parameter α_0 controls how close to the workspace edge the robot can be.

$$c_0(s, a, s') = e^{-\alpha_0 dst(x_r.pose, wkspc)} \tag{2}$$

- $c_1(s, a, s')$ represents the intrusiveness, calculated based on distance from the robot to the human's head[1]. Note that this is in direct conflict with the observation reward. The platform-specific parameter α_1 controls the rate at which distance decreases the robot's intrusiveness.

$$c_1(s, a, s') = e^{-\alpha_1 ||x_r.pose - x_{h\_head}||} \tag{3}$$

- $c_2(s, a, s')$ represents the platform-specific power consumption of each of the robot's actions.

III. Planning Method

As we are solving for total accumulated reward over a fixed trajectory, we solve the SMDP over a finite horizon with undiscounted rewards. We begin by defining reward and cost functions calculated for only a state and action by taking expectations over resulting states and action durations:

$$r(s, a) = \sum_{s'} \left[p(s'|s, a) \sum_{\tau} p(\tau|s, a, s') r(s, a, s') \tau \right] \tag{4}$$

$$c_i(s, a) = \sum_{s'} \left[p(s'|s, a) \sum_{\tau} p(\tau|s, a, s') c_i(s, a, s') \tau \right]. \tag{5}$$

Next, we frame the problem as a CMDP, represented by the tuple $(S, s_0, A(s), p(s'|s, a), r(s, a), \mathbf{c}(s, a), \mathbf{d})$, where $c_i(s, a) \in \mathbf{c}$ is a cost function (i.e. Equation 5), and $d_i \in \mathbf{d}$ is a constraint value associated with c_i. The goal of a CMDP is to maximize the expected total reward subject to a set of constraints defined by the expected total costs:

$$\begin{aligned} &\max_{\pi} u_r^{\pi}(s_0) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{N} r(s_t, a_t) | s_0 \right] \\ &s.t.\, u_c^{\pi}(s_0) = \mathbb{E}_{\pi} \left[\sum_{t=0}^{N} c_k(s_t, a_t) | s_0 \right] \leq d_k \quad \forall k. \end{aligned} \tag{6}$$

We solve the CMDP by reformulating the above problem as a linear program (see [5] for details).

IV. Human Motion Prediction

Rather than solve the autonomous observation CMDP for a single human trajectory, we can leverage human motion prediction to consider a set of likely human trajectories. Specifically, we base our approach on Fisac et al.'s confidence-aware human motion prediction [6]. We define the human state x_H as a 6-DOF pose $x_H = [\mathbf{p}_H, \dot{\mathbf{p}}_H]$, and the human's actions as a 6-DOF acceleration $u_H = [\ddot{\mathbf{p}}_H]$, with transition dynamics defined as $x_H^{t+1} = f(x_H^t, u_H^t)$. For a given task that the robot is observing, we define the human's current goal $g \in \{x_{g0}, x_{g1}, \dots\}$ as a pose from the set of goal poses defined by a task schedule (e.g. for an equipment transfer task, goals would include poses for the equipment pick-up and drop-off locations). We can then determine a human motion distribution as follows:

$$\begin{aligned} &P(x_H^{t+1} | x_H^t, \beta, g) = \\ &\qquad \sum_{u_H^t} \mathbb{1}\{x_H^{t+1} = f_H(x_H^t, u_H^t)\} P(u_H^t | x_H^t, \beta, g), \end{aligned} \tag{7}$$

where $P(u_H^t | x_H^t, \beta, g)$ is a noisy rationality model with model confidence β. Further, the robot can update its joint 111belief in the model confidence and the human's current goal $b^t(\beta, g)$ with online observations, as explained in [6].

We incorporate motion prediction into the autonomous observation problem by sampling a set of N human trajectories $\mathbf{x}_h$ from Equation 7, over which we calculate expected rewards and costs $\tilde{r}(s_t, a_t) = \frac{1}{N} \sum_{x_h} r(s_t(x_h), a_t)$ and $\tilde{c}_k(s_t, a_t) = \frac{1}{N} \sum_{x_h} c_k(s_t(x_h), a_t)$. We then substitute $\tilde{r}(s_t, a_t)$ and $\tilde{c}_k(s_t, a_t)$ for $r(s_t, a_t)$ and $c_k(s_t, a_t)$ in Equation 6, and solve for an optimal policy over the set of trajectories.

V. Results and Future Work

To date, we have implemented the autonomous observation problem in a case study using NASA's Astrobee robot to observe humans in a simulation of the ISS. We include a video[2] showing visualizations of the robot's policy under different constraint specifications for an inspection task with a known human trajectory. Our current work involves implementing the human motion prediction component, with the goal of testing the system in a more realistic use case observing real humans.

Acknowledgement

This work is supported in part by an Early Career Faculty grant from NASA's Space Technology Research Grants Program.

References

[1] T. Smith, J. Barlow, M. Bualat, T. Fong, C. Provencher, H. Sanchez, and E. Smith, "Astrobee: A new platform for free-flying robotics on the international space station," 2016.

[2] "CIMON - the intelligent astronaut assistant," https://www.dlr.de/dlr/en/desktopdefault.aspx/tabid-10212/332_read-26307/#/gallery/29911, published: 2018-03-02.

[3] M. Ishida and K. Shimonomura, "Marker based camera pose estimation for underwater robots," in *2012 IEEE/SICE International Symposium on System Integration (SII)*. IEEE, 2012, pp. 629–634.

[4] D. M. Roijers, P. Vamplew, S. Whiteson, and R. Dazeley, "A survey of multi-objective sequential decision-making," *Journal of Artificial Intelligence Research*, vol. 48, pp. 67–113, 2013.

[5] E. Altman, *Constrained Markov decision processes*. CRC Press, 1999.

[6] J. Fisac, A. Bajcsy, S. Herbert, D. Fridovich-Keil, S. Wang, C. Tomlin, and A. Dragan, "Probabilistically safe robot planning with confidence-based human predictions," in *Proceedings of Robotics: Science and Systems*, 2018.

[7] Q. Hu and W. Yue, *Markov decision processes with their applications*. Springer Science & Business Media, 2007, vol. 14.

[1] Taken together, the costs c_0 and c_1 account for human proxemics.

[2] `https://youtu.be/PC-qKuJ_g10`