An Adjustable Autonomy Paradigm
for Adapting to Expert-Novice Differences*

Bennie Lewis, Bulent Tastan, and Gita Sukthankar

Abstract—Multi-robot manipulation tasks are challenging for robots to complete in an entirely autonomous way due to the perceptual and cognitive requirements of grasp planning, necessitating the development of specialized user interfaces. Yet even for humans, the task is sufficiently complex that a high level of performance variability exists between a novice and an expert’s ability to teleoperate the robots in a sufficiently tightly coupled fashion to manipulate objects without dropping them. The ultimate success of the task relies on the skill level of the human operator to manage and coordinate the robot team. Although most systems focus their effort on forging a unified connection between the robots and the operator, less attention has been spent on the problem of identifying and adapting to the human operator’s skill level. In this paper, we present a method for modeling the human operator and adjusting the autonomy levels of the robots based on the operator’s skill level. This added functionality serves as a crucial mechanism toward making human operators of any skill level a vital asset to the team even when their teleoperation performance is uneven.

I. INTRODUCTION

Multi-robot systems can be very useful both for performing jobs that are beyond the capability of a single robot and speeding task completion through parallelization of effort [1]. Yet managing a robot team can be overwhelming for a single human operator, and improved telepresence is not necessarily a solution for this problem since the operator must maintain situational awareness over the whole team, rather than a single robot. The use of adjustable autonomy to reduce operator workload has shown promise in many multi-robot tasks since the operator’s effort and attention is used sparingly during critical sections of the task [2], [3].

In this paper, we present an adjustable autonomy approach for the challenging problem of multi-robot manipulation. The robots execute a lift and delivery task, under the guidance of a human operator. The teleoperation interface must support the user who is directing the robot’s navigation, manipulating objects with an arm and gripper, and coordinating the robots to jointly deliver objects to the goal. Failure to manage and coordinate pickups can lead to dropped objects and slow task completion times. Previous work in this area of multi-robot user interfaces has focused on improving the operator’s use of time and effort by detecting neglected robots [4] and improving coordination through the use of teamwork proxies [5].

However, one issue is that the same interface may not work equally well for users with different skill levels; it may not be the case that “one size fits all”. A possible approach to this problem is to allow the users to configure the interface through the use of programmable macros [6]. Here, we suggest that these expert-novice differences can be automatically detected after a short period of use and used to guide command decisions from the adjustable autonomy module. However, rather than simply mapping user expertise onto a single axis of competence, we model the user’s expertise on the separate task components of navigation, manipulation, and coordination. Based on previous user experiences, we have observed that many of the operators perform extremely well during one section of the task, while doing poorly on another. Having multiple axes of competence allows us to model users that fit this profile and increase robot autonomy to bolster the operator’s weaknesses.

In this paper we describe an adaptive user interface for adjusting the autonomy of the robots based on the operator’s skill level on three separate axes of competence. We present a paradigm for learning a model of the user’s competences from a short example teleoperation trace. In our multi-robot manipulation task, the human operator coordinates a team of two mobile robots to lift objects using an arm and gripper for transport to the goal location. The household environment contains a assortment of small and large objects, some of which can be transported by a single robot and others that require both robots to lift. Figure 1 shows a picture of the team of robots cooperatively moving an object that cannot be carried by a single robot. This cooperative pickup task is an important component of many potential applications of multi-robot systems, including cooperative assembly [7], home service robot teams [8], urban search and rescue [9], and patient recovery robot teams.

II. RELATED WORK

Four general approaches for improving human-robot interactions are:
1) improving visualization of the environment to reduce the cognitive load on the human operator [10];
2) building a multi-modal user interface that facilitates the tasking of robots [11];
3) creating adjustably autonomous robots that can operate effectively when the operator’s attention is elsewhere [12];
4) imbuing the robot with knowledge of human social conventions [13].

The guiding principle behind the first two approaches is the reduction of operator effort through good user design. In particular, 3D user interfaces can provide a more natural
metaphor for interactions with the physical world. Ricks, Nielsen, and Goodrich [14] present an ecological interface paradigm that fuses video, map, and robot pose information into a 3-D mixed-reality display. Results from their user studies show that the 3-D interface improves robot control, robustness in the attendance of delay, awareness of the camera orientation with respect to the robot, and the ability to perform search tasks while navigating the robot.

Operator neglect was identified as an important factor by Crandall et al. [15] who used an analysis of neglect and interaction time to predict the performance of a team of robots controlled by a single human. Wang and Lewis [5] theorize that in multi-robot control problems where tasks and robots are largely independent the operator sequentially neglects robots until their performance deteriorates sufficiently to require new operator input. This leads to poor performance in tasks with higher coordination demands, such as when the robots have differing sensing capabilities. Introducing a teamwork proxy [16] that can enable the robots to coordinate among themselves was shown to successfully increase robot autonomy and decrease demands on the human operator. Operator neglect can also be detected using hidden state estimation techniques [4], [17] and compensated for by the robots.

Adjustable autonomy, having the robots alter their level of autonomy in a situationally-dependent manner, has been used successfully in human-robot teams [18], [19]. In this paradigm, the robots reason about the tradeoffs between disturbing the human user vs. the risk of task errors. Here, rather than focusing on the user’s interruption threshold or distraction level, autonomy is adjusted based on the user’s capability to perform different aspects of the task. Our adaptive user interface component analyzes human operator’s skill level based on a short teleoperation segment and modifies the level of robot autonomy.

Earlier work in this area has studied how an interface can be adapted to the user’s profiles and preferences. For example, Kawamura et al. [20] developed an agent-based architecture for an adaptive human-robot interface, and Ahmad et al. [21] have done work on adaptive user interfaces in educational systems. Adaptive intelligent tutoring systems modify the performance of the ITS in response to a model of the learner’s abilities [22]. However unlike adaptive intelligent tutoring systems, our user interface models but does not attempt to improve the user’s teleoperation skills. We believe that the problem of attempting to train the users in addition to compensating for their weaknesses, is an interesting area for future work.

III. ROBOT PLATFORM

To examine this problem of multi-robot manipulation, we constructed a pair of inexpensive robots by mounting a robotic arm and gripper on a mobile wheeled base. The Home and Urban Intelligent Explorer (HU-IE) system is designed to be proficient at picking up light objects in a household environment with either carpets or hard floors. Having the arms on separate robots makes the pickup task more challenging but allows the user to parallelize large sections of the delivery task. Our robot includes the following components: an iRobot Create, Acer Aspire One netbook, the NXT 2.0 Robotics kit, a Logitech Communicate STX webcam, Turtlebot shelves, and Tetrix Robotics parts. The total cost per robot is around US $1000. Figure 2 shows the robot architecture.

A. Base

The iRobot Create has a differential drive that allows left and right wheel speeds to be independently specified and two bump sensors for detecting physical collisions. In addition to the internal Create sensors, we added an ultrasonic sensor mounted on the claw of the robot to determine the distance between the claw and the pickup object along with an accelerometer to measure the arm angle. A small webcam mounted on the robot arm presents a first-person perspective to the user during teleoperation. An Acer netbook (Intel Atom 1.6 GHz processor with Windows 7) functions as a relay forwarding sensor information from the Create sensors and webcam to the user interface.

B. Manipulator

The arm on the HU-IE robot was created using the LEGO NXT Robotic Kit. It is 1.2 feet long and extends 8 inches in front of the robot. The arm is actuated using three motors and has an operating range of \(-45^\circ\) to \(90^\circ\) in elevation. At the end of the arm is a four tong claw with rubber grips capable of grasping objects sized for a human hand. Textrix Robotic Metal parts are used to bolt the arm to the iRobot Create and serve as the rigid structure of the arm. A NXT intelligent brick, containing a 32-bit ARM7 microprocessor, is used to control the arm and communicate with all the sensors and actuators. Commands from the user interface are sent directly to the arm via Bluetooth, bypassing the Acer netbook.
C. Mapping

The robots’ workspace is monitored using a separately mounted Microsoft Kinect sensor. The Kinect provides RGB-D data directly to the user interface which uses it to track and display the location of the objects in the area. The position of the robots, based on the internal Create odometry, is marked on an occupancy grid and verified with the Kinect sensor. A modified blob detection technique is used to detect the other objects in the environment.

IV. USER INTERFACE

The user views the environment and interacts with the robot team through our user interface running on a separate Dell XPS M1530 laptop computer (Figure 3). In this paper, we evaluate an adaptive version of the user interface that learns a model of expert-novice differences for the various aspects of the teleoperation task vs. a non-adaptive version. The baseline user interface provides the user with a mirror mode for simultaneously controlling both the robots in which the second robot simultaneously executes a modified version of the commands that the user has issued to the actively controlled robot. This enables the robots to cooperatively lift objects and drive in tandem to the delivery location.

The operator controls the robots using an Xbox 360 Gamepad controller as follows. The trigger buttons on the Xbox 360 controller are used to toggle between the two robots and to activate the mirror mode in the unmanaged robot. The A, B, X, Y buttons are used to drive the mobile base. The right button halts the actively managed robot. The left and right analog sticks control the elevation and azimuth, respectively, of the robot arm. The claw grip is controlled by the D-pad on the Xbox 360 controller.

A. Adaptive Interface Component

Layered on top of the basic user interface is an adaptive interface component that adjusts the robots’ autonomy based on a learned model of the user’s teleoperation competence. An assessment of the user’s teleoperation performance is performed offline and loaded into the adaptive interface component (Figure 4). The adaptive section of the user interface is structured as a multi-agent system containing the following elements:

- **Attribute Component:** Imports the attribute report generated offline describing the human operator’s competence on the three task axes of navigation, manipulation, and cooperation.
- **Operator Interface Agent:** Adjusts the commands passed to the robots based on the user model.
- **HU-IE Interface Agent:** Handles interactions with the robots.
- **Human Input Component:** Handles interactions with the human operator.
- **Status Component:** Gathers and updates the status information from the robots to be displayed on the user interface.

All adjustable autonomy decisions occur within the Operator Interface Agent, which takes the offline attribute report describing the human operator’s competence on the three task axes and modifies the teleoperation commands sent to the robots. In general, the lower the human operator’s skill level, the more the agent filters the commands that are passed to the robots.

B. User Modeling

To construct a model of expert-novice differences in teleoperation performance, we collected example teleoperation sequences from twelve users and clustered the data using a semi-supervised version of k-means. The goal of this process was to learn a model of user competence on the three axes of navigation, manipulation, and collaboration. We selected these three axes as being both an accurate representative of our previous experiences with users and well-suited to
inform adjustable autonomy decisions for the multi-robot manipulation task.

To model navigation proficiency we extracted the following features from the raw trace: 1) task completion time; 2) number of seconds the robots spent moving in each cardinal direction; 3) number of seconds robots were halted; 4) number of times the user reversed driving direction. For classifying manipulation competence, the features used were: 1) task completion time 2) number of backward and right-left robot movements 3) number of seconds the arm spent at high, mid, and low elevations 4) number of claw command switches. Backward and right-left movements were particularly significant since they were rarely used by expert users who were able to drive forward and lift the item in one smooth motion, without reverses and changes of direction. The features for classifying robot coordination include the same features used for manipulation plus the percentage of time the user controlled both robots.

We observed the performance of the users on a simplified teleoperation task and rated them as being either confident or not confident on an axis of performance. The results of k-means clustering with $k = 2$ and a Euclidean distance measure proved to be a good fit for our data. The accuracy on separating the training data set was 100% for the navigation axis, 91% for the manipulation axis, and 83% for the coordination axis.

C. Adjustable Autonomy

Based on the learned model of expert-novice differences on the three axes of teleoperation proficiency, the adaptive version of the user interfaces selectively modifies the autonomy of the robots. Users who are less confident on the navigation axis receive more help during sections of the task that involve driving the robots. Two additional functions are invoked:

**Auto goal return:** When a human operator has successfully picked up an object, based on the ultrasonic sensor readings and robot arm accelerometer, the Operator Interface Agent commands the robot to drive the object to the goal area. The A* algorithm is used to find the shortest path to the goal, while avoiding obstacles marked in the occupancy grid.

**Nearest object seeking:** Once an object is delivered to the goal, the Operator Interface Agent detects the nearest object and starts driving the robot in that direction. Any time that the robot is under autonomous operation, the human operator can retake control of the HU-IE robot by canceling the drive command. For novice human operators only, the system will reactivate the drive command during robot idle times. If the user is classified as being confident at navigation, the system does not reactivate the drive command.

For users that are classified as less confident at the manipulation sections of the task, the adaptive user interface autonomously adjusts the arm and the claw to help the user using the functions:

**Auto arm adjustment:** The robot arm needs to be at a certain angle relative to the target object for a successful grasp and lift. Based on arm accelerometer sensor data and Kinect object detection, the adaptive user interface
attempts to calculate the angle required for a successful pickup and adjusts the arm accordingly when an object is within a certain radius of the robot. The Operator Interface Agent observes the incoming commands, adds the required adjustments to the end of the command string, and displays it to the user before sending it to the robot.

Auto claw adjustment: If the ultrasonic sensor indicates that the grasp will not be successful, the mobile base and claw are autonomously adjusted to improve the grasp. Note that even though it is possible to autonomously calculate reasonable base, arm, and claw positions for grasping objects an expert human user can still outperform fully autonomous operation. Users who perform poorly on the coordination axis are experiencing difficulty in maneuvering the robots together and performing object lifts with both arms simultaneously. The adaptive user interface attempts to adjust the arm, claw, and base of both robots when they are within close proximity of the same pickup object using the auto arm adjustment and auto claw adjustment functions. This behavior is also triggered if the arms of both robots are not positioned evenly. A video of the system can be viewed at: http://youtu.be/hrFa12C0784.

V. EXPERIMENTAL METHODOLOGY

Our experiments were designed to evaluate the human operators’ ability to complete a set of indoor multi-robot manipulation scenarios under both the adaptive and non-adaptive version of the user interface. 20 users (8 male, 12 female) between the ages of 20 and 35 participated in the study. Before the user interface evaluation scenarios, all users were given 10 minutes of practice time and asked to complete three skill assessment tasks designed to measure the teleoperation performance on the axes of navigation, manipulation, and cooperation. Several of the subjects had prior experience playing Xbox games, but none of them had previous robotics experience.

Teleoperating Assessment Task 1: Each participant was allotted ten minutes to navigate a single robot through an obstacle course; the results of this task were used to classify the user’s navigation skill.

Teleoperating Assessment Task 2: Each participant was allotted ten minutes to lift a single small object; the results of this task were used to classify the user’s manipulation skill.

Teleoperating Assessment Task 3: Each participant was allotted ten minutes to lift a large box (shown in Figure 1)); the results of this task were used to classify the user’s cooperation skill.

Scenario 1: For the first scenario, the participant had to use the two robots to search the area and transport small objects (movable by a single robot) to the goal basket within 15 minutes. The environment contained three piles with five round shaped objects (shown in the left and center panels of Figure 5). The participant performed this scenario twice in randomized order, once with the adaptive interface and once with the baseline version.

Scenario 2: For the second task, the participants had to use the two HU-IE robots to search the area and transport awkward objects that required bimanual manipulation to the goal basket within 15 minutes. There were three piles with bimanual objects in this scenario (shown in the right panel of Figure 5). The participant performed this scenario twice in randomized order, once with the adaptive interface and once with the baseline version.

VI. RESULTS

In the results, we compare the performance of the adaptive vs. the non-adaptive version of the user interface. Figure 6 presents a comparison of the times required for each participant to complete Scenario 1 (small objects) and Scenario 2 (bimanual manipulation) under both experimental conditions. Table I summarizes the completion time results. We confirm that the improvements in completion time is statistically significant under a paired two-tailed t-test at the $p < 0.01$ level for both Scenario 1 and 2.

Figure 7 presents a comparison of the object drops by each participant in Scenario 1 (small objects) and Scenario 2 (bimanual manipulation) under both experimental conditions, the adaptive and non-adaptive user interface. Table II summarizes the number of dropped objects in each condition. We confirm that the reductions in dropped objects is statistically significant under a paired two-tailed t-test at the $p < 0.01$ level for both Scenario 1 and 2.

The figures show that for all of the participants (other than subject #10) the adaptive component improves the human operator’s performance measured by both task completion time and reductions in dropped objects. Our post-questionnaire indicated that 90% of the users had a strong preference for adaptive vs. the non-adaptive version of the user interface, and the remaining 10% expressed no preference between the two conditions.

Table III shows the results of the user modeling component of the system. The classifier learned from previous teleoperation traces identified half of the users as being expert

### Table I

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Adaptive Time ±σ (sec)</th>
<th>Non-adaptive Time ±σ (sec)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>340.9 ± 86.8</td>
<td>408.0 ± 79.8</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>600.7 ± 117.7</td>
<td>716.4 ± 150.7</td>
<td>yes</td>
</tr>
</tbody>
</table>

### Table II

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Adaptive Drops ±σ</th>
<th>Non-adaptive Drops ±σ</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.10 ± 1.45</td>
<td>3.25 ± 1.74</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>3.85 ± 2.58</td>
<td>7.85 ± 3.15</td>
<td>yes</td>
</tr>
</tbody>
</table>
Fig. 5. The two robots operate within a 6.3′ × 6.4′ household area and move objects from various piles to the goal area. Scenario 1 (left, middle) contains piles of small objects that can be moved with a single robot, whereas Scenario 2 (right) contains objects that require bimanual manipulation.

Fig. 6. Time to complete Scenario 1 (left) and Scenario 2 (right) in minutes for each subject (x-axis). All of the participants (except subject #10) experience time improvements with the adaptive version of the user interface.

Fig. 7. Number of objects dropped by each subject (x-axis) in Scenario 1 (left) and Scenario 2 (right). All of the participants (except subject #10) experience reductions in dropped objects with the adaptive version of the user interface.

Fig. 8. Expert/novice differences in frequency of command utilization for navigation, manipulation, and collaboration. Beginners (blue) utilize the stop command more frequently than experts (red) both when driving the robot base or moving the arm. They open the claw more frequently than experts who require fewer attempts to lift objects. In contrast, experts issue the close claw and forward drive commands more frequently than the beginners.
at navigation and manipulation, and slightly fewer as being experts at the cooperative sections of the task. Figure 8 shows the relative distribution of commands issued by experts vs. novices using the non-adaptive version of the interface.

Several interesting facts emerge: 1) novices more frequently issue _stop_ commands for the robot base, whereas experts more frequently use _forward_; 2) novices open the claw more often than expert users, probably following object drops; 3) experts more regularly issue the _up_ command to the robot arm, whereas novices more frequently stop the arm in its trajectory. The classifier is able to utilize these differences in command distribution to accurately learn a model of expert/novice differences along the three teleoperation axes. In most cases, the users’ self-reported level of confidence on each axis agreed with the classifier. However, we believe that relying strictly on self-reports of expertise in undesirable, particularly in situations where the users’ have greater external motivation to claim expertise.

VII. CONCLUSION AND FUTURE WORK

Synchronizing coordination and delegating task assignments across multiple robots can be a difficult task for even an expert human operator. Multi-robot manipulation tasks are particularly sensitive to poor coordination since tight temporal coupling is required to avoid object drops. Yet capable human operators can easily outperform a fully autonomous system since they are able to more reliably solve grasp planning problems from limited sensor data. Adjustable autonomy paradigms show particular promise in this domain since they free the operator to focus attention on critical task segments.

In this paper, we demonstrate the utility of an adaptive user interface that adjusts the robots’ autonomy based on expert-novice differences. A user model of teleoperation competence on three axes of performance (navigation, manipulation, and coordination) is learned from short example tasks. The adaptive user interface modifies the robots’ autonomy in a task specific way, based on the operator’s skill level. In our user study, the proposed user shows statistically significant improvements in reducing the task completion times and dropped objects. An interesting avenue for future work is applying the same user modeling techniques as part of a teleoperation training system to instruct users in the principles of robot teleoperation.

REFERENCES