Developing an Adaptive Robotic Assistant for Close Proximity Human-Robot Collaboration in Space

Developing an Adaptive Robotic Assistant for Close Proximity Human-Robot Collaboration in Space

Przemyslaw Lasota[*], Stefanos Nikolaidis[†] and Julie Shah [‡]

Massachusetts Institue of Technology, Cambridge, MA, 02139, U.S.A

In this paper, we present a framework for an adaptive and risk-aware robot motion planning and control, and discuss how such a framework could handle uncertainty in human workers' actions and robot localization. We build on our prior investigation, where we describe how uncertainty in human actions can be modeled using the *entropy rate* in a Markov Decision Process. We then describe how we can incorporate this model of uncertainty into simulations of a simple collaborative system, involving one human worker and one robotic assistant, to produce risk-aware robot motions. Next, we highlight the difficulties associated with localization uncertainty in a space environment and describe how we can incorporate this uncertainty into an adaptive system as well. Expected advantages of an adaptive system are described, including increases in overall efficiency due to reductions in idle time, increases in concurrent motion, faster task execution, as well as subjective improvements in the worker's satisfaction with the assistant and reduced worker stress and fatigue. A pilot experiment designed to evaluate the benefits of introducing risk-aware motion planning is described. It is found that human-robot teams in which the robot utilizes risk-aware motion planning show on average 24% more concurrent motion and execute the task 13% faster, while simultaneously improving safety by having a 19.9% larger mean separation distance between the human and robot workers. Finally, possible future system developments and user studies are discussed.

I. Introduction

The introduction of robotic assistants into previously human-only domains is a topic generating a significant amount of interest, both on Earth for manufacturing applications, as well as in space. Robonaut 2, for example, has become a permanent resident of the International Space Station in the beginning of 2011, and is envisioned to aid operations on the station's Destiny laboratory. While confined to the interior of the station for now, there are plans to continue development of Robonaut to make it carry out key support roles, working alongside astronauts both inside the station and during EVAs.[4]

Incorporating robots into workplaces, especially those in which humans and robots would work together in close proximity, poses safety challenges that may be addressed through human-centered design of robot autonomy and control. One such challenge is the inherent uncertainty in the actions of a human worker. While some tasks require a strict order of actions to be followed, in many human-only domains, the exact sequence of actions to be taken is left to the worker's discretion. While robots can manage strictly-structured tasks well by performing pre-programmed motions repetitively, they are not capable of adjusting to the uncertainty introduced by a human co-worker who is not following a strict sequence of actions. New applications for human-robot collaboration will require robots that act in a risk-aware manner that accommodates the uncertainty in the human's actions and motions.

As recent research in the field has shown, there are many different ways of dealing with such uncertainties. One example of recent work presents a system which predicts human workspace occupancy based on early stages of the human workers movement, and then modifies the robots task selection and trajectory generation based on this information.[9] In work focused on improving assisted teleoperation through policy blending

[*]Graduate Student, MIT Department of Aeronautics and Astronautics
[†]Graduate Student, MIT Department of Aeronautics and Astronautics
[‡]Assistant Professor, MIT Department of Aeronautics and Astronautics, AIAA Professional Member.

of user input and prediction, the humans goals and motion trajectories are instead predicted through a formulation based on Inverse Reinforcement Learning.[5,6]

Incorporating robotic assistants in a space environment introduces additional difficulties, including uncertainty in robot localization.[2,8] In order to make the introduction of robotic assistants into uncertain domains possible, a method of adapting the robots to various uncertainties needs to be developed.

II. Uncertainty in Human Actions

Modeling Human-Robot Interaction

The first step in developing an adaptive robotic assistant is to model the uncertainty in the human's next action. In our ongoing research we utilize a Markov Decision Process (MDP) to computationally encode a teaming model that captures knowledge about the role of the robot and the human team member.[10]

We describe the teaming model as a Markov Decision Process, a tuple $\{S, A, T, R\}$, where:

- S is a finite set of states of the world; it models the set of world environment configurations.
- A is a finite set of actions; this is the set of actions the robot can execute.
- $T : S \times A \longrightarrow \Pi(S)$ is the state-transition function, which, for each world state and action, gives a probability distribution over world states; the state transition function models the variability in human action. For a given robot action a, the human's next choice of action yields a stochastic transition from state s to a state s'. We write the probability of this transition as $T(s, a, s')$. In this formulation, human behavior is the cause of randomness in our model, although this can be extended to include stochasticity from the environment or the robot actions as well.
- $R : S \times A \longrightarrow \mathbf{R}$ is the reward function, giving the expected immediate reward gained by taking each action in each state. We write $R(s, a)$ for the expected reward for taking action a in state s.

The policy π of the robot is the assignment of an action $\pi(s)$ at every state s. The optimal policy π^* can be calculated using dynamic programming.[11] Under this formulation, the role of the robot is represented by the optimal policy π^*, whereas the robot's knowledge of the role of the human co-worker is represented by the transition probabilities T.

Quantitative Evaluation of Predictable, Convergent Joint Action

During team training, we expect the human and robot to perform similar patterns of actions as the human and robot converge on a collaboration strategy. This means that the same states will be visited frequently and the robot uncertainty about the human's action selection will decrease.

To evaluate the convergence of the robot's computational teaming model and the human mental model, we assume a uniform prior and compute the *entropy rate*[7] of the Markov chain (Eq. 1). The Markov chain is induced by specifying a policy π in the MDP framework. For the policy π we use the robot actions that match human preference, as it is elicited by the human after training with the robot. Additionally, we use the states $s \in S$ that match the preferred sequence of configurations to task completion. For a finite state Markov chain X with initial state s_0 and transition probability matrix T the entropy rate is always well defined.[7] It is equal to the sum of the entropies of the transition probabilities $T(s, \pi(s), s')$, for all $s \in S$, weighted by the probability of occurrence of each state according to the stationary distribution μ of the chain (Equation 1).

$$H(X) = -\sum_{s \in S} \mu(s) \sum_{s' \in S} T(s, \pi(s), s') log\left[T(s, \pi(s), s')\right] \tag{1}$$

Interestingly, the conditional entropy, given by Eq. 1, also represents the robot's uncertainty about the human's action selection. Post-hoc analysis of the human subject experiments[10] indicates that this measure decreases as human and robot train together, and increases when the human deviates from the robot's probabilistic model of human action-intent (Fig.1-Left). The entropy rate measure has also been shown to produce different results (of statistical significance) for various interactive planning techniques (Fig.1-Right), and to correlate to objective and subjective measures of team performance.[10] This measure can be generalized to encode situations where the human has multiple preferences or acts stochastically. We propose that these results provide first support that entropy rate may be used as a component of a dynamic assessment of risk, and may be used to generate risk-aware robot motions and action selections.

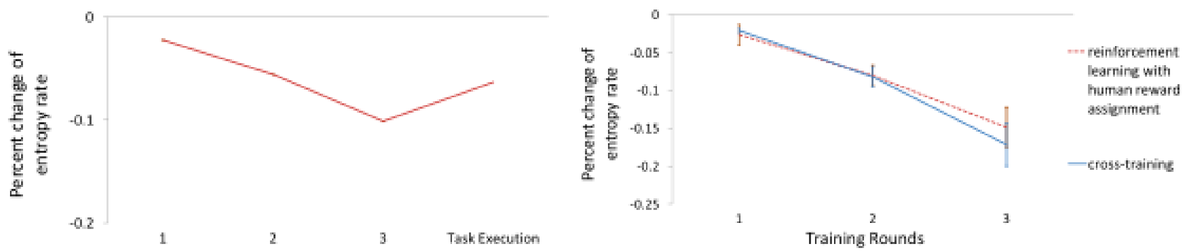

Figure 1. (Left) From prior human subject experiments[10] - in this trial the participant changed strategies for working with the robot from training to execution. The entropy rate decreases over all three rounds of interactive planning, and then sharply increases at execution as the person acts out a different strategy than planned. (Right) The entropy rate measure was also shown to produce different results, of statistical significance, for human-inspired interactive planning through cross-training versus training through traditional interactive reinforcement learning.[10] These differences correlated with improvements in objective and subjective measure of team performance for cross-training.

Adaptive Motion Planning

In this work, we take the next step to use these uncertainty models to adapt robot motions. Once we know the probability distributions of human actions, we can determine with what probability the human worker will occupy various locations in the workspace shared with the robot. Take, for example, a sample workspace shown in Figure 2.

If we know with high probability that the human worker will be placing a screw at the third location from the left next, we can utilize task and human motion models to determine what portion of the shared workspace will be occupied by the worker while he or she executes the task. In the case of the sample human-robot collaboration workspace shown in Figure 2, we can simplify this procedure by approximating the obstructed space by a cylinder extending from the person's shoulder down to the target location on the table, as shown by the virtual representation of the workspace depicted in Figure 3. Once the anticipated obstructed space is determined, the robot can choose actions which maneuver around this space.

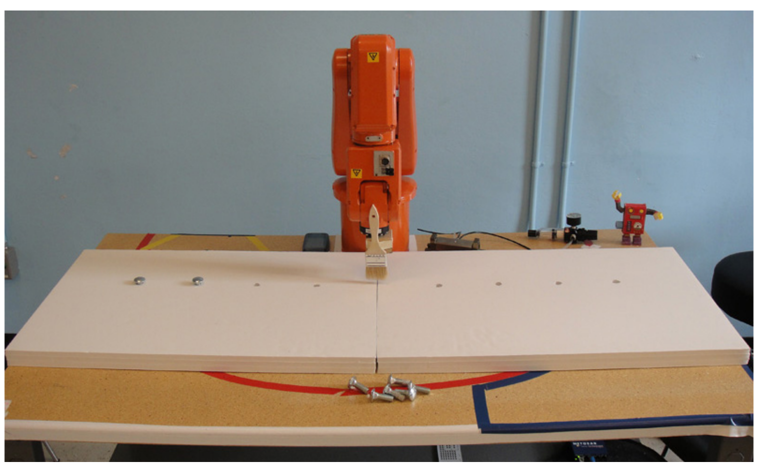

Figure 2. Sample human-robot collaboration workspace

III. Uncertainty in Localization

While the described approach is applied to adapt to uncertainty in the human's next action, the basic method is by no means limited to just this one type of uncertainty. If robotic assistants like Robonaut 2 are to be used in space, for example, several other uncertainties can arise. Due to environmental conditions and technological limitations (e.g. gyro bias, thermal bias, and attitude estimation errors), localization technology in a space environment comes with a significant amount of uncertainty. We plan to make use

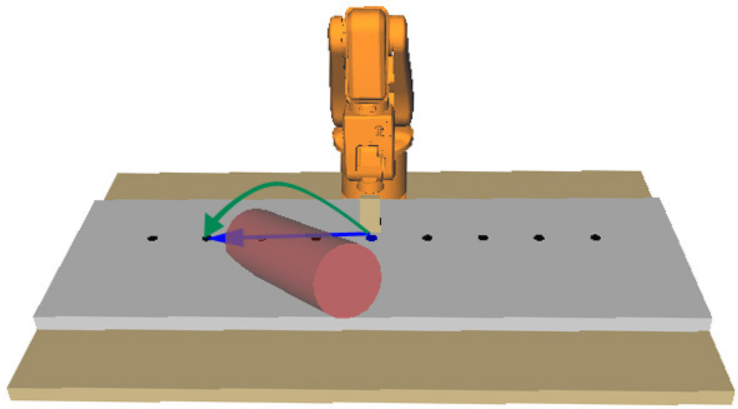

Figure 3. Virtual counterpart to the workspace of Figure 2

of a high fidelity localization uncertainty model, developed specifically for space application, to adapt robot motion planning. This will be done by using the localization probabilities given by the model along with the probabilistic models of which portions of the workspace the human worker is expected to occupy, as described in the previous section. These two sources of information can be used to calculate with how much confidence the robot will be able to successfully navigate around the human worker.

IV. Potential Benefits

There are several reasons why incorporating such adaptations is beneficial. A robot which does not adapt to a human worker and simply performs a pre-set sequence has to stop any time the human worker is in the way of the robot's next task. Additionally, precedence complications could arise when the human worker performs a task which needs to be done prior to a certain robot action. For example, if the human places screws to be drilled by the robot in a sequence other than the robot's pre-programmed plan, the robot will have to sit idle until the human places a screw at the robot's anticipated drilling location. These problems could potentially lead to significant decreases in efficiency, especially if the pre-programmed sequence the robot is using is particularly different from the worker's preferred order of actions. Consequently, we expect that having a robot which can successfully predict the human's next action will lead to significant decreases in robot idle time, resulting in improved system efficiency.

Incorporating risk-aware motion planning on top of a system capable of predicting human actions produces additional benefits. A risk-aware robot is capable of not only moving to the right location at the right time, but also moving there in a way which avoids the portion of the workspace which the human worker is anticipated to occupy. We predict that having a robot which moves in such a manner will lead to significant increases in concurrent motion of the human-robot team, as the two partners are less likely to utilize interfering paths. By having more concurrent motion, the team will execute the task more efficiently, leading to potential decreases in task execution time.

Beyond objective improvements in efficiency, having an adaptive robot also has the potential to increase the human worker's comfort with the robotic assistant. We expect that by adapting to the human's next most likely action and using risk-aware motion planning, the number of times in which the robot will reach toward the same space as the human and have to come to an abrupt stop will decrease, thus benefiting human trust in the system.

Incorporating task localization uncertainty will ensure that the system will provide all of these benefits by making the system more robust. This is especially true in settings where localization uncertainty can have a significant impact, such as space environments, as mentioned in the introduction section. Using a probabilistic model for anticipating human actions as well as human and robot localization will lead to executions which minimize interference with the human, allowing all the previously mentioned benefits to be realized.

V. Pilot Experiment

While there are many possible benefits of having a robotic co-worker capable of predicting human actions and planning risk-aware motions, if one were to test a system with both of these capabilities, it would be difficult to determine the individual benefits of each component. Consequently, a small pilot study was designed in order to determine the benefit of using risk-aware motions alone while keeping the prediction capability of the robot constant. In order to achieve this, we assumed the system is capable of predicting the human's actions with perfect accuracy, and then observed the performance of two types of human-robot teams: one in which the robot used risk-aware motion planning, and one in which the robot moves to each location using a pre-programmed shortest path.

In order to emulate a system which predicts the human's next action with perfect accuracy, the screw placement sequence was predetermined and the subjects were told to place the screws in this specific sequence. The subjects were split into two conditions: one group worked with a robot which anticipated the human worker's next action and adapted by using risk-aware motion planning to move around the human in the workspace (adaptive condition); the other group worked with a robot which did not attempt to adapt its motions and simply moved to the targets using a shortest path motion (non-adaptive condition). In both groups, a safety procedure was running in the background which decreased the speed of the robot in an exponential manner when the separation distance between the human and robot fell below a certain threshold, and stopped the robot completely if this distance fell below an even lower designated threshold.

In the adaptive condition, the robot's motions needed to avoid the human worker. The risk-aware motion plans needed to achieve this were constructed offline using the Constrained Bi-directional Rapidly-Exploring Random Tree (CBiRRT) algorithm from the CoMPS suite in OpenRAVE.[1,3] For each possible pair of robot action and anticipated human location, a virtual environment was constructed with a cylinder placed at the anticipated human location to simulate expected human workspace occupancy, as show in Figure 3. A trajectory was then computed and then saved in a database to be used by the adaptive system. This offline motion planning method was used in order to ensure repeatability of motions, as the CBiRRT algorithm uses a Rapidly-Exploring Random Tree (RRT) which inherently produces different motions each time.

The robot used for this study was an ABB IRB-120. A PhaseSpace motion capture system was used to track the position of the human in the workspace to determine what actions the human worker is taking, as well as to provide the safety system described above with information needed to slow down and stop the robot when necessary.

The task performed by the subjects consisted of placing screws at eight locations on a table while the robot went over each placed screw with a brush to emulate applying a sealant. The workspace is shown in Figure 2. In order to see how an adaptive, risk-aware system differs from the non-adaptive system, let us take the example configuration shown in Figure 2. Imagine that the human worker has just finished placing the second screw, and is expected to place the third screw in the location directly to the right of the second screw. Based on this information, the system will construct the virtual environment shown in 3. If the robot already applied sealant to the first screw, its next action is to move to the second screw and apply sealant there. A non-adaptive robot will take the shortest-path route to the second screw, represented by the blue arrow in the figure. This will likely cause the robot to have to stop abruptly as the human worker reaches through its planned path. An adaptive, risk-aware robot will avoid such a conflict by selecting a path to its goal which moves around the region of expected human occupancy. One example of such a path is shown in the figure as a green arrow.

After the experiment, the subjects filled out a brief questionnaire to assess their satisfaction with the robot as a co-worker and to determine how safe they felt during task execution. The questionnaire contained seven questions to which they provided answers using a standard five-point Likert scale. The questions asked were as follows:

- I trusted the robot to do the right thing at the right time.
- The robot did not understand how I wanted to do the task.
- The robot and I performed the task well together.
- I felt safe when working with the robot.
- The robot moved too fast for my comfort.

- The robot came too close to me for my comfort.
- I trusted the robot would not harm me.

The two main hypotheses in this experiment were:

- $H1$. The adaptive system will lead to more concurrent motion, a higher average robot speed, and a shorter total execution time when compared to the non-adaptive system.
- $H2$. Subjects who worked with the adaptive system will be more satisfied with its performance as a co-worker and feel more comfortable and safer working with it than subjects in the non-adaptive condition.

VI. Results and Discussion

A total of five subjects participated in the pilot experiment. The subjects were randomly assigned to each condition, with three to the adaptive and two to the non-adaptive condition. The sequence selected for the study was a simple placing of the screws in order from left to right. This sequence was selected, because it provides an even mix of conflicting and non-conflicting motions and is easy to remember and execute.

There were three sources of data: data recorded by the system, a video of the task execution, and post experiment questionnaire responses. The system recorded information about robot speed and separation distance during task execution. The video was used to determine total task execution time, the percentage of concurrent motion, as well as for qualitative analysis of the human worker's response to the robot. The questionnaire assessed the human workers' satisfaction with the robot as a co-worker and how safe they felt working with it.

As expected, the mean separation distance between the human and robot during task execution was higher in the adaptive condition, as the robot was actively evading regions which were expected to be occupied by the human. The average separation distance for the non-adaptive condition was 20.2 cm while in the adaptive condition this distance was 19.9% higher at 24.2 cm.

The robot speed is defined with the use of a speed multiplier given by the safety system, which is a value in the range of 0 to 1. When the separation distance falls below the speed reduction threshold, an exponential function outputs a number in this range based on the remaining separation distance. The product of the preset robot speed, which is set to a constant for all motions in this task, and the speed multiplier is used to set the robot's actual speed. The mean velocity multiplier was 10.2% higher for the adaptive condition (0.97 for adaptive, 0.88 for non-adaptive), which indicates that not only was the overall separation distance higher, but also that the threshold distance which triggers a reduction in speed was reached less often by the adaptive system.

The adaptive system also performed better in terms of percentage of concurrent motion. The percentage of concurrent motion was calculated by analyzing the video of the task execution and dividing the number of seconds of concurrent motion by the time of task execution during which concurrent motion was possible. The portion during which concurrent motion was possible was defined to be from when the robot motion starts to when the human finishes placing the last screw. This is because it is possible for the human to get ahead of the robot and finish the task early while the robot still has a few screws to apply sealant to. During this portion of task execution, concurrent motion is not possible, since the human is done with his tasks, and so this segment is ignored in concurrent motion calculations. Based on these definitions, the percentage of concurrent motion for the non-adaptive condition was 68% which is 24% lower than the 92% obtained in the adaptive condition. This result indicates that the adaptive system's capability of avoiding locations where the human is expected to move to next allows the human worker to move toward successive target locations more freely and with less hesitation. This result was confirmed by reviewing the video recordings of the experiment and observing how the subjects in the two groups performed the task. One can see that the subjects in the non-adaptive condition initially attempt to work concurrently, but quickly switch to a strategy in which they wait and attempt to time their motions to evade the robot. This timing strategy is not present during task execution in the adaptive condition, with the subjects reaching toward successive target locations without waiting for the robot.

The results of the aforementioned metrics are summarized in table 1 below:

	Mean Separation Distance (cm)	Mean Velocity Multiplier	Percent of Concurrent Motion (%)
Adaptive	24.2	0.97	92
Non-Adaptive	20.2	0.88	68

Table 1: Comparison of team fluency metrics for both conditions

The adaptive system also performs better in another key metric: total task execution time. Despite the adaptive system's wide motions being much less direct than the straight-line paths of the non-adaptive system (see Figure 3), the human-robot teams in which the adaptive system was used completed the task on average in 41s, which is 13% faster than teams with the non-adaptive system at 47s, as can be seen in Figure 4 below. This shows that that the benefit of allowing for more concurrent motion and having the robot slow down less often outweighs the penalty of having the robot take less direct paths toward its targets.

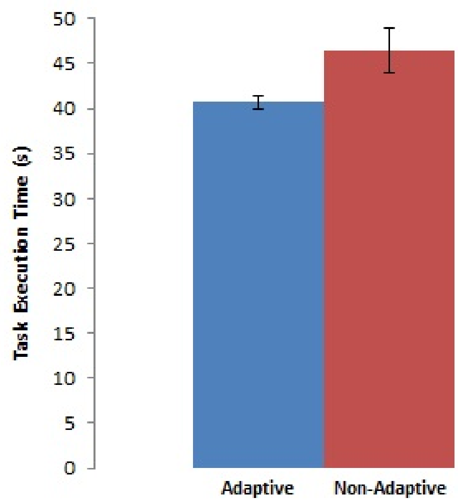

Figure 4. Comparison of total task execution time for the two conditions

The fact that teams which used the adaptive system completed the task faster is a very important result, as this indicates that a robotic co-worker can be made safer to work with, by having it maintain a larger separation distance, while simultaneously increasing the overall efficiency of the team.

The results discussed thus far all provide support for our first hypothesis, showing that there is a clear benefit to having the robot use risk-aware motions when working with a human partner.

In order to see if there is support for the second hypothesis, which states that human subjects working with the adaptive system will be more satisfied with their robotic co-worker and feel more safe and comfortable, the post experiment questionnaires were evaluated. In terms of the subjects' satisfaction with the robotic co-worker, the survey responses did not provide a clear result. While subjects indicated the adaptive system better understood how they wanted to perform the task when compared to the non-adaptive counterpart, they also indicated they trust the non-adaptive system more to do the right thing at the right time. The two systems received the same rating when the subjects were asked if they and the robot worked well together. Based on these three questions and the corresponding responses, it appears there was no perceived difference in how well the robot performed as a teammate. One possible explanation for this trend, in light of the previously mentioned execution strategies employed by the subjects in the two conditions, is that the strategy used by the subjects in the non-adaptive condition, namely to wait and carefully time their motions to avoid the robot, is a quick subconscious adjustment which is not consciously perceived by the subject as a burden.

When looking at the second aspect of the second hypothesis, perceived comfort and safety when working with the robot, the non-adaptive system actually received higher marks than the adaptive system. In the adaptive condition, subjects indicated the robot moved too fast more often, whereas in the non-adaptive condition they trusted the robot would not harm them and they reported they felt safe working with the robot. Both systems received the same ranking when asked if the robot came too close for the person's comfort. These results suggest that overall the adaptive system was perceived as less safe by the subjects. This is an interesting trend given that, in the adaptive condition, the robot kept a larger distance from the subjects and actively avoided producing conflicting motions. One possible explanation for why the subjects in the adaptive condition reported feeling less safe, is that the wide motions used to avoid the human were less predictable than the short straight-line paths of the non-adaptive system. A robot which moves less predictably can potentially be perceived as less safe, even if the motions are technically safer due to a larger separation distance and less conflicting motions.

VII. Conclusions and Future Work

Based on the results of the pilot study, one can see that the introduction of risk-aware motion planning can lead to significant increases in team efficiency by encouraging more concurrent motion which leads to shorter task execution times. Now that the effect of risk-aware motion planning decoupled from the potential benefits of predicting human actions has been studied, the next step is to develop a full adaptive system which combines human action prediction and risk-aware motion planning while compensating for localization uncertainty. Once this system is developed, we intend to run a large-scale user study to verify its anticipated benefits. We propose an experiment in which subjects will work on the same task as in the pilot study, but this time a virtual environment will be constructed in real time to simulate probability distributions over likely human actions and plan risk-aware robot motions. The robot's motion and task selection will be determined by an algorithm which utilizes the human action and localization uncertainties.

Subjects will be assigned to one of two conditions, one in which the robot adapts based on action prediction, and one in which it simply follows a pre-programmed sequence. Additionally, the group of subjects in which the robot adapts will be divided into several sub-groups of varying levels of uncertainty in human actions as well as uncertainty in localization of both the human and robot workers. While the human worker and robot work together on the task, several key parameters will be measured, including: robot idle time, the amount of concurrent motion, the number of times the robot was forced to stop due to interference with the human, and total execution time. Additionally, subjective measures of human comfort and satisfaction with the robotic assistant will be measured with the use of a questionnaire. The data collected will then be analyzed to determine the merit of the adaptive system as compared to a standard non-adaptive robot. As an additional benefit, we anticipate the results will be used to inform the design of tracking and localization systems for human-robot collaboration in space environments.

References

[1] Dmitry Berenson, Siddhartha S Srinivasa, Dave Ferguson, and James J Kuffner. Manipulation planning on constraint manifolds. In *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*, pages 625–632. IEEE, 2009.

[2] Antonio Chella, Ignazio Infantino, and M. Donatella Guarino. Localisation and description of the movements of a robotic arm for space applications by active contour techniques on image sequence. In *Proceeding of the 6th International Symposium on Artificial Intelligence and Robotics & Automation in Space: i-SAIRAS 2001,*, Quebec, Canada, 2001.

[3] Rosen Diankov and James Kuffner. Openrave: A planning architecture for autonomous robotics. *Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-08-34*, page 79, 2008.

[4] Myron A. Diftler, Joshua Mehling, Muhammad E. Abdallah, Nicolaus A. Radford, Lyndon B. Bridgwater, Adam M. Sanders, Roger Scott Askew, D. Marty Linn, John D. Yamokoski, Frank Permenter, Brian K. Hargrave, Robert Platt, Robert T. Savely, and Robert O. Ambrose. Robonaut 2 - the first humanoid robot in space. In *IEEE International Conference on Robotics and Automation, ICRA 2011, Shanghai, China, 9-13 May 2011*, pages 2178–2183. IEEE, 2011.

[5] Anca Dragan and Siddhartha Srinivasa. Formalizing assistive teleoperation. 2012.

[6] Anca D Dragan and Siddhartha S Srinivasa. Assistive teleoperation for manipulation tasks. In *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, pages 123–124. ACM, 2012.

[7] L. Ekroot and T.M. Cover. The entropy of markov trajectories. *Information Theory, IEEE Transactions on*, 39(4):1418 –1421, jul 1993.

[8] A. Kraussling. A novel approach to the mobile robot localization problem using tracking methods. In *Proceedings of the 13th IASTED International Conference on Robotics and Applications*, RA '07, pages 107–112, Anaheim, CA, USA, 2007. ACTA Press.

[9] Jim Mainprice and Dmitry Berenson. A human-robot collaborative manipulation planning framework that reasons on early prediction of human motion. In *Intelligent Robots and Systems, 2013. IROS'13. IEEE International Conference on.* IEEE, 2013.

[10] Stefanos Nikolaidis and Julie Shah. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In *To Appear in the IEEE/ACM International Conference on Human-Robot Interaction*, March 2013.

[11] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach.* Pearson Education, 2003.