

Field Trial for Simultaneous Teleoperation of Mobile Social Robots

Dylan F. Glas

Takayuki Kanda

Hiroshi Ishiguro

Norihiro Hagita

ATR Intelligent Robotics and Communication Laboratory
2-2-2 Hikaridai, Keihanna Science City,
Kyoto, Japan

{dylan, kanda, ishiguro, hagita}@atr.jp

ABSTRACT

Simultaneous teleoperation of mobile, social robots presents unique challenges, combining the real-time demands of conversation with the prioritized scheduling of navigational tasks. We have developed a system in which a single operator can effectively control four mobile robots performing both conversation and navigation. We compare the teleoperation requirements for mobile, social robots with those of traditional robot systems, and we identify metrics for evaluating task difficulty and operator performance for teleoperation of mobile social robots. As a proof of concept, we present an integrated priority model combining real-time conversational demands and non-real-time navigational demands for operator attention, and in a pioneering study, we apply the model and metrics in a demonstration of our multi-robot system through real-world field trials in a shopping arcade.

Categories and Subject Descriptors

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

General Terms

Design, Human Factors.

Keywords

Simultaneous teleoperation, Adjustable autonomy, Prioritized control, Teleoperation of social robots

1. INTRODUCTION

The day is quickly approaching when we will see mobile robots around us in our everyday environments. Field trials have been conducted with robots in museums, expositions, schools, train stations, and other places [1, 12, 14, 22, 24]. Institutions such as hospitals have also begun to use mobile robots [17], and projects such as DustBot and URUS [19] have begun exploring the ways robots can operate alongside humans in urban environments.

The operation of robots within real social environments (Figure 1) is a significant divergence from traditional fields such as manufacturing, military, or space robotics, where robots operate in isolated or dangerous environments and are far removed from the world of everyday life. This change in role, operating environment, and relationship between robots and humans calls for a

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'09, March 11–13, 2009, La Jolla, California, USA.

Copyright 2009 ACM 978-1-60558-404-1/09/03...\$5.00.

fresh reassessment of the fundamental requirements and priorities for teleoperation of mobile, social robotic systems.

For traditional robotic systems, significant progress has been made in the formulization of concepts such as fan-out, interaction time, neglect time, and interaction efficiency [2, 18, 25] to model the number of robots that can be controlled simultaneously for tasks such as search and navigation.



Figure 1: Customers interacting with robots in our field trial

Yet, little progress has been made in remote operation for interactive social robots. The Wizard-of-Oz (WOZ) technique has been used in research [10, 27], but this has been based on the assumption that one operator controls one robot. As their recognition capabilities and autonomy increase, simultaneous teleoperation of multiple mobile, social robots is becoming a real possibility.

This is a field, however, in which we have only begun to break the ice. What are the requirements for these systems? What new roles will the operator play? How should our approach to multiple-robot teleoperation differ from traditional approaches?

Many technological hurdles lie between the state of the art today and a future where fully-autonomous robots play roles in society. Partial autonomy and teleoperation can help bridge this gap, both for research and for real applications. A human operator can take the place of technology which is not yet mature, transitioning over time from an active, controlling role to a passive, monitoring role as the robot's autonomy and recognition capabilities improve.

2. MOBILE SOCIAL ROBOTS

2.1 Unique Characteristics

The first question to address is, what are the key unique points we must consider in the teleoperation of mobile social robots?

Safety of People: The safety of people is perhaps the most important issue. This is, of course, a concern with many robot systems; however, the people who work with military or manufacturing

robots tend to be specialists with training. Social robots must be 100% safe when interacting with the general public.

Safety of Robots and Environment: The importance of individual robots and their environment must also be reconsidered. In some robot scenarios, such as for “dull, dirty, or dangerous” tasks, “number of robots lost” is used as a performance metric. However, in a social environment, loss of control of a robot could harm people or damage the environment in addition to losing the robot; for a “social” robot, such disruptive and dangerous situations must be avoided.

Predictability of Events: Conversational interactions tend to follow patterns making it sometimes possible to anticipate the need for an operator. This enables us to proactively schedule behaviors to avoid conflicts for the operator’s attention [6].

However, at the same time, navigational tasks must be performed, where the need for an operator tends to be based on stochastic failures and cannot be anticipated. Thus an operator must attend to a mix of predictable and unpredictable events.

Real-Time Responsiveness: Social interactions are conducted in real time, and sometimes an operator’s other tasks must be interrupted to maintain real-time responsiveness. Users become frustrated with robots which do not respond within a few seconds [23].

Although the metric of “neglect tolerance” [3,26], for example, is a useful tool for modeling robot downtime when delays in operator response are acceptable, such concepts cannot be easily applied to the social interaction domain due to its real-time nature.

2.2 Research Objectives

The objectives of this research can be summarized as follows:

1. To define the characteristics of and operator’s tasks for *mobile* social robots
2. To design a priority model among these tasks by considering real-time needs
3. To enable an operator to control multiple *mobile* social robots based on this priority model
4. To identify metrics for task difficulty and operator performance that capture the real-time aspect of *mobile* social robots
5. Finally, to assess the practical value of the task definition, priority model, and metrics in a field trial, and retrieve the parameters needed for task difficulty metrics to enable further mathematical modeling of *mobile* social robots

Above all, it is important that the knowledge be verified through *real-world testing* in field trials like [13]. This project is thus a pioneering effort for the purpose of establishing a new model of requirements and priorities for mobile, social robots.

3. DESIGN CONSIDERATIONS

3.1 Requirements for Autonomy

3.1.1 Error Detection

Because the system sometimes operates autonomously, it must be able to rely on automatic error detection to ensure safety. The need for this is twofold: first, the system must default to a safe state in case an operator is not immediately available, and second, the system should communicate the error to the operator, who should correct the error as soon as possible.

Previous studies in collaborative control have explored such cooperation between an operator and an automatic error detection

system [4, 5]. Kaupp *et al* have also recently done research in this direction [15], using robot-initiated notifications of error states to reduce teleoperation workload for navigational tasks.

A point essential to the discussion in this paper is that we have implemented in our robot system the ability to reliably detect critical errors, enabling the operator to focus on higher-level tasks.

3.1.2 Safe Default Actions

Since the availability of an operator is not guaranteed, it is essential that in the absence of an operator, the robot’s default actions must be inherently safe; that is, the robot cannot simply report the error to an operator. For most navigational tasks, stopping locomotion when an error is detected is a safe default action, allowing the robot to wait safely for commands from the operator.

3.2 Operator Responsibilities

3.2.1 Assisting Recognition

For social robots, human interaction is essential to their function. So long as the technology for recognition of speech, emotion, and gesture in real-world situations is insufficient for smooth human-robot interaction, a primary function of the operator will be assisting recognition. This function must be performed in real time, or conversational interactions will fail because people will not wait for a robot that is not responding to them.

3.2.2 Correcting Errors

Another responsibility of the operator is correcting errors in navigation. If we can assume that error detection for critical situations is reliable, and that the robot can put itself into a safe state autonomously, this is not a time-critical task. Error states should be corrected sooner rather than later, but the cost of a delayed response is simply lower efficiency, so error correction carries a lower priority than recognition assistance (ignoring which could potentially cause an interaction failure).

3.2.3 Supervisory Monitoring

A third type of task is supervisory monitoring of the robots to detect other kinds of errors. Although we assume error detection is reliable for critical errors, there is always the potential for non-critical errors, for example a robot mistakenly initiating a conversation with a door. When there are no higher-priority conditions to concentrate on, it is useful for the operator to monitor the robots in a supervisory role to ensure smooth operation.

3.2.4 Attention Management

One of the problems inherent in this definition of operator responsibilities is that the task of continuous monitoring always conflicts with the other tasks.

This was confirmed by our observations in early trials, where we observed that the operator was often too absorbed in one specific task to be able to monitor other robots in need of high-priority assistance. In our implementation we propose the use of an autonomous switching mechanism to help in resolving this conflict. The concept of attention management aids has also been investigated for navigational tasks by Goodrich *et al*. [9].

3.3 Priority Model

Each robot has several possible conversational and error states which demand an operator’s attention. These states vary in both severity and immediacy, and they need to be prioritized for consistent and effective operation of the robot. The states considered

Table 1: Robot states and their priorities

State	Severity	Immediacy	Priority	Consequences if operator is absent
Navigation-Normal	Low	Low	0	Failures unlikely
Conversation-Normal	Low	Low	1	Operator could miss an out-of-turn utterance
Navigation-Stopped	Medium	Low	2	Robot stops and cannot perform services
Conversation-Pre-critical	Medium	Medium	3	Operator could misunderstand conversational context
Navigation-Unsafe	High	Medium	4	Robot cannot perform services; could be mistaken for human by other robots; could be re-associated incorrectly
Conversation-Critical	High	High	5	Robot cannot respond to the customer; interaction failure
Hardware-Critical	Highest	High	6	System could crash; potential damage to robot

here are summarized in Table 1 and represented graphically in Figure 2, and the meaning of each state is defined below.

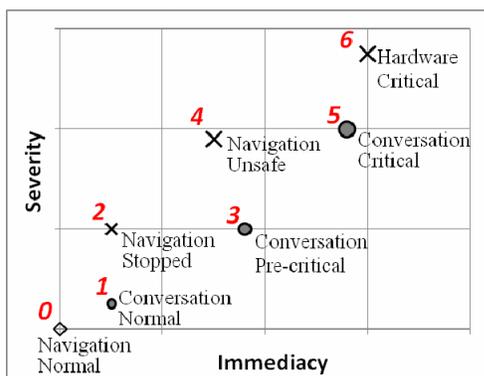


Figure 2: Priorities of robot states

Navigation-Normal is the case where the robot is patrolling with no error conditions, and it is the lowest priority, as it is non-interactive and thus does not require real-time monitoring.

Conversation-Normal is next in priority; if the operator has time, it is more valuable to monitor conversation than navigation, but the need for recognition support is not anticipated.

The next priority level is *Navigation-Stopped*. This state can indicate any of several errors, such as a persistent obstruction or failure of a sensor, but all with the same consequence – the robot has stopped and cannot approach people, and thus efficiency is lost.

The next higher-priority condition is what we call *Conversation-Pre-critical*, when the need for an operator is anticipated within the next 5-10 seconds. Listening to the conversation at this point can help establish situational awareness [21] to help the operator with recognition when the critical section does begin.

Tracking failures (described in Section 4.1) are categorized as *Navigation-Unsafe*, because a tracking failure can have more consequences than simply stopping a robot. Other robots might mistakenly identify the lost robot as a human and attempt conversation with it, or the system may re-associate the robot to the wrong position, for example to a human who is standing still. This could cause problems and confusion later, so fixing this error is moderately high priority.

Conversation-Critical is even higher priority, as it requires the operator’s recognition support at the risk of conversation failure.

Finally, a *Hardware-Critical* warning is the most serious condition, as neglecting to respond to this situation could cause damage to the robot. For example, a low-battery condition could cause a system crash, with the risk of hard drive damage.

3.4 Metrics for Task Difficulty

We previously proposed two metrics for measuring the difficulty of *conversational* tasks for social robots [6], and to this we add a metric for characterizing *navigational* task difficulty.

Situation Coverage (SC): This measure of conversational task difficulty describes the fraction of conversational situations for which the robot would autonomously execute the correct behaviors given perfect sensor recognition. For “covered” situations, the operator acts like a simple sensor, but “non-covered” situations the operator needs to think to resolve the situation.

Critical Time Ratio (CTR): Another conversational task difficulty metric, this is the ratio of the duration of time spent in critical sections of a conversation, to the total conversation time.

Mean Time Between Failures (MTBF): This metric characterizes task difficulty for navigational tasks. Fong et al. [4] proposed the term “Mean Time Between Interventions”, describing the “mean time a human-robot system operates nominally (i.e. outside of intervention)”. Here, we will use MTBF, describing the average duration of a non-error state, to characterize the system reliability independent of operator intervention.

4. IMPLEMENTATION

The system we implemented can be represented as a three-layer architecture, consisting of (1) an environmental sensor network for robot localization and human tracking, (2) the four robots, and (3) the remote operation console, as depicted in Figure 3.

4.1 Infrastructure

Environmental Sensor Network

To track human and robot positions, we used a sensor network embedded in the environment. Although the task of people-tracking has been achieved using purely on-board sensors [20], the use of an environmental sensor network has many benefits.

First, people can be tracked robustly over a wide area, regardless of the positions of the robots or occlusions of their sensors. People often crowd around the robots, so this is a serious concern.

Second, a global reference can be established for coordinate systems (so robots are localized relative to each other) as well as a globally consistent set of human ID’s (so that robots can identify which people have already interacted with other robots).

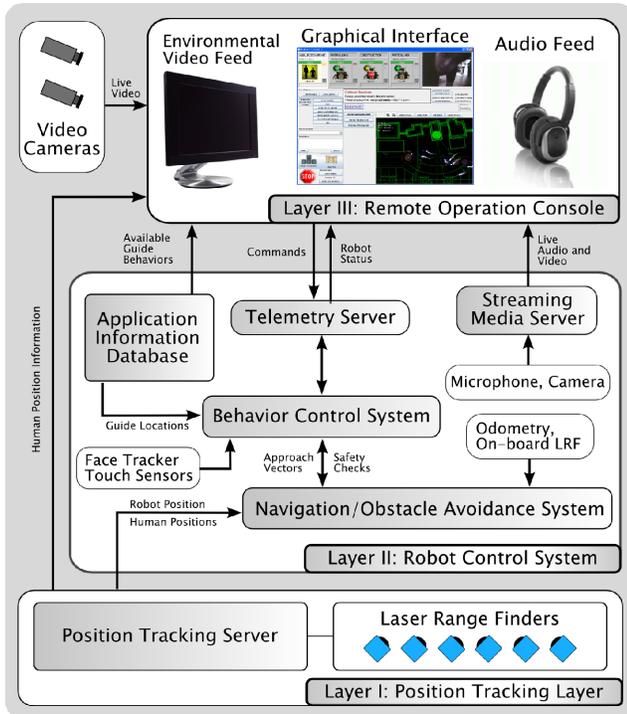


Figure 3: System architecture overview

Finding people

To track human and robot positions, we used a network of six SICK LMS-200 laser range finders, positioned around the perimeter of the area. They were set to a detection range of 80m with precision of 1cm, each scanning an angular area of 180° at a resolution of 0.5°, providing readings of 361 data points every 26ms.

The laser range finders were mounted 85cm from the ground, a height chosen so the sensors could see above ground clutter and obstacles such as benches and luggage. Also, at long range, the scan beams are spaced quite far apart (over 8cm at a distance of 10m) and detection of small features like legs is difficult. Detection of larger targets, like a torso, is more robust at these distances.

For detection and tracking of people, a technique based on the algorithm described in [7] was used. Particle filters are used for estimating positions and velocities for each person, and a contour-analysis technique is used to estimate the direction in which a person is facing. This orientation angle can be useful to the robot in determining likely approach candidates.

Robot localization

The people-tracking system described above is also used for robot localization. Global localization is not currently performed by this system – the operator assigns each robot’s initial position. The localization system then assumes that the robot is one of the “people” being tracked, and it associates the robot with the nearest observed human position, within a threshold distance.

The particle filter being used to track that human is then replaced with a new particle filter, incorporating the robot’s reported wheel velocities in its motion model and using different shape parameters for its likelihood model. When this robot tracking model can no longer be associated with any detected features, the model is removed and a “tracking failure” occurs.

This position data is used to correct the robot’s estimated x-y position. It cannot, however, be directly used to correct the robot’s orientation. Due to shape and reflectivity issues, the robot’s orientation is difficult to track using the human-tracking algorithm, so orientation corrections are instead performed by using a Kalman filter to estimate angular deviations between the robot’s observed trajectory and an expected trajectory based on its odometry data.

4.2 Robots’ autonomy

Finding obstacles

To ensure safety during autonomous locomotion, each robot has a triple-redundant collision avoidance system. The robot uses information from the human tracking system to determine the distance to the nearest person. Each robot also has a Hokuyo URG laser range finder mounted on its base, able to detect obstacles at a height of 12cm, up to a range of over 3m from the robot. When either of these two systems detects an obstacle or human within a threshold distance of 80cm from the front of the robot, locomotion is automatically stopped.

Finally, as an emergency backup system, the robot has bumper sensors mounted around its base which cause the robot to stop locomotion immediately when triggered. In our field trial, this backup system was never actually triggered.

Note that most obstacles are people, who move quickly. However, if the obstacle persists for longer than two seconds, it is likely to be a stopped robot, an unexpected object such as a piece of luggage, or a fixed part of the environment. In such situations, the robot turns and attempts to move to a new point far from the other robots. After three more seconds, if the obstruction still persists and the robot is unable to move away from it, the robot enters an “Obstruction” state and waits for operator assistance.

Conversation

The robot’s conversational interactions generally consist of offering guidance, recommending shops, explaining directions, and general greeting and chatting behaviors. Each behavior combines speech with gestures and motions. The transitions between these behaviors may be fixed, random, based on sensor data, or dependent upon speech recognition data provided by the operator.

Since the presence of an operator is not guaranteed, each critical conversation behavior has a timeout threshold, after which the robot transitions to another behavior. This allows the robot to continue interacting both when an operator is unavailable and when customers do not respond to its questions.

In an example interaction, the robot might approach a couple, greet them, introduce itself, and offer to shake hands. Once they shake hands, the robot notifies the operator that it will need speech recognition assistance and begins telling the customers about an interesting event that evening. The customers say nothing to this, so after the timeout limit is reached, the robot asks them if they are thirsty, because it is hot outside. The customers say no, so the robot asks if they would like guidance to anywhere in particular. The customers ask where they can buy souvenirs. The operator hears this request and clicks on the speech recognition button marked “souvenirs,” and the robot randomly selects a souvenir shop and explains how to get there. The customers then thank the robot, which says goodbye and returns to patrol mode.

4.3 Teleoperation interface

The teleoperation software used by the operator was based on the operation console used in our previous experiment [6]. Its basic structure, shown in Figure 4, includes a row of panels at the top, which display critical information about each robot, and which the operator can click to switch to that robot. The icons on these panels reflect the robot’s current state.

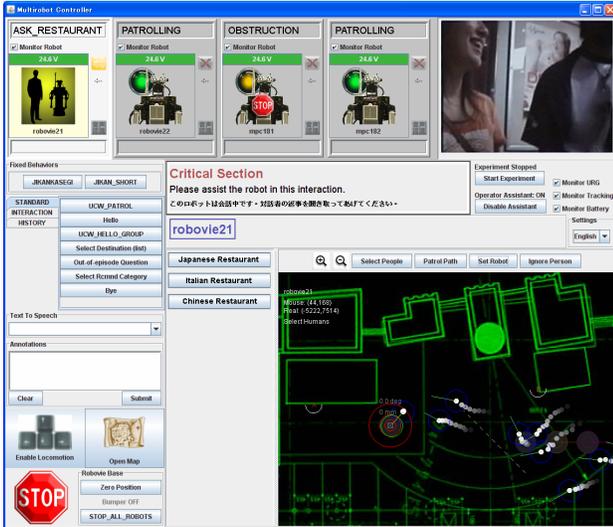


Figure 4: Teleoperation console

The video display in the upper right shows the view from the selected robot’s camera. The lower section of the screen contains controls for sending commands to the currently-selected robot and correcting recognition errors. There is also a map showing the robot’s position, orientation, current approach target, current patrol path, and the positions of humans and robots nearby. Controls on this map can be used to correct errors in the robot’s position and orientation.

A guidance panel is also provided in the center of the screen. This panel gives instructions on how to respond appropriately to the robot’s current condition. For example, if the robot is in the critical section of a conversation, this panel explains that the operator should listen to the customer’s response to the robot’s question and click the appropriate button below. If the robot is in an obstructed state, this panel explains how to manually maneuver the robot until it is again free to move.

4.4 Managing Operator’s Attention

As noted in Section 3.2.4, the operator’s tasks of recognition support and error correction conflict with the continuous task of system monitoring. Deciding which robot to devote attention to is a task in and of itself, which the operator may neglect when busy.

Thus we developed an autonomous switching system which continuously monitors all of the robots and automatically switches the operator’s view to the robot which is highest-priority for the operator’s attention. This enables the operator to ignore other robots and focus solely on helping the current robot.

4.4.1 Error Detection

Reliable error detection is essential in order for an automatic switching system to be trusted with the responsibility to control an operator’s attention. The system detects four error states.

Low Battery: The robot continually monitors its battery charge at a resolution of 0.1 V, sending that data to the teleoperation console every time it changes.

Sensor Failure: If the time elapsed since the last update from the robot’s onboard laser range finder exceeds a threshold of 300 ms, an error alert is sent to the teleoperation console.

Tracking Failure: A timeout warning is sent when location data is not received from the localization system, updated every 27 ms.

Obstruction: Information about the distance and direction of obstacles from the robot is sent to the teleoperation console continuously; however, the “Obstruction” state is only registered after the robot has attempted to escape the obstructed state on its own and failed, transitioning to a *Navigation-Stopped* status.

4.4.2 Switching Algorithm

The switching algorithm follows the state priority order shown in Table 1, assigning the operator to the robot with highest priority, constrained by the following three rules.

First, established operator-robot assignments are only interrupted when another robot is in a state of higher (not equal) criticality. This prevents the system from switching repeatedly between two robots in the same state.

Second, if multiple robots have equal criticality, the robot which has been neglected for the longest time is chosen.

Third, switching is temporarily disabled while the operator is engaged in an active task, inferred from mouse motion. This prevents the system from switching robots just as an operator is about to send a command, causing the command to be sent to the wrong robot. Three tasks were defined to be uninterruptible: behavior selection, position correction, and direct locomotion control.

5. FIELD TRIAL

5.1 Scenario

Environment: To demonstrate our system’s performance in a real-world setting, we conducted a field trial in a shopping arcade. We operated four Robovie-II humanoid communication robots in an area in front of clothing and accessory shops. The area had light pedestrian traffic but was not too crowded, giving the robots many opportunities both to interact with people and to use their navigational ability to approach people from a distance.

The field trials took place in a corridor roughly 5m wide and over 20m long. Laser range finders were placed around the perimeter to minimize occlusions, as shown in Figure 5.

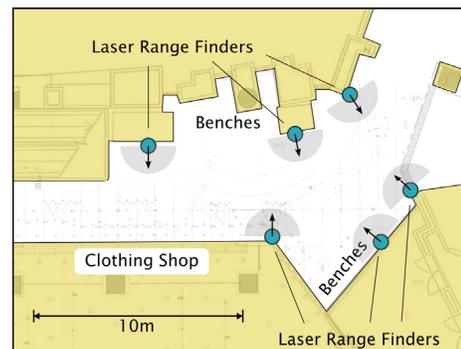


Figure 5: Map of the field trial environment

Robot Task: The task assigned to the robots included both navigation and conversation. Robots patrolled the area, looking for someone to approach. When a robot successfully approached a customer, it would greet them, recommend shops or restaurants, offer directions, and chat about random topics. After the customer left, the robot would resume patrolling to look for a new customer.

Operator Task: During each trial, the operator assisted the robots with speech recognition support during conversation by clicking on buttons that corresponded to the words customers spoke. The operator also assisted with obstacle avoidance, by manually driving the robots away from obstacles. Finally, the operator corrected tracking errors by driving the robots around to determine their positions, and then clicking on the correct position on the map.

The video included with this paper was recorded during this field trial and shows examples of people interacting with the robots.

5.2 Procedure

5.2.1 Conditions

In our field trials, we compared the performance of a single operator controlling four mobile social robots under two conditions.

Manual Switching Mode: In the first condition, the operator actively monitored all robots, manually switching between them to correct error states and perform manual speech recognition.

Automatic Switching Mode: In the second condition, the system automatically switched the operator’s focus to the robot with highest-criticality status. The states of the other robots were hidden, to help the operator focus on one robot at a time.

In both conditions, operators were instructed to try to keep all four robots in active operation (i.e. not lost, stuck, or inactive) as much of the time as possible, but to give priority to critical conversational interactions over navigational corrections.

Two staff members helped to protect the robots from abusive children and monitor the sensor network, as events such as a customer moving an LRF out of alignment can cause misdetections.

5.2.2 Participants

Three participants were used for these field trials; all were male university students working as interns in our lab. None were involved with this or other teleoperation research.

The participants were given 90 minutes of orientation, introducing the robot’s capabilities, the localization system, and the various displays and functions of the teleoperation console. This was followed by 30 minutes of practice time each, operating first one robot, then four, in both manual and automatic switching mode.

5.2.3 Trials

Each participant then operated the robots in real interactions with customers for three 10-minute trials under each condition (manual vs. automatic switching). Trials were counterbalanced, alternating between manual and automatic switching modes to avoid time-of-day bias.

Results from six trials were discarded due to technical trouble – two due to problems with the tracking server, and four due to an error in the switching mechanism. In these cases, we continued to run trials until each participant had completed three valid trials under each condition.

5.3 Metrics

5.3.1 Metrics for task difficulty

The metrics SC, CTR, and MTBF, described in Section 3.4, were used in the analysis, as were two additional metrics related to the amount of conversational activity during each trial:

Number of people: If more people encounter the robots, simultaneous conversations are more likely to occur, increasing task difficulty. Localization errors are also more difficult to correct, as detected people can be confused with robots.

Number of critical sections: This indicates the degree to which the robots interacted with people. Long interactions can contain many critical sections, so the number of critical sections reflects task difficulty more accurately than the number of interactions.

5.3.2 Metrics for objective performance

Operator presence during critical sections: This is a measure of the operator’s ability to manage multiple robots in conversation. An operator is considered “present” if available at any time during the critical section. This value should be as high as possible.

Downtime: This reflects the success of the operator in managing the robot’s navigational tasks. Mau *et al.* defined downtime as “the time during which the robot is waiting for the human to finish [correcting its error state]” [16]. Downtime was calculated as the sum of *Navigation-Stopped* and *Navigation-Unsafe* time, expressed as a percentage of the total operational time of the robot.

5.3.3 Metrics for workload

To evaluate the operator’s workload, we used the NASA-TLX scale (Task Load Index) [11], a tool for assessing subjective workload based on six factors: mental demand, physical demand, temporal demand, operator performance, frustration, and effort. Participants filled out TLX evaluation forms after every trial.

5.4 Results

5.4.1 Task Performance

Regarding operator presence during critical sections (Figure 6), a two-way repeated-measures Analysis of Variance (ANOVA) was conducted, where there were two within-participant factors, *switching mode* and *trial-repetition*. A significant main effect was revealed in the *switching mode* ($F(1,2)=33.964, p<.05$), whereas no significance was found in *trial-repetition* ($F(2,4)=.478, p=.651$), nor in interaction within these factors ($F(2,4)=.012, p=.988$). Thus, autonomous switching *did* improve the task performance in terms of operator presence during critical sections.

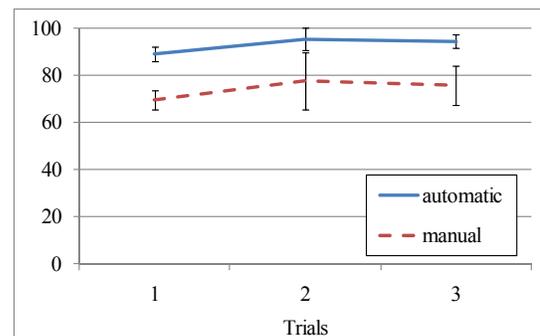


Figure 6: % of critical sections with an operator present

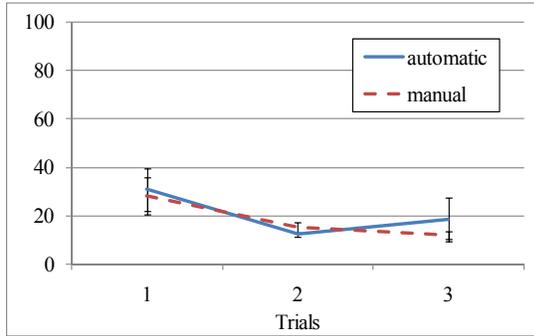


Figure 7: % Downtime

For the result of “% of downtime” (Figure 7), a two-way repeated-measures ANOVA was conducted, where there were two within-participant factors, *switching mode* and *trial-repetition*. There was no significant main effect in *switching mode* or *trial-repetition*, or in the interaction of those two factors. ($F(1,2)=.199, p=.699, F(2,4)=2.333, p=.213, F(2,4)=.712, p=.544$). Thus, we cannot say whether switching mode affects task performance for navigational tasks. At least, this data gives no indication that automatic switching sacrifices navigational performance.

5.4.2 Task Load

Regarding the NASA-TLX workload index (Figure 8), a two-way repeated-measures ANOVA was conducted, where there were two within-participant factors, *switching mode* and *trial-repetition*. A significant main effect was revealed in the *trial-repetition* ($F(2,4)=7.478, p<.05$), whereas no significance was found in the *switching mode* ($F(1,2)=.435, p=.578$) or interaction within these factors ($F(2,4)=1.030, p=.436$). Multiple comparison with the Bonferroni method was conducted for *trial-repetition*, which suggests an almost-significant difference between trial 1 and trial 2 ($p=.084$). This suggests a possible training effect.

5.4.3 Task Difficulty

Task difficulty in both conditions was similar, as measured by number of people, number of critical sections, MTBF, SC, and CTR. A comparison between these factors is shown in Table 2.

Note that SC was assumed to be 100% for all trials. People tended not to ask complex questions in this environment, and we did not see any instance where the operator needed to override the behavior flow, for example to manually enter new utterances.

For each task difficulty metric, a two-way repeated-measures ANOVA was conducted, where there were two within-participant factors, *switching mode* and *trial-repetition*. There was no significant main effect in *switching mode* or *trial-repetition*, or in interaction of these two factors.

Table 2: Task difficulty

(Numbers in parentheses represent standard error)

	automatic	manual
Number of people (per 10 min.)	166.4 (28.6)	154.6 (16.6)
Number of critical sections	20.8 (3.9)	20.6 (4.0)
Mean Time Between Failures (s)	68.1 (16.4)	59.9 (9.7)
Situation Coverage	100%	100%
Critical Time Ratio (%)	19.7 (2.4)	21.7 (2.5)

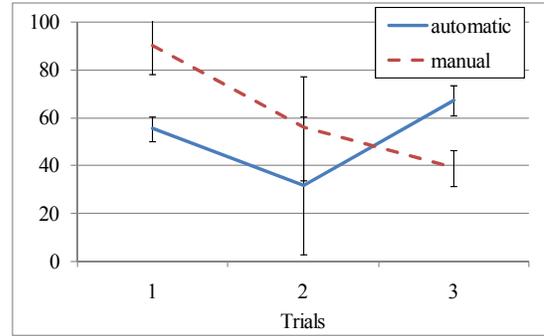


Figure 8: NASA-TLX Scores

5.4.4 Observations

How robots interacted with people

Roughly 60%-80% of the people responded to the robot when it initiated conversation with them. Usually interactions lasted for a few exchanges of turns. When the robot asked, “Would you like more information about this shopping area?” people’s responses included, “dinner”, “we are going to eat”, “hello”, “can you speak?”, “no thanks”. Sometimes, they did not respond to the robot, but talked to others saying things like “it’s talking!”, “let’s take a picture”, or “I’m curious how it works”.

How operators interacted with robots

Operators in the autonomous switching mode were generally calm and could successfully handle the requested tasks, although sometimes they seemed frustrated when the interface switched them to another robot without warning. In the beginning, the operators strongly preferred the automatic switching mode, with one of them actually begging us, “please let me use autonomous mode!”

In manual mode, the operators tended to do well in low-conflict situations, but faced with conflicting tasks, they began making mistakes, such as becoming so absorbed in the task of localizing a robot that they fail to notice that other robots were in conversation with customers and needed help. At the beginning, an operator commented, “it was difficult to control four robots at the same time. I did not understand what was going on.”

However, these differences seemed to become smaller after they got accustomed to operating the robots. One operator commented that “manually switching between the robots came to be enjoyable, even though it kept me busy; I learned to operate the system, and it became easy to operate the robots in manual mode as well.”

6. DISCUSSION

6.1 Limitations

As this study represented a highly challenging, unprecedented task, our focus was primarily on solving the main issues. Many issues remain to be explored and covered in more depth.

One limitation of this study was that the number of participants (3) was small. However, operating these robots is still a difficult task, and it is important for the operators to have some knowledge of robotics and an understanding of the system. The operator has the ability to override safety checks, and thus can seriously damage the system. It is currently difficult to run trials for a large number of participants, but in the future, we aim to make the system safer and more robust, enabling trials with more participants.

We also did not measure customer impressions of their interactions with the robot, which depends heavily on the design of the interactive behavior sequences, a topic beyond the scope of this paper. The robot's interactions with customers were also quite simple; responding to complex situations would take much more of the operator's time, an aspect not investigated in this study.

6.2 Conclusions and Future Work

We have presented a pioneering achievement in teleoperation of mobile social robots, a system in which a single operator has successfully controlled four mobile, social robots in both conversational and navigational tasks. We have achieved this by automating safety-critical tasks and providing reliable error detection. Additionally, we have proposed a prioritized task model for operator attention control, which was demonstrated to improve operator performance, compared with purely manual control.

The primary achievements of this demonstration were the definitions of key problems, parameterizations of task difficulty, and the presentation of early results based on this knowledge, positioning us to explore these issues further in future studies. In the future, we hope to further investigate and formalize the unique requirements for teleoperation of social robots, and to improve the operator's capacity for supporting advanced conversational tasks.

7. ACKNOWLEDGMENTS

This research was supported by the Ministry of Internal Affairs and Communications of Japan.

8. REFERENCES

- [1] Burgard, W. *et al.*: The interactive museum tour-guide robot, *National Conference on Artificial Intelligence*, pp. 11-18, 1998.
- [2] Crandall, J.W. and Cummings, M. L., Developing Performance Metrics for the Supervisory Control of Multiple Robots, *Human-Robot Interaction (HRI2007)*, pp. 33-40, 2007.
- [3] Crandall, J.W., Goodrich, M.A., Olsen, D.R., Jr., Nielsen, C.W., Validating human-robot interaction schemes in multi-tasking environments, *IEEE Trans. on Systems, Man and Cybernetics*, Part A, 35 (4), pp. 438- 449, 2005
- [4] Fong, T. *et al.*: A Preliminary Study of Peer-to-Peer Human-Robot Interaction, *Int. Conf. on Systems, Man, and Cybernetics*, 2006
- [5] Fong, T., Thorpe, C., and Baur, C., Robot as partner: vehicle teleoperation with collaborative control, *In Proceedings from the 2002 NRL Workshop on Multi-Robot Systems*, 2002.
- [6] Glas, D., Kanda, T., Ishiguro, H., and Hagita, N., Simultaneous Teleoperation of Multiple Social Robots, *Human-Robot Interaction (HRI2008)*, pp.311-318, 2008.
- [7] Glas, D., *et al.*, Laser Tracking of Human Body Motion Using Adaptive Shape Modeling, *Int. Conf. on Intelligent Robots and Systems (IROS 2007)*, pp. 602-608, 2007.
- [8] Goodrich, M.A., and Schultz, A.C., Human-Robot Interaction: A Survey, *Foundations and Trends in Human-Computer Interaction*, Vol. 1, No. 3 (2007) 203-275
- [9] Goodrich, M. A., McLain, T. W., Anderson, J. D., Sun, J., Crandall, J.W., Managing Autonomy in Robot Teams: Observations from Four Experiments, *Human-Robot Interaction (HRI2007)*, pp. 25-32, 2007.
- [10] Green, A. *et al.*: Applying the Wizard-of-Oz Framework to Cooperative Service Discovery and Configuration, *Int. Workshop on Robot and Human Interactive Communication*, 2004.
- [11] Hart, S. G., Staveland, L. E., Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, In *Human Mental Workload*, P. Hancock, N. Meshkati (Eds.) pp. 139 – 183, 1988.
- [12] Hayashi, K. *et al.*: Humanoid robots as a passive-social medium - a field experiment at a train station, *Human-Robot Interaction (HRI2007)*, pp. 137-144, 2007.
- [13] Hill, S. G. and Bodt, B., A Field Experiment of Autonomous Mobility: Operator Workload for One and Two Robots, *Human-Robot Interaction (HRI2007)*, pp. 169-176, 2007.
- [14] Kanda, T., *et al.*, "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.
- [15] Kaupp, T. and Makarenko, A., Decision-Theoretic Human-Robot Communication, *Human-Robot Interaction (HRI2008)*, pp. 89-96
- [16] Mau, S., Dolan, J., Scheduling for Humans in Multirobot Supervisory Control, *Int. Conf. on Intelligent Robots and Systems (IROS 2007)*, pp.1637-1643, 2007.
- [17] Mutlu, B. *et al.*, Robots in Organizations: The Role of Workflow, Social, and Environmental Factors in Human-Robot Interaction, *HRI2008*, pp. 287-294, 2008.
- [18] Olsen, D.R., Jr., & Goodrich, M. A., Metrics for Evaluating Human-Robot Interactions. *PERMIS 2003*, September 2003.
- [19] Sanfeliu A. and Andrade J., Ubiquitous networking robotics in urban settings, in Proc. *IEEE/RSJ IROS Workshop on Network Robot Systems*, pp. 14-18, 2006.
- [20] Schulz, D., Burgard, W., Fox, D., and Cremers, A., People Tracking with Mobile Robots Using Sample-Based Joint Probabilistic Data Association Filters, *International Journal of Robotics Research*, Vol. 22, No. 2, 99-116 (2003)
- [21] Sellner, B. P. *et al.*: Attaining Situational Awareness for Sliding Autonomy, *Human-Robot Interaction (HRI2006)*, pp. 80-87, 2006.
- [22] Shiomi, M., *et al.*, Interactive Humanoid Robots for a Science Museum, *Intelligent Systems*, vol. 22, no. 2, pp. 25-32, 2007.
- [23] Shiwa, T., *et al.*, How Quickly Should Communication Robots Respond?, *Human-Robot Interaction (HRI2008)*, pp.153-160, 2008.
- [24] Siegwart, R. *et al.*: Robox at Expo.02: A Large Scale Installation of Personal Robots. *Robotics and Autonomous Systems*, 42, 203-222, 2003.
- [25] Steinfeld A., *et al.*: Common Metrics for Human-Robot Interaction, *Human-Robot Interaction (HRI2006)*, pp. 33-40, 2006.
- [26] Wang, J., Wang, H., Lewis, M., Assessing Cooperation in Human Control of Heterogeneous Robots, *Human-Robot Interaction (HRI2008)*, pp. 9-15, 2008
- [27] Woods, S. *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