

# Simultaneous Teleoperation of Multiple Social Robots

Dylan F. Glas<sup>1</sup>

Takayuki Kanda<sup>1</sup>

Hiroshi Ishiguro<sup>2</sup>

Norihiro Hagita<sup>1</sup>

<sup>1</sup>Intelligent Robotics and Communication Laboratories  
ATR

2-2-2 Hikaridai, Keihanna Science City  
Kyoto, 619-0288 Japan

{dylan,kanda,hagita}@atr.jp

<sup>2</sup>Intelligent Robotics Laboratory  
Osaka University

2-1 Yamada-oka, Suita  
Osaka, 565-0871 Japan

ishiguro@ams.eng.osaka-u.ac.jp

## ABSTRACT

Teleoperation of multiple robots has been studied extensively for applications such as robot navigation; however, this concept has never been applied to the field of social robots. To explore the unique challenges posed by the remote operation of multiple social robots, we have implemented a system in which a single operator simultaneously controls up to four robots, all engaging in communication interactions with users. We present a user interface designed for operating a single robot while monitoring several others in the background, then we propose methods for characterizing task difficulty and introduce a technique for improving multiple-robot performance by reducing the number of conflicts between robots demanding the operator's attention. Finally, we demonstrate the success of our system in laboratory trials based on real-world interactions.

## Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces - *Interaction styles*; I.2.9 [Artificial Intelligence]: Robotics

## General Terms

Algorithms, Design, and Human Factors.

## Keywords

Adjustable autonomy, communication robots, human-robot interaction, multiple robots, social robots, supervisory control

## 1. INTRODUCTION

As rapid progress is being made on all frontiers of robotics technology, many of the key components necessary for developing socially-situated autonomous robot systems are falling into place. Field trials of social robots placed in real-world environments such as museums [1, 12, 15], schools [7, 10, 11], and train stations [8] have shown great success, provided valuable research opportunities, and provided insight into real-world social phenomena which cannot be observed in the lab.

However, as inspiring and exciting as it is to see robots operating in the field, the inescapable reality is that social dynamics are

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HRI'08, March 12–15, 2008, Amsterdam, Netherlands.

Copyright 2008 ACM 978-1-60558-017-3/08/03...\$5.00.



Figure 1 : A robot providing route guidance in a mall

complex, and today's technology is not yet truly capable of supporting a fully-autonomous robot playing a real role in society. Any robot will eventually find itself in unanticipated circumstances, where failure to respond appropriately could lead to socially awkward, money-losing, or even dangerous situations.

A field trial we recently conducted at a Japanese shopping mall illustrates an example of a social robot application. We placed a humanoid robot in a central public space in the shopping mall for several hours a day, where it chatted with visitors and provided information and route guidance to locations within the mall. Customers were excited by the engaging interactions, and people crowded around the robot every day, waiting for a chance to talk with it (Figure 1).

Although a large part of the attraction of social robots is their ability to “understand” natural language and engage people interactively, this task is still largely beyond the capabilities of today's robots to achieve without a human operator. Thus, in this trial, we used the “Wizard of Oz” method, employing a human operator to control the robot remotely.

Field trials of robotics technology in social settings have often involved some degree of remote control [6, 17]. This is acceptable in the context of research, where technologies are generally novel and untested, but in the commercial world, a robot requiring one or more full-time operators would offer little advantage over simply hiring a human employee to do the same work. Capable as these robots may be, so long as the operator-to-robot ratio remains 1:1 or greater, they will not be ready to leave the laboratory.

In other fields of robotics, such as search-and-rescue or space exploration, teleoperation and the operation of multiple robots is an active field of research, encompassing many important topics such as partial autonomy, adjustable autonomy [3, 5], controlling interfaces [4, 14], fan-out [2], workload [9], and situational awareness for monitoring [13]. Yet this concept has not yet been explored in the field of social human-robot interaction.

**Table 1. Differences in teleoperation between navigation (fundamental tasks for mobile robots [16]) and social interaction.**

	Navigation	Social interaction (this study)	New problems in social interaction
Operator’s role	Obstacle avoidance. Giving current position, path, goals.	Understanding of user’s intention and providing required service	
Source of input to operator	Scenery + Map	Audition (+scenery)	Cannot monitor multiple sources
Operator’s output (low level control)	Velocity	Utterance, gesture, +(body orientation and position)	Typing, controlling many DOFs for gesturing very slow
Operator’s output (abstracted control)	Position (destination)	Behavior (combination of utterance and gesture)	Difficult to prepare for minor cases in advance
Consequence of ignoring errors caused by autonomy	Crash into obstacle, or lose the robot	Person might get lost, buy wrong product, or receive wrong service.	Definitely we should not ignore errors in either case.
Can robots wait after an error detected?	Yes	No. Users might soon leave if a robot stops.	An operator should take control of the robot immediately.
Can robots anticipate the timing of possible error?	No	Yes	Most errors are from speech recognition, often after the robot asks a question.

In this paper, we address the unique challenges of multi-robot operation in the case of social robots. We have implemented a multiple-robot teleoperation system enabling a single operator to monitor and control several communication robots at once, directing their behaviors as they answer questions and provide information. We developed an interface which enables an operator to control one robot while simultaneously monitoring several others, and we employ a technique we call “proactive timing control” to reduce conflicts between the robots for the operator’s attention. We have tested this system in laboratory trials and demonstrated that a single operator is able to successfully control up to four robots at once, as they simultaneously engage in human-robot communication interactions.

## 2. DESIGN CONSIDERATIONS

In order to develop an efficient system for multi-robot teleoperation, we will examine several aspects of the operator-robot system. As extensive research has been done on teleoperation within the field of robot navigation, we have summarized how the issues in teleoperation for social interaction differ from those regarding teleoperation for navigation (Table 1). Although there are some overlaps, such as the issue of adjustment of autonomy, the biggest difference seems to be the time-critical aspect of social interaction.

The role of the operator in a social robot system is generally similar to that in other semi-autonomous systems, in that the operator directs high-level goals and identifies errors that the system cannot detect on its own. Social robots today can reliably perform many functions autonomously, such as detecting and tracking humans or performing output-related tasks such as speaking and gesturing. It is in the recognition and interpretation of verbal and nonverbal communication and the ability to make common-sense judgments that an operator can add the greatest value. Particularly when dealing with unexpected utterances and interruptions, a human operator’s intuitive ability to understand social situations and quickly determine an appropriate response is indispensable. By choosing appropriate high-level behaviors or handling the situation through direct low-level control, the presence of an operator can enable a robot to perform effectively in situations the robot was never designed to handle.

### 2.1 Adjusting and Using Autonomy

Before addressing the integration of multi-robot systems, it is important to consider how an operator will interact with the robot, monitoring situations and taking over control when necessary.

#### 2.1.1 Controlling a Robot

Improvements in the efficiency of robot control can be made possible through layers of abstraction. For example, an operator could specify the individual joint angles for the robot’s arm at a low level, or achieve the same result by giving the robot a high-level command, e.g. “point to the left”. Most robot systems already incorporate this kind of abstraction to some degree. Joint angles can be grouped into poses, poses grouped into motions, motions and utterances grouped into behaviors, behaviors and transition rules grouped into episodes, and so on. A similar abstraction can be found in the case of navigation, referred as adjustable autonomy [3, 5].

As layers of abstraction are added to the system, the robot usually becomes able to function with a higher degree of autonomy, thus reducing the workload for the operator. When high-level functions are not prepared for a situation, the operator can use low-level functions instead. For example, if there is no behavior prepared for giving directions to a Japanese restaurant, an operator might directly type phrases for the robot to say and control the arms manually to point the way.

#### 2.1.2 Correcting a Robot’s Recognition

An operator can also choose to correct a robot’s sensory recognition errors, rather than completely taking over control of its behaviors. For example, an operator observes a scene where a user says the words “Japanese restaurant”, but the speech recognizer fails to pick it up. If the robot has behaviors in place to react to those words, the operator can correct the robot’s speech recognition results and allow the robot to complete the interaction as usual. This is less effort than taking over behavior control in order to generate a guiding behavior for directing the user to a Japanese restaurant.

## 2.2 Managing Operator Attention

In a multi-robot teleoperated system, the operator’s attention is a scarce resource. It is important to provide the right balance of information to enable the operator to efficiently monitor several robots at once.

### 2.2.1 Status Monitoring

One of the most important roles of the operator is to monitor and react to unexpected problems in a timely manner. Consider an error such as a speech recognition failure or a user unexpectedly speaking out of turn. If the error is caught immediately, the operator can manually redirect the behavior flow, and the interaction will still seem natural. If the error is missed, the interaction may

become more confusing, awkward and frustrating the longer the error remains uncorrected.

Enabling the operator to quickly discover errors like these is largely an issue of user interface design. State information about each of the robots should be delivered to the operator through background channels in such a way as to make errors easily recognizable.

### 2.2.2 Timing Control

Another way to help manage the operator's attention efficiently is to indicate situations in which the risk of error is high.

To this end, we divided each interaction into *critical sections*, where there is a high risk of error and thus a high likelihood that operator assistance will be needed, and *non-critical sections*, which can safely be performed autonomously. Critical sections include tasks such as asking questions or other behaviors where a response from the user is expected, whereas non-critical sections include behaviors such as greeting, talking, and giving directions.

In a multi-robot system, an operator can use this information to efficiently focus attention on the robot or robots with the highest risk of interaction error. The challenge here comes when two or more robots enter critical sections at the same time. We will consider two possibilities for handling such conflicts, to which we will refer as *reactive* and *proactive* timing control.

#### 2.2.2.1 Reactive Timing Control

The first solution, *reactive timing control*, is for each robot to notify the operator of the critical section, then to proceed in its interactions regardless of the state of the operator or other robot(s). If the interaction reaches a point where the robot is unable to respond to the user without operator intervention, the robot must stall for time until the operator is available.

For example, if a user asks a robot for information, but the robot cannot understand the request, the robot can ask the user to wait, buying time until an operator becomes available to answer the question. Such a behavior might leave a negative impression on the user, since the user is waiting for an answer to the question and may become impatient.

#### 2.2.2.2 Proactive Timing Control

A second possible solution to this problem is *proactive timing control*, in which interactions are planned in order to prevent such conflicts from arising at all. One means of achieving this is for each robot to send a reservation request to the operator before a critical section begins. If the operator accepts, the robot can proceed to the critical section. Otherwise, the robot performs other behaviors in order to delay entry into the critical section.

From a user's perspective, this technique is fundamentally different from reactive control. In the reactive case, the delaying behaviors are executed *after* the user's "turn" in the conversation, that is, after the user has made a request or asked a question. There, the user is understood to have the initiative, and the robot is expected to react.

In the proactive case, however, the delaying behavior is executed *before* the user has spoken, while it is still the robot's "turn" to speak. The robot has not yet relinquished the initiative, and thus the extra behaviors naturally integrate into the flow of conversation.

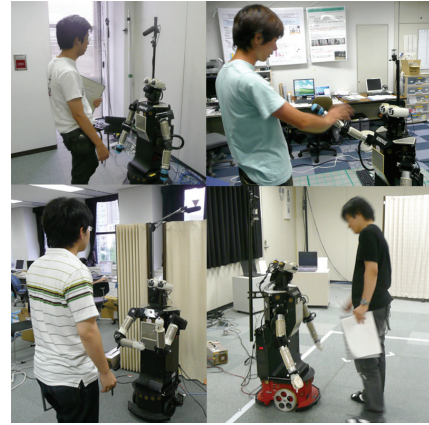


Figure 2: Four Robovie II's operated simultaneously

## 2.3 Task Difficulty

Finally, it is useful to have some means of quantifying the capability of a system to perform a task autonomously, so that relative difficulty of tasks and systems can be compared. Several fundamental metrics describing factors that bias task performance are well summarized by Steinfeld et al. [16], while task-related metrics for social interaction have not yet been explored well. Here, we describe two metrics important for social interaction in service tasks.

### 2.3.1 Situation Coverage

The first metric we propose is Situation Coverage (SC). We define a situation to be "covered" if the system would autonomously execute the correct behavior given perfect sensor inputs. For example, if a user asks the robot where to buy some broccoli, but the robot is not programmed to react to the word "broccoli", the situation is considered not to be covered, even if the robot does have the ability to direct the user to a supermarket.

Hence, this value is an *a posteriori* evaluation of task difficulty which takes into account both the capabilities of the system and the expectations of users. Given the complexity of real social situations, it is usually impractical to attempt to achieve 100% SC. Instead, the strategy for use of partial autonomy should be to design the robot for anticipated situations, perhaps achieving an SC of 70%, and then to consider the remaining 30% to be "unknown", relying on operator assistance for those situations.

The significance of SC in the context of multi-robot operation is that it represents an upper bound to the capacity of the system to operate autonomously. A system with an SC of 70% will be able to successfully complete no more than 70% of its tasks autonomously, and will require operator intervention at least 30% of the time.

### 2.3.2 Critical Time Ratio

A second metric that is useful for evaluating degree of autonomy is the Critical Time Ratio (CTR). This is defined as the ratio of the amount of time spent in critical sections to the total duration of an interaction. For tasks with a low CTR, the likelihood of two robots entering a critical section at the same time is correspondingly low, and thus timing control behaviors will seldom be necessary. Tasks with a high CTR are more likely to conflict, which can lead to higher wait times for users and a heavier workload on the operator.

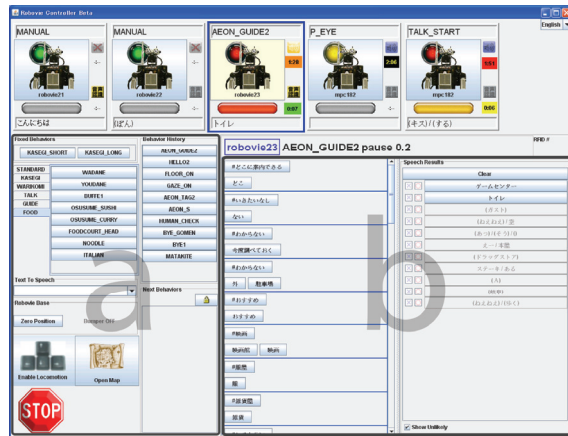


Figure 3: User Interface for Multiple-robot Control

### 3. IMPLEMENTATION

The teleoperation system we implemented consisted of four Robovie II humanoid robots (Figure 2). The robots were placed in separate areas of our laboratory, and video cameras were set up to provide live video feeds to a monitor next to the operating console.

#### 3.1 Robot System

##### 3.1.1 Robot Platform

Robovie II is a humanoid robot platform developed for human-robot interaction research. It is capable of humanlike expressions with a head that can be moved with 3 DOF (degrees of freedom), arms with 4 DOF each, eye cameras with 2 DOF each, and a wheeled base for locomotion.

Each robot has two color CCD cameras, one omnidirectional camera, a microphone, and several touch sensors. The microphone was connected to a speech recognition system. An environmental sensor network for tracking humans was simulated by a keyboard interface, with one key indicating that a subject was present in front of the robot, and another key indicating that the subject had left.

##### 3.1.2 Behavior Control

The robot control software used in our system was adapted from behaviors and data developed for our field trial in the shopping mall.

The design of the software framework encapsulates robot behaviors (short sequences of motions and utterances) into discrete modules. Sets of rules called “episodes” then define the transitions between these modules based on sensor inputs and robot state. With this framework, if we theoretically assume no errors in sensor recognition and user behavior only within the limits of SC, it is possible for the robot to execute any length of behavior chains with full autonomy.

To give a simple example, when a robot in an idling state detects a person approaching, the episode rules may trigger a transition from the idling behavior to a greeting behavior. After greeting the person, the next behavior might be to offer route guidance and wait for a response. The transition rules would then choose the next behavior based on input from the speech recognition system. If the person asked for directions to a bookstore, and if we assume the speech recognition system correctly recognized the word “bookstore”, the system would then transition to the module for giving directions to the bookstore.

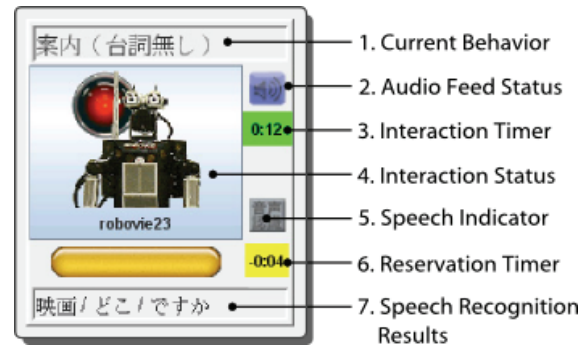


Figure 4: Background Robot Display

#### 3.2 User Interface

The operator’s interface software was developed to enable the operator to control one robot (referred to here as the “active” robot) while monitoring the others in the background. One challenge of developing this interface was to provide just the right amount of information about the background robots to enable the operator to identify errors and unexpected situations, but not so much information as to be overwhelming.

The interface console (Figure 3) is divided into two areas. The panels in the top area represent all connected robots, and the bottom area contains a more detailed display of the active robot.

##### 3.2.1 Active Robot Interface

The main area of the interface is designed for controlling the active robot. Within this space, the controls on the left side (area “a” in Figure 3) allow the operator to directly control the robot, and the controls on the right side (area “b” in Figure 3) enable the operator to correct or confirm results returned by the robot’s speech recognition system.

###### 3.2.1.1 Direct Robot Control

One function of this interface is to provide both low- and high-level control, as described in Section 2.1.1. The left side of the main screen area (area “a” in Figure 3) is used for controlling the robot. The top half of the panel allows the operator to select behaviors by category. The bottom half includes a text box for directly entering phrases for the robot to speak, as well as buttons for locomotion control and a map with which the operator can quickly choose from 136 possible guide destinations. Finally, an emergency stop button allows the user to abort the robot’s current behavior and stop all motors.

###### 3.2.1.2 Sensor Data Correction

The right side of the active robot control panel (area “b” in Figure 3) is used for correcting sensor recognition results, as described in Section 2.1.2. This area contains two columns of controls dealing with speech recognition.

The column in the center shows all possible key words and synonyms expected by the current behavior. These buttons allow the operator to override the speech recognition system and enter a speech result directly.

The column on the right allows the operator to designate speech recognition results as correct or incorrect, or to select an utterance from recent history, which can be useful when the user speaks out of turn.

### 3.2.2 Background Robot Display

The panels along the top of the console represent all of the robots currently available to the operator. Each panel presents information for the operator to monitor, in order to identify problems or potential problems quickly and easily. Figure 4 shows a detailed view of this display.

#### 3.2.2.1 Status Monitoring

Several kinds of status information are provided to help the operator watch for and anticipate interaction failures (see Section 2.2.1). The Current Behavior display (1) at the top of the panel shows the name of the behavior the robot is currently executing. Below it, the Audio Feed Status button (2) indicates whether the robot's streaming audio server is online and connected. The Interaction Timer (3) measures the length of the robot's interaction with a single user, enabling the operator to identify at a glance how long each of the robots in the system has been interacting. Below that, the Speech Indicator (5) lights up when activity is detected from the speech recognition system. Speech Recognition Results (7) are displayed at the bottom of the robot status panel. These can be valuable in case the operator is unable to switch to a robot in time to hear what the user is saying.

Audio and video are also indispensable resources for detecting failures. The operator can listen to a live audio stream from the currently selected robot through headphones. Video has not been integrated with the console yet, so video feeds observing all four robots are displayed on an adjacent monitor. In the future, additional sensor information, such as human position estimates or customer data, could also be presented on an additional screen.

#### 3.2.2.2 Reactive Timing Control

The Interaction Status display (4) is a picture of the robot with a green, red, or yellow light behind it. The color of the light behind the robot indicates whether the robot's state is "idle" (green, not interacting with a person), "active" (yellow, interacting with a person), or "critical" (red, meaning that the robot has entered a critical section). In order to attract the operator's attention, the red light is accompanied by a flashing red background if this robot is not the active robot. An audible beep is also sounded when any robot enters a critical section. This notification is a part of reactive timing control, as described in Section 2.2.2.1.

#### 3.2.2.3 Proactive Timing Control

The Reservation Timer (6) and the button next to it are part of a proactive timing control system, as described in Section 2.2.2.2. In our implementation, the robot notifies the operator several seconds before a critical section is scheduled to begin. The robot's Interaction Status icon flashes red, and the Reservation Timer begins counting down to the point where the operator must choose whether to commit to the robot or not. By pressing the button, the operator can accept the reservation, and the robot will continue to the critical section. If the operator has not pressed the button by the time the counter reaches zero, the robot will execute other behaviors to delay the entry point to the critical section, resetting the Reservation Timer appropriately. The system allows the operator to commit to a robot other than the current active robot, enabling the operator to plan ahead and commit to a second robot while finishing an interaction with a first.

### 3.2.3 An Example Interaction

Here we will describe an example of a typical multi-robot control session from our experiment. In this example, the operator is controlling three robots, and the system is using reactive time control, *i.e.* there is no attempt to prevent conflicts between robots demanding the operator's attention at the same time.

First, Robot 1 detects a person approaching. As it begins a greeting behavior, its Interaction Status light changes to yellow and the Countdown Timer on the robot's status panel begins counting down until the robot expects the human to speak.

The operator clicks on the robot's status panel to choose Robot 1 as the active robot, and the bottom half of the user interface refreshes to show Robot 1's current status, behavior history, and speech recognition results. The audio stream from Robot 1 is also piped to the operator's headphones, and the operator listens in as Robot 1 introduces itself, "My name is Robovie, and my job is giving directions. Where would you like to go?"

At this point, the operator notices that a person has approached Robot 3 as well. However, the operator stays focused on Robot 1, as its Countdown Timer is just reaching zero. The customer asks where to find an ATM. Unfortunately, due to background noise, Robot 1's speech recognition was unable to pick up the word "ATM", so the operator goes to the expected phrases panel and clicks on "ATM". Robot 1 then begins giving directions to the customer, and the operator quickly switches to Robot 3, whose countdown timer has almost reached zero.

By this time, a customer has approached Robot 2 and begins asking directions while the operator is still busy helping Robot 3. Robot 2's Interaction Status light flashes red. By the time the operator finishes helping Robot 3, the customer talking to Robot 2 has already finished speaking. Robot 2's speech recognition system has picked up the word "hamburger", which is displayed on its Speech Results display, but the robot has no mapping between that word and a location in the mall. The operator quickly switches to Robot 2, opens the map, and clicks on a restaurant that specializes in hamburgers. Robot 2 then gives directions to that restaurant, as the Interaction Status indicators for Robots 1 and 3 return to green.

## 4. EXPERIMENT

The primary goal of this study was to demonstrate whether our system could achieve the unprecedented task of a single operator simultaneously controlling multiple robots engaging in basic social interactions with humans. The secondary goal was to identify aspects of the system which merit further exploration, particularly those aspects which are unique to the operation of multiple social robots.

### 4.1 Scenario

In order to test the effectiveness of our system, we chose the task of route guidance as a realistic example of the kind of task a robot might be assigned to perform. It is easy to imagine a business such as a shopping mall, museum, or theme park placing a robot in a high-visibility location such as a central information booth. This task also lies in an interesting middle-ground between full predictability and open-endedness, and it provides a level of interactivity not found in primarily one-way interactions such as guiding visitors in a museum.

## 4.2 Experiment Design

### 4.2.1 Experiment 1

The first experiment was designed to be particularly difficult for the operator in order to push the limits of the system. To make the Critical Time Ratio as high as possible, only the minimum set of behaviors necessary to complete the route guidance interaction was used. The robot greeted the customer, who then asked for directions. The robot then provided directions, and the interaction was finished. To further ensure a high CTR, the interactions for Experiment 1 were run using reactive timing control.

### 4.2.2 Experiment 2

The second experiment was designed to test our system under less stringent, more realistic conditions. Interactions in this experiment were modeled directly on actual interactions from our field trial at the shopping mall, where the robot was given several topics to chat about each day, such as special events at the shopping mall, the robot's experiences at the mall's restaurants, etc. The robot randomly selected one of these chat behaviors for each customer.

For this experiment, we used the same behaviors as part of a proactive timing control system. When the robot began introducing itself to the visitor, it sent a reservation request to the operator. If the operator accepted the reservation, the robot proceeded to offer route guidance as usual. Otherwise, the robot executed one of the chatting or information providing behaviors instead, to delay the critical section entry point. When that behavior finished, the robot proceeded to the critical section.

It is important to note that these behaviors were not merely time-killing behaviors. When the robot spoke about these topics in the field trial, they were relevant to the customers, who enjoyed their interactions with the robot.

### 4.2.3 Experimental Procedure

We followed identical procedures for the two experiments. All trials were run in our laboratory, and the same four subjects, all of whom were research members of our laboratory, were used for all trials. The subjects were asked to imitate the visitors of shopping mall whom they had observed interacting with the robot. All subjects and the operator had been involved in controlling or observing the robot's interactions at the shopping mall trial.

To make the interaction more realistic, it would have been ideal to actually run the four robots in the shopping mall; however, the shopping mall cannot accept immature technology that could provide wrong information to a real customer. Thus, it was necessary to perform this experiment in our lab environment.

### 4.2.4 Conditions

Each experiment was run under five conditions. For the first four conditions, the operator controlled one, two, three, and four robots respectively. These will be referred to as the 1R, 2R, 3R, and 4R conditions for brevity. For the fifth condition, to which we will refer as the A condition, robots were run in a fully-autonomous mode with no operator intervention.

Originally we had intended to use a set of real interactions taken from the field study, but the SC for that sample (nearly 100%) provided too little of a challenge. Instead, we created a more difficult set of questions to reduce the SC to around 50%.

Forty-eight interactions were performed for each condition. Trials for the first four conditions were run in eight blocks of 24

interactions each, with breaks of at least five minutes between sessions. The trials were conducted in a pseudorandom balanced order to avoid any bias over time. The autonomous conditions were tested separately, in blocks of 24 interactions.

Each of the four subjects interacted with the robots a minimum of six times for each trial. Start times were staggered by 10 seconds each, and subjects proceeded at their own pace. Some finished earlier than others, so to prevent this from affecting the data (e.g., if one subject finished early, a four-robot trial would have become no different from a three-robot trial), the subjects continued interacting with the robots until every subject had finished.

### 4.2.5 Evaluation

There are many possible measures of performance success, and Steinfeld et al. summarized this aspect well regarding the case of non-task-oriented social interaction [16]. Here, we chose to focus on task-oriented social interactions from the user's perspective ("user" referring to the customer rather than the operator).

The subjects were asked to fill out a short questionnaire after every interaction, which measured *Interaction Success* (whether or not the robot had factually answered their question), and *User Satisfaction* (whether they felt any dissatisfaction with the interaction, measured on a 1-7 scale, with 7 representing no dissatisfaction at all, and 1 representing strong dissatisfaction).

## 5. RESULTS AND DISCUSSION

### 5.1 Results

#### 5.1.1 Experiment 1

In the first experiment, only reactive timing control was used, which made simultaneous operation of the robots challenging for the operator. In the 3R and 4R cases, there were conflicts between robots demanding the operator's attention simultaneously, and the operator relied more heavily on the autonomous system.

**Task difficulty:** The questions the subjects were to ask the robots had been selected for an approximate SC of 50%. The ASR (automatic speech recognition) system performed well in our quiet laboratory conditions, and the fully-autonomous trial showed a remarkable 42% success rate. The average CTR value for this experiment was 31.4%.

**Interaction success:** The success rate for the 1R and 2R cases was 100% (Figure 5). The 3R and 4R cases were 90% successful, with five failures each. Six of them were due to speech recognition errors, three to operator errors due to rushing, and one failure was due to the operator's misunderstanding of a question. A chi-square test revealed significant difference among conditions ( $\chi^2(4)=85.232$ ,  $p<.01$ ). The residual analysis revealed that 1R and 2R had a significantly smaller number of failures than A ( $p<.01$ ).

**Users' satisfaction:** The users' satisfaction ratings for the 1R – 4R conditions were all quite high (Figure 6), when compared with the fully-autonomous condition. According to subjects' comments, the slight decrease visible in the 3R and 4R conditions was due to both an increased failure rate and longer wait times. A repeated-measure ANOVA revealed a significant difference in the main effect of number of robots ( $F(4,232)=30.381$ ,  $p<.001$ ). A Bonferroni test revealed that 1R, 2R, 3R, and 4R were significantly better than A ( $p<.001$ ), 1R was significantly better than 4R ( $p<.05$ ), and 2R was marginally better than 4R ( $p<.1$ ).

**Operator supervision in critical section:** As the number of robots increased, the operator had less time available to assist each robot. Figure 7 shows the amount of time the operator spent monitoring each robot during critical section periods, expressed as a fraction of the total critical section duration for all robots in that condition. It can be seen that this value dropped quickly as additional robots were added, up to the 4R condition, in which over 45% of the critical section time across all robots was not monitored by the operator.

**Manual error correction:** There were a number of cases in which the robot, due to errors in speech recognition, began to guide the user to the wrong destination. In the 1R and 2R cases, the operator was able to correct all the errors, usually before the robot began giving directions. In the 3R case three errors occurred, of which the operator corrected one, and in the 4R case this increased to six errors, of which the operator corrected two.

### 5.1.2 Experiment 2

In Experiment 2, two conditions were changed. Proactive timing control was used, and a more realistic conversation flow was used for the interaction, increasing total interaction time and resulting in a lower CTR. Operator supervision in the critical section was high, and performance was improved in both interaction success and manual error correction.

**Task difficulty:** The SC was around 50%, as in Experiment 1. Due to the longer average interaction length, the CTR was much lower than in Experiment 1, with an average value of 17.5%.

**Interaction success:** The interaction success rate did not vary significantly between the 1R – 4R conditions. Success rates were all above 95%. There were only three failures, one in the 2R case and two in the 3R case. One of the 3R “failures” was due to the operator not knowing the answer to the question. The remaining two failures were caused by careless operator errors due to rushing. As in Experiment 1, there was a significant difference when compared with the autonomous case, which was successful only 35% of the time ( $\chi^2(4)=125.894$ ,  $p<.01$ , residual analysis: 1R, 2R, and 4R to A:  $p<.01$ , 3R to A:  $p<.05$ ).

**Users’ satisfaction:** User satisfaction did not vary significantly between the 1R – 4R conditions. A repeated-measures ANOVA revealed the significant difference in the main effect of number of robots ( $F(4,232)=41.357$ ,  $p<.001$ ). A Bonferroni test revealed that 1R, 2R, 3R, and 4R were significantly better than A ( $p<.001$ ).

**Operator supervision in critical section:** Due to the proactive timing control and lower CTR, the operator availability during critical sections was over 90% for every condition in this experiment (Figure 7). Although a slight decrease in availability is visible as the number of robots increases, the effect is not nearly as strong as in Experiment 1. We attribute the improved success rate in the 3R and 4R cases to this increase in operator availability.

**Manual error correction:** The robot started a total of 7 incorrect guiding behaviors due to erroneous speech recognition, but the operator was able to identify and correct all of them.

## 5.2 Discussion

We were, frankly speaking, quite surprised by the positive results of this experiment and the operator’s success in controlling four robots. We had initially expected three robots to be a challenge and four to be nearly impossible, even with proactive timing control. With reactive control, we expected that the operator would

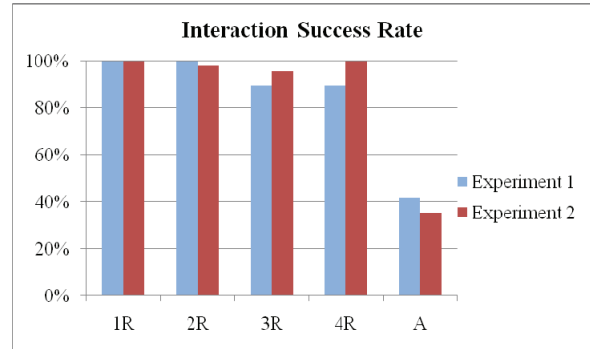


Figure 5: Interaction Success Rate

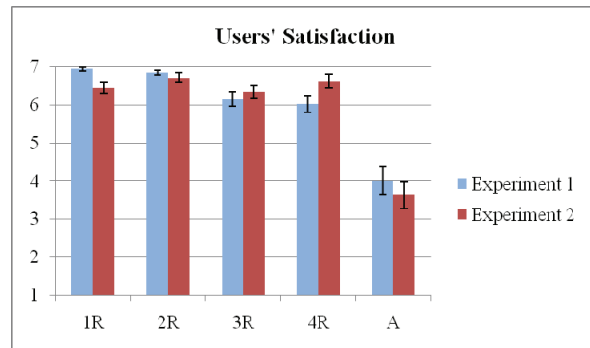


Figure 6: Users’ Satisfaction

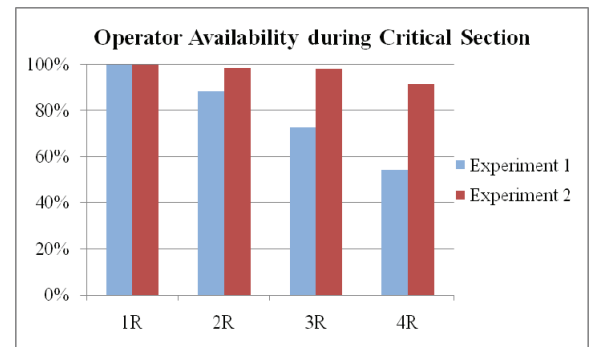


Figure 7: Operator Availability during Critical Section

only be able to control one robot. Thus, fortunately or unfortunately, we could not observe the upper boundary for the number of robots one operator can control in the full-interaction case. Figure 7, showing the operator availability during the critical section, seems to indicate that beyond four robots, the operator will only be available for half of the critical section time or less. In that case, autonomous system failures are likely to increase dramatically, adversely affecting Interaction Success.

### 5.2.1 Maximum number of robots

For Experiment 2, we consider the four-robot case to have been a success. For a challenging case like Experiment 1, two robots might be the maximum if there is a low tolerance for failure. In a real application, without a continuous stream of customers, the CTR could be even lower than in Experiment 2, increasing the maximum number of robots even further.

### 5.2.2 User Interface

Of the data displayed in the robot status panels, the type of information the operator found most useful was that which provided long-range situational perspective. The most useful displays turned out to be the Current Behavior display and the Interaction Timer, both of which provide information about the progress of the interaction and enable the operator to plan ahead.

An even higher-level interaction plan such as a flowchart might be even more useful to an operator, since the operator adds value to an interaction primarily through situational understanding.

### 5.2.3 Automatic Error Detection

Although automatic error detection has not yet been implemented in our system, it has great potential for improving the efficiency of a multi-robot system. Without such functionality, the operator is required to take the initiative in searching for errors, based on whatever information is available through the interface. Automatic error detection would enable the system to take the initiative instead, with the operator reacting to the system's requests to investigate suspected errors.

### 5.2.4 Degrees of Criticality

It might be useful to consider additional levels of criticality, or error risk. The Interaction Status indicator displays three levels of criticality: "Idle" is a very low level, since the robot is not interacting. "Active" is a slightly higher level of criticality, an indication that the operator should be aware of the robot's status and monitor it in the background. "Critical" is the highest level, and it demands the operator's attention immediately. Automatic error detection or other estimators of error risk could be used to create more refined and accurate levels of criticality.

## 6. CONCLUSIONS

In this study, we have successfully demonstrated a system in which a single operator is able to simultaneously control up to four humanoid robots engaging in social interactions. For an artificially difficult interaction, we showed that the operator can easily control two robots, and can control more with some performance degradation. For a more realistic interaction, and using a technique we call "proactive timing control" to reduce the number of conflicts between robots demanding the operator's attention, we were able to achieve a task success rate of over 95%, even in the four-robot case. In both cases, task success and user satisfaction in every condition were far superior to those attainable by the same system operating in a fully-autonomous mode. Most importantly, we have tested this system using an actual task often performed by our robots in the field, suggesting that this technology can be immediately put to use in real-world field trials. This study only marks the beginning of teleoperation for multiple social robots, and several of the areas addressed in this paper merit further in-depth research.

## 7. ACKNOWLEDGMENTS

We wish to thank the staff of the AEON Takano-hara Shopping Center for their kind cooperation. This research was supported by the Ministry of Internal Affairs and Communications of Japan.

## 8. REFERENCES

- [1] W. Burgard et al.: "The interactive museum tour-guide robot," *National Conference on Artificial Intelligence*, pp. 11-18, 1998.

- [2] J. W. Crandall, M. L. Cummings: Developing Performance Metrics for the Supervisory Control of Multiple Robots, *ACM/IEEE 2nd Annual Conference on Human-Robot Interaction (HRI2007)*, pp. 33-40, 2007.
- [3] J. W. Crandall and M. A. Goodrich: "Characterizing efficiency of human robot interaction: a case study of shared-control teleoperation," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2002)*, pp. 1290-1295, 2002.
- [4] T. Fong and C. Thorpe: "Vehicle Teleoperation Interfaces," *Autonomous Robots*, Vol. 11, No. 1, pp. 9-18, 2001.
- [5] M. A. Goodrich, T. W. McLain, J. D. Anderson, J. Sun, J. W. Crandall, Managing Autonomy in Robot Teams: Observations from Four Experiments, *ACM/IEEE 2nd Annual Conference on Human-Robot Interaction (HRI2007)*, pp. 25-32, 2007.
- [6] A. Green et al.: "Applying the Wizard-of-Oz Framework to Co-operative Service Discovery and Configuration," *Proc. IEEE Int. Workshop on Robot and Human Interactive Communication*, 2004.
- [7] R. Gockley, J. Forlizzi, R. Simmons, Interactions with a Moody Robot, *ACM/IEEE 1st Annual Conference on Human-Robot Interaction (HRI2006)*, pp. 186-193, 2006.
- [8] K. Hayashi et al.: Humanoid robots as a passive-social medium - a field experiment at a train station, *ACM/IEEE 2nd Annual Conference on Human-Robot Interaction (HRI2007)*, pp. 137-144, 2007.
- [9] S. G. Hill, B. Bodt, A Field Experiment of Autonomous Mobility: Operator Workload for One and Two Robots, *ACM/IEEE 2nd Annual Conference on Human-Robot Interaction (HRI2007)*, pp. 169-176, 2007.
- [10] T. Kanda, T. Hirano, D. Eaton, H. Ishiguro, "Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial," *Human Computer Interaction*, Vol. 19, No. 1-2, pp. 61-84, 2004.
- [11] J. R. Movellan et al.: The RUBI Project: A Progress Report, *ACM/IEEE 2nd Annual Conference on Human-Robot Interaction (HRI2007)*, pp. 333-339, 2007.
- [12] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita: Interactive Humanoid Robots for a Science Museum, *IEEE Intelligent Systems*, vol. 22, no. 2, pp. 25-32, Mar/Apr, 2007.
- [13] B. P. Sellner et al.: "Attaining Situational Awareness for Sliding Autonomy," *ACM/IEEE 1st Annual Conference on Human-Robot Interaction (HRI2006)*, pp. 80-87, 2006.
- [14] N. E. Sian et al.: "Whole Body Teleoperation of a Humanoid Robot Integrating Operator's Intention and Robot's Autonomy," *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS2003)*, pp. 1651-1656, 2003.
- [15] R. Siegwart et al.: "Robox at Expo.02: A Large Scale Installation of Personal Robots". *Robotics and Autonomous Systems*, 42, 203-222, 2003.
- [16] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, M. Goodrich, Common Metrics for Human-Robot Interaction, *ACM/IEEE 1st Annual Conference on Human-Robot Interaction (HRI2006)*, pp. 33-40, 2006.
- [17] S. Woods et al.: "Comparing Human Robot Interaction Scenarios Using Live and Video Based Methods," Towards a Novel Methodological Approach, *Int. Workshop on Advanced Motion Control*, 2006.