

A Teleoperation Approach for Mobile Social Robots Incorporating Automatic Gaze Control and 3D Spatial Visualization

Andres Mora, Dylan F. Glas, Takayuki Kanda, and Norihiro Hagita

Abstract— The teleoperation of mobile social robots requires operators to understand facial gestures and other non-verbal communication from a person interacting with the robot. It is also critical for the operator to comprehend the surrounding environment, in order to facilitate both navigation and human-robot interaction. Allowing the operator to control the robot's gaze direction can help the operator observe a person's non-verbal communication; however, manually actuating gaze increases the operator's workload and conflicts with the use of the robot's camera for navigation. To address these problems, the authors developed a teleoperation system which combines automatic control of the robot's gaze and a 3D graphical representation of the surrounding environment, such as location of items and configuration of a shop. A study where a robot plays the role of a shopkeeper was conducted to validate the effectiveness of the proposed gaze control technique and control interface. It was demonstrated that the combination of automatic gaze control and representations of spatial relationships improved the quality of the robot's interaction with the customer.

Index Terms— spatial relationships, workload, teleoperation, human-robot interaction, partial autonomy, social mobile robot.

I. INTRODUCTION

MOBILE social robots are expected to be used in everyday environments such as shopping malls, elderly care centers, museums, etc. While it would be ideal for such robots to be deployed with full autonomy, we have found that a small amount of teleoperation enables such social robots to provide useful services even before key technologies, such as the capability of understanding speech in noisy environments, are technically feasible [1],[2]. There are also safety and legal reasons that would require partial supervision from human operators. Teleoperation is also actively used in laboratory studies, using Wizard-of-Oz (WOZ) methods [3]-[5], and teleoperation is also a mechanism which could be used for teaching robots to perform social tasks [6].

The development of social robots which incorporate teleoperation brings together two very different branches of human-robot interaction (HRI). One branch is social HRI, which focuses on studying psychological aspects of conversational interactions between people and robots. In this

study, such interactions are examined in a shopping scenario, so we use the term "customer" to refer to a person who engages in social interaction with a robot. This term has some similarity with Scholtz's role of "peer" [7] in that it represents the human side in a face-to-face interaction with a robot. But it differs from the "peer" or "teammate" role in that the human and robot are not collaborating to achieve a single goal; rather, the human's role is as a service receiver in the interaction, in contrast to the robot's role as a service provider. The other branch is HRI for teleoperation, which typically focuses on issues like the workload of the *operator* (person remotely controlling the robot), situation awareness, and shared autonomy [8] for the remote operation of non-social robots.

Little research has explored different techniques for teleoperating mobile social robots, leaving many questions unanswered. What new requirements exist for social robots? What new techniques can aid a teleoperator in controlling social robots effectively?

Keeping track of a person's face is fundamental for social interactions, as it provides the operator with awareness of the customer's state, including facial expressions and gestures. Yet, manually actuating this task requires a large amount of effort by an operator. An automatic gaze control technique was implemented to keep the customer within the robot's field of view and relieve the operator from this routine task.

For *mobile* social robots, however, navigation is another important concern. Teleoperation for mobile robots is often conducted based on a video feed from the robot's camera. However, this use conflicts with the proposed automatic gaze control technique.

Indeed, the field of view of most cameras is narrow and at any one time, the video can be showing the customer or the environment (e.g. the area in front of the robot) but not both. Thus, if the proposed automatic gaze control is always engaged, the operator cannot see video of the area in front of the robot. This effectively limits the operator's understanding of the robot's position and surroundings.

To overcome this difficulty, a 3D graphical user interface (GUI) was created to represent the robot's environment which augments the operator's understanding of spatial relationships. In this paper, we establish that a teleoperation system for mobile social robots must provide the operator with an appropriate representation of spatial relationships when automatic gaze control is used.

A. Mora, D. Glas, T. Kanda and N. Hagita are with the Advanced Telecommunications Research Institute International 2-2-2 Hikaridai, Keihanna Science City, Kyoto, Japan, 619-0288, e-mail: amoravar@asu.edu

II. RELATED WORKS

A. Teleoperation for navigation tasks

For mobile robots that have to accomplish navigation tasks in order to carry out missions such as search and rescue, military tasks or space exploration, there are two opposite approaches along the ends of a spectrum: being completely teleoperated by humans [9]-[11] or being fully automated [12], [13]. Some of the aspects of research on teleoperation involve increasing and maintaining the level of situational awareness of the operator [14], [15], combining mixed and virtual reality techniques to help the operator improve the navigation of the robot [16], and the design of the Graphical User Interface (GUI) to be used to remotely operate the robot.

Particular to the design of GUIs for navigational robots, a number of studies have been done regarding the way to present information [17], [18]. One notable finding could be summarized as the need to combine different types of information altogether [19], [20]. Specifically, such works have studied how the navigation of the robot improves with a GUI that integrates a video feed and map data within a 3D environment, in contrast to video-based only or map-based only GUI.

Although existing knowledge in this domain has proven useful, further understanding of the requirements governing the teleoperation of mobile social robots is imperative. The teleoperation of social robots requires observation of new kinds of information (e.g. gestures, facial expressions, tone of voice, relative positioning) [21] as well as to address new problems in actuation that may arise (controlling conversation, gaze direction, and gestures; following someone via locomotion or gaze control) [22]. Our approach to solve these issues is presented in later sections of this paper.

B. Teleoperation of social robots

In practice, the WOZ methodology in HRI involves the remote control of a robot system. In that respect, it appears to be similar to teleoperation. However, in such studies the teleoperation system itself is typically seen as a means to an end, and not as a research topic in itself.

In the work carried out by Kuzuoka [23], focus is given to the “ecology” among operators and customers. In Kuzuoka’s study the idea of the operator acquiring all the information through a video-only interface is investigated, and no map information is provided. It reports the fact that what the operator utilizes (in this case, a three-screen based GUI) is not necessarily a good factor for the interaction with a customer e.g. due to the robot’s lack of natural motion.

C. Natural interaction with social robots

In this study, our focus is to enable a “context-sensitive” interaction between a human and social robots, where the robots’ interactions go beyond simple question-and-answer or command-response interactions.

In the scope of this paper, the importance is the adaptability of the robot to the customer’s context, including location, surrounding objects, target of attention, and subtle reactions (see our watch shop scenario in Section V as an example of

such interaction). There have been a number of studies with social robots conducted for natural interaction. There are many aspects to be studied, such as knowledge of non-verbal behaviors, like natural way of gazing [16], [19], proximity behavior [14], [15], the way of social dialog [4], and social patterns [25]. These studies are certainly useful for future social robots; however, the context of users was often out of focus in this type of studies. Some previous studies in robotics have aimed to recognize users’ context, like a way to recognize joint attention behavior [1], attention [26], engagement [27] and perspective [24]. Although new techniques are constantly being developed, the robots’ capabilities in context-sensitive interactions have remained highly limited.

III. DESIGN PRINCIPLES

Previous work on the teleoperation of mobile robots has been mainly focused on navigational robots, whereas little is known about the teleoperation needs of mobile social robots. The basic design of our teleoperation system was created according to this previous knowledge on teleoperation for navigational robots. This section introduces the authors’ proposed techniques for the teleoperation of mobile social robots and the guidelines on which these techniques are based.

A. Guidelines for Navigational Robotics

Research on the teleoperation of mobile robots, using traditional 2D GUIs, has shown that distributing information on different locations of the interface may result in an increased workload and decreased performance of the operator [20]. These results may be caused by poor situation awareness of the operator. Situation awareness can be referred to as the level of understanding of the operator with respect to the environment around the robot that allows the operator to provide accurate instructions to the robot [28].

In [18], a study compares the usefulness of combining map and video information in a navigation task by comparing a side-by-side 2D representation and an integrated 3D representation. This study reports that the integration of map and video information in a 3D-based GUI positively affected the performance of the operator during navigation of the robot. However, the scope of this study is only a navigational task and it does not address important issues such as observing facial gestures of a customer and how they would affect the performance of an operator.

From a design perspective, Nielsen et al. [20] summarize that to improve situation awareness in human-robot systems it is recommended to: a) use a map, b) fuse sensor information, c) minimize the use of multiple windows and d) provide more spatial information to the operator.

Based on these recommendations, the authors have implemented a GUI that incorporates laser range data, a video feed, a 3D model of the robot used in this research and a 3D representation of the environment where the robot is located.

B. Proposed Techniques

In addition to these guidelines, two fundamental mechanisms for facilitating the teleoperation of a mobile

social robot are proposed: automatic gaze control and visualization of spatial relationships. The first one helps relieve the operator from continuously having to direct the camera towards the customer and the second one helps the operator retain the awareness that may be lost by providing the operator with autonomy.

1) Automatic Gaze Control

A critical requirement for the teleoperation of mobile social robots is to allow the operator to observe the facial expressions and gestures of the customer. Typically, this information is provided to the operator as a video feed coming from a camera pointing to the object or location of interest; in this way, the operator can understand the intentions of the customer. However, the actuation required by the operator to maintain the customer within the field of view of the robot's camera may increase the workload of the operator, especially when the customer may continuously move inside the environment. Thus, the automation of such task would become useful to reduce the effect of this workload on the performance of the operator. A feature called "automatic gaze control" is proposed to allow the system to automatically control the robot's gaze (i.e. camera direction) to follow a person's location and the person's face. The operator then is able to observe the facial expressions and gestures of the person interacting with the robot without the tedious responsibility of maintaining the robot's gaze direction manually.

2) Visualization of Spatial Relationships

We expect the automatic gaze control to reduce the operator's workload and improve the operator's awareness of the customer, but the problem remains that navigational tasks also require use of the robot's camera.

Some of our early tests showed that the use of automatic gaze control can even be disorienting to the operator. Operators sometimes reported that they lost track of spatial relationships around the robot after using automatic gaze control for a while. For example, as a customer moved around the robot, the spatial relationship between the customer and robot changed continuously; however, as the camera direction was controlled by the system, the operator was "out of the loop", and not highly aware of how much the camera is moving to keep the customer centered. When the operator then attempted to navigate the robot to another product, this often resulted in a moment of confusion. The operator seemed to require time to regain awareness of spatial relationships, resulting in the robot behaving strangely, e.g. spinning around (many operators intentionally did this to visually orient themselves within the room) or moving in a direction that did not make sense.

Therefore, we propose partially decoupling the problem of visual awareness for social interaction from the problem of visual awareness for navigation. By incorporating a graphical visualization of spatial relationships between the robot and objects in the environment, we can provide the operator with a second source of visual information to complement the information from the video feed. This should function somewhat like peripheral vision, enabling an operator to

maintain an internal awareness of the robot's position in its environment and easily transition between using the robot's camera for social interaction and using it for navigation.

One consideration we would like to emphasize is that to support mobile social interaction, it is important to display not only fixed objects such as walls, furniture, and products, but also the dynamic locations of customers in the shop, to enable an operator to easily locate people and interact with them.

Using graphical visualization of such spatial relationships in conjunction with a video feed should improve the operator's overall perception of the environment, by releasing the operator from the need to create a mental map of the objects in the environment, since they are represented on the GUI.

Through combination of the design recommendations presented in [18] with our proposed techniques, we expect that an operator will be more effectively able to control a robot, ultimately resulting in an improved human-robot interaction.

IV. SYSTEM IMPLEMENTATION

Given that our approach incorporates shared autonomy, implementation is necessary on both the robot side and operator side. This section presents how the concepts of visualizing spatial relationships and automatic gaze control are carried out within the proposed teleoperation system.

A. Robot side

The robot platform used in this study is the Robovie 2 humanoid communication robot. It comprises a mobile base (Pioneer 3) and an upper body that has two arms, each with 4 degrees of freedom (DOF) and a head with 3DOF. The arms can be used to point at the objects of interest as well for other gestures that complement its utterances. The head has a camera, a microphone and a speaker to allow an operator to gather information about the environment and the person the robot is interacting with. Robovie has two laser range sensors attached to its mobile base (about 10[cm] from the ground), one in the front and one in the back, in order to cover almost 360[deg] around the robot to detect obstacles.

1) Environmental Human Tracking Sensor System

A tracking system using laser range finders (LRF's) embedded in the environment was used to track the positions of people and localize the robot in the room. Six SICK LMS-200 laser range finders were placed around the perimeter of the room to minimize occlusions. They were set to a detection range of 80[m] with precision of 1[cm], each scanning an angular area of 180 [deg] at a resolution of 0.5[deg], providing readings of 361 data points every 26[ms].

The LRF's were mounted 85[cm] from the ground, a height chosen so the sensors could see above clutter and obstacles such as benches and luggage. Another reason for this placement was that at long range, the scan beams are spaced quite far apart (over 8[cm] apart at a range of 10[m]) and detection of small features like legs is difficult. Detection of larger targets, like a torso, is more robust at these distances.

The sensors were connected directly to a central data acquisition PC in another room, which then streamed all sensor data to the tracking server. The tracking server

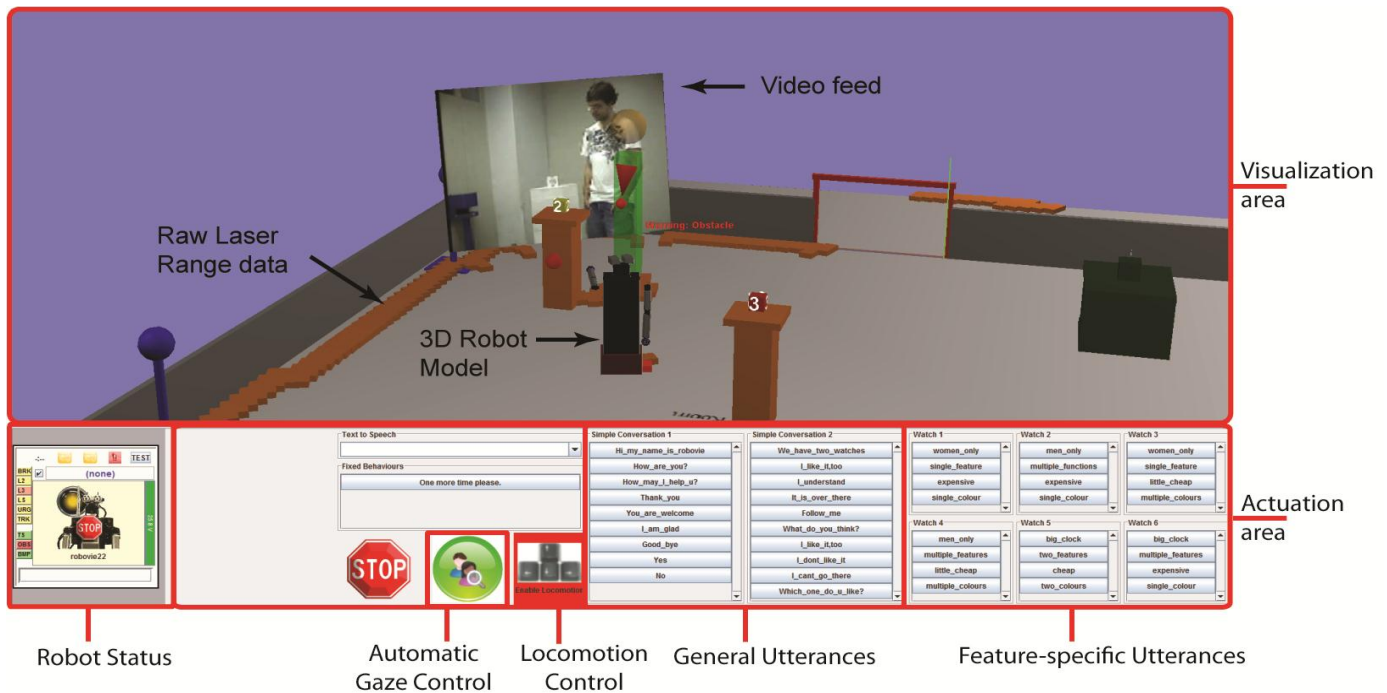


Figure 1. GUI with the implemented visualization of spatial relationships and automatic gaze control.

performed background subtraction on the scan data to remove fixed environment features, then combined the foreground data from all sensors. Particle filters were used to track each entity (human or robot) in the environment according to the algorithm presented in [27], and the system was used to correct the robot's localization according to the method described in [29]. The accuracy of this system varies according to sensor placement but has been measured at ± 6 [cm] in field deployments.

2) Automatic Gaze Control System

The proposed automatic gaze control system follows the face and upper body of a person once the subject has been identified by the environmental human tracking sensor system.

The position of the person (in 2D coordinates) is continuously obtained from the environmental human tracking sensor system. The height at which the robot's camera gazes the person is determined by the use of trigonometry and considering the distance separating the robot and the person interacting with it. This relationship is bounded by an angle ranging between 57[deg] and 60[deg] at a minimum distance of 1[m] from the person.

The automatic gaze control is enabled through the graphical user interface presented in this study. The operator clicks on the representation of the person of interest in the GUI and the system determines the angle at which the camera should point to, in order to maintain the person's face in focus and allow the operator to observe the person's facial expressions and gestures.

B. Operator Side

The data gathered by the robot's on-board sensors and by the environmental sensory system (human tracker module) are presented to the operator through a 3D-based interface.

The proposed GUI combines the two factors discussed in

Section III-B, and aims to allow the operator to identify and locate a person and objects of interest quickly, as well as to establish social distances accurately. Figure 1 shows an instance of the proposed system's GUI. The interface is divided in two sections: a visualization area (top) where a video feed is combined with a 3D model of the controlled robot and range data from laser sensors, and an actuation area (bottom).

1) Visualization

The visualization comprises three main elements: map and object representations, video feed and robot representation.

Map and Object Representations

The map representation of the environment was generated using the *a priori* known locations of objects (desks, watch stands, etc.) within the environment. These objects do not move in order to make the environment a static one. 3D computer-generated models of walls, environmental laser sensors, stands and tables represent the different objects of interest in the environment. The laser range data representation is shown as small blocks on the ground.

Video feed

The GUI incorporates a video screen into the 3D environment, the movement of which is synchronized to the movement of the head of the robot.

The video screen presents the image of the area at which the robot is looking.

In addition to helping the operator understand the environment in which the robot is located and avoid obstacles, video feedback can help the operator understand the intention of the person interacting directly with the robot.

Robot representation

It is important for the operator to understand the position, orientation and gestures of the teleoperated robot. In order to satisfy this requirement, a 3D model of Robovie II was



Figure 2. Photos of the experimental environment. The operator worked from a control room (left) while the robot interacted with evaluators in a separate room (right).

implemented. This 3D model can represent the different movements of the limbs, head and position and orientation of the robot within the 3D environment. The operator observes the environment from a tethered point of view anchored 3[m] behind the head of the 3D model representation of the robot. In addition, the status of the robot and safety warnings are displayed. Information regarding the status of the robot such as battery and identification of the robot are presented in the lower left corner of the GUI as presented in Figure 1. Obstacles are shown spatially on the floor as yellow and red points and they represent the level of danger of navigating the robot in a particular direction. Yellow points represent obstacles that are in the vicinity of the robot but that would not cause any danger to the robot or the customer and red points represent obstacles that would do so. Safety warnings are also shown to bring the operator's attention to possible dangers during the navigation of the robot. These safety warnings are shown on top of the head of the robot's representation and as a drop-down message from the top of the 3D environment visualization. These warnings are intended to help the operator navigate more smoothly and avoid collisions with obstacles or people.

2) Actuation

The main actuation tasks the operator can perform are locomotion, pointing, utterances, and gaze control.

Locomotion

The robot is able to move forward and rotate to the left and to the right around its own z-axis in order to reach a desired location. The operator drives the robot using the keyboard's arrow keys.

Pointing

In addition to these translation commands, the operator can also point to a given position or object.

The operator right-clicks a location or an object on the 3D environment and selects one out of two utterances the robot can say: "this one" or "that one". Both of these actions can be performed through the use of the GUI or using a mouse and a keyboard.

Utterances

There are two different sets of utterances given to the operator: general and feature-specific. The general utterances are those utterances designed to help the operator have a smoother interaction with the person, i.e. "would you like to see some other product?". The feature-specific utterances have been designed to allow the operator to give specific

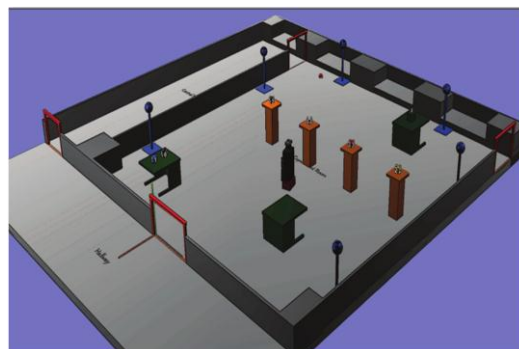


Figure 3. 3D view of the watch shop.

information about an object of interest to the person the robot is interacting with, i.e. "this product costs 5,000 yen". Both types of utterances are accessed by the operator by clicking on the button having the desired utterance's label. Some of the utterances are accompanied by head and arm gestures to make the robot more expressive.

Gaze

The operator manually controls the gaze of the robot by clicking on the video screen and dragging it to the direction where the operator wants the robot to look. The operator can enable the automatic gaze control by simply pressing a button on the GUI (Figure 1). When the automatic gaze control is used, the robot uses the data obtained from the human tracker module to calculate the locations of the robot and the customer. These data are then used to compute the robot's target gaze direction.

V. EXPERIMENT

An experiment was conducted to validate the combined effect of the visualization of spatial relationships and the automatic gaze control in the teleoperation of a mobile social robot. Whereas the automatic gaze control is expected to reduce the operator's workload and help the operator better understand the facial expressions and gestures of the person interacting with the robot, the visualization of spatial relationships is expected to help the operator better understand the locations of the robot and objects in the environment.

A. Scenario

The scenario chosen for the experiment had a Robovie 2 playing the role of a shopkeeper in an experiment room set up to represent a watch shop. The robot was controlled by an operator in another room, as shown in Figure 2.

In this scenario, various clocks and watches are located on stands and tables, Figure 3 shows an example of one of the configurations for the location of each of the six watches and clocks. The robot would navigate within the shop showing customers different watches at different locations. A collection of six external laser range finders was used to localize the robot and customers in the environment.

B. Hypothesis and Prediction

The automatic gaze function is designed to help operators look at the customer's face. Regarding this functionality, we propose three hypotheses:

First, we expect that the lower effort required to control the

robot's gaze will allow operators to direct the robot's gaze to the customer more easily:

(A1) Operators will more often direct the robot's gaze towards the customer when automatic gaze control is available.

Next, we expect automatic gaze control to reduce the operator's perceived workload, as it reduces the number of concurrent tasks the operator must perform. Hence:

(A2) The use of automatic gaze control will reduce the operator's perceived workload.

Gaze control tasks can also consume an operator's attention, distracting from or interrupting other actuation tasks such as locomotion and conversation. This can result in slow or inefficient operation. When gaze control is handled by autonomy, the operator can focus exclusively on these other tasks, enabling more efficient operation of the robot with less wasted time. Thus:

(A3) The use of automatic gaze control will result in shorter interaction length.

Furthermore, due to the operator's reduced workload and a better awareness of the customer's face and behavior thanks to the frequent gaze control, we expect that:

(A4) When automatic gaze control is available, customers will report higher satisfaction with the interactions.

Next, the visualization of spatial relationships was prepared to support operators in gaining awareness of spatial relationships around the robot. Thus, we believe that operators will be more aware of the surrounding situation, such as positions of the robot, customers, and watches. We make the following hypotheses regarding spatial visualization:

(V1) Visualization of spatial relationships will increase the operator's awareness of the robot's surroundings.

We expect that this increased awareness will require less effort by the operator to observe the environment by controlling the robot's locomotion or gaze direction, making operation easier:

(V2) Visualization of spatial relationships will reduce the operator's perceived workload.

We also believe that this will enable the operator to be more efficient in operating the robot, with less wasted time, thus:

(V3) Visualization of spatial relationships will result in shorter interactions.

As a consequence of this reduced workload and more efficient operation, we expect that the operator's overall performance will be improved, and that this will be visible in customer satisfaction:

(V4) Visualization of spatial relationships will result in higher satisfaction from the customer.

We are also interested in the combination of the two factors, and thus we conducted the experiment having both factors at the same time.

C. Procedure

There were 29 undergraduate students (15 female and 14 male, average 22 years old) who participated as operators. In addition, two undergraduate students (1 female and 1 male)

acted as evaluators, playing the role of customers in the interactions and providing evaluations from the viewpoint of the customer.

The operator participants were given an explanation of the task during the experiment, as well as a training session to practice using the GUI to control the robot. They were allowed to ask questions during this practice time to confirm their understanding of the different features of the GUI and their role in the experiment. The operator participants were located in a separate room from the location where the robot was, and they never directly observed the room until the end of the experiment.

To prevent the operators from learning the positions of objects in the room over time, it was necessary to change the layout of the objects after every trial. Five layouts were created: one for the training session and one for each of the four trials. In these layouts, watches were placed an average of 2.3[m] from the center of the room, with a standard deviation of 1.2[m]. In preliminary trials, we had observed that placing watches in close proximity to each other increased the difficulty of the teleoperation task, so we attempted to make each of the layouts used in the experiment similar in difficulty, by placing one pair of watches within 0.8[m] of each other, and spacing the remaining four watches around the room, at least 1.2[m] apart. The order of the layouts in each experiment was counterbalanced to avoid associating the layouts with any of the experimental conditions.

1) Operator's Role

The role of the participant working as an operator was to control the robot to behave as a shopkeeper at the simulated watch shop. The operator's tasks included locating a customer who is wandering inside the watch shop, approach the customer and show and talk about the different watches or clocks to the customer based on the customer's non-verbal expressions. Based on a customer's facial expressions, for example, the operator should identify the interest or lack thereof in a given watch or clock and introduce different features of the current watch or guide the customer to another watch that may be of more interest to the customer.

2) Customer's Role

Each of the evaluators behaved as a customer for each session. Thus, for each session, there were two customers visiting the shop. The customer was instructed to walk into the watch shop and wander around until the robot approached him/her. The customers were instructed to communicate both facial and body gestures and spoken language; though no specific instruction was given to use a particular cue. They were told what a good interaction should be like and to observe the behavior of the robot to later make their evaluations. They did not share their evaluation criteria. There was no scripted conversation; instead, the customer was given a situation and a watch that should be the target one. An example of a situation is that the customer will participate in a wedding and is interested in buying a watch.

In order to make each interaction equivalent, the customer is also instructed to wait until at least 3 different watches have been presented to make a purchase. If none of the watches that

have been presented within those 3 watches is the targeted one, the customer will wait until the robot presents the target one and finally purchase it.

Rather than employing novice participants to play the role of the customer in the interactions, we assigned two trained evaluators to participate as customers in all trials. This enabled us to obtain a consistent measurement of the robot's quality of service between trials. The use of trained evaluators was also necessary to ensure consistent behavior of the customers, as variations in customer behavior would have the effect of changing the task difficulty. For example, a customer who mainly uses facial expressions and body language to communicate would require much more visual attention from the operator than a customer who is primarily verbal in communication.

D. Conditions

A 2x2 within-participants experimental design was used with the following conditions:

- **Automatic Gaze Control** factor
 - Autogaze; in this condition, there is a button that enables the automatic tracking of the customers. This can be turned off by either pressing the button again, or manually moving the robot's head (via the GUI).
 - No-Autogaze; in this condition, the button is disabled, and the only way to control the robot's gaze (presumably to track and observe the customer) is direct manual control via the GUI.
- **Visualization of Spatial Relationships** factor
 - Spatial-Visualization; this condition adds 3D models of the objects (static, located in the room) and also avatar(s) of the persons (customers, keeping track of their current location; an example is provided in Figure 4).
 - No-Spatial-Visualization; in this condition, only the URG laser sensor raw data are shown, along with a 3D model of the robot, and the video feed coming from one of the robot's cameras (an example is provided in Figure 5).

E. Evaluation

A combination of subjective and objective techniques was employed to measure the performance of the operators in each condition as presented below.

- **Gaze time** was measured as the time the robot's gaze direction was actuated to face toward the customers or anywhere else (e.g. to seek for the location of the watch) either with manual control or automatic control.
- After each condition, subjective evaluations from the operator participants were provided to score on a 7-point Likert scale asking **the operator's awareness of surroundings** measured with an average of 7-point Likert scale items, asking for the awareness of the location of the robot, each customer, and each of the watches.
- **The operator's perceived workload** was evaluated using a NASA-TLX test [30] that the operator had to complete after each condition. The result of this test is

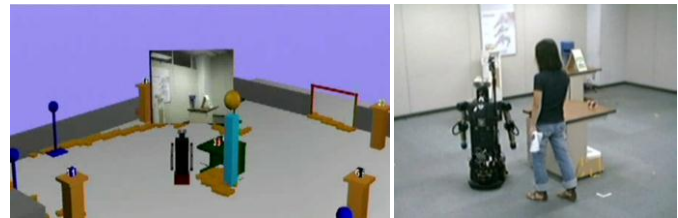


Figure 4. Interface in the Spatial-Visualization condition. The positions of the customer and the objects in the room are displayed in addition to range data and video from the robot's eye camera.



Figure 5. Interface in the No-Spatial-Visualization condition, where only laser range finder data and video from the robot's eye camera are displayed. Unless the robot is looking directly at the customer, it is difficult to identify where the customer is standing.

in a range between 0 and 100 points. Lower values represent lower workload, whereas higher values represent higher workload. In the context of this study, an operator was typically in a situation where he/she would suffer from overflow of tasks; thus, we consider that lower workload represents a situation where he/she were less suffered from such overflowing situation, and thus had a potential to respond to the customer in a more efficient way.

- As an indicator of how well the performance of the operator was, the authors timed the total **interaction length** of each condition. In our study, if an operator is efficient enough, the interaction length is supposed to be short. Customers were waiting for appropriate information to be provided, while operators were asked to identify the customers' interest and to choose the information contents to be presented by the robot. Clumsy operation and failure in identifying this situation would result in consumption of redundant time. Since the customers were waiting for information, such loss of time would result in a less engaging interaction and would lead to making customers bored.
- After each condition, subjective evaluations from the customer participants were provided to score on a 7-point Likert scale asking "how satisfactory was the robot's service?", i.e. **customer satisfaction**. In this scale, higher values represent higher satisfaction.

F. Hypothesis Testing

The results presented in Figure 6 share the following format: the blue dotted series represent the condition No-Spatial-Visualization, the red continuous series correspond to the Spatial-Visualization condition for the spatial relationships factor. The x-axis represents the "No-Autogaze" and "Autogaze" conditions corresponding to the experimental factor "automatic gaze control". A two-way repeated measures Analysis of Variance (ANOVA) was conducted with two within-subject factors, visual relationships and automatic gaze control, for all the results presented in this section.

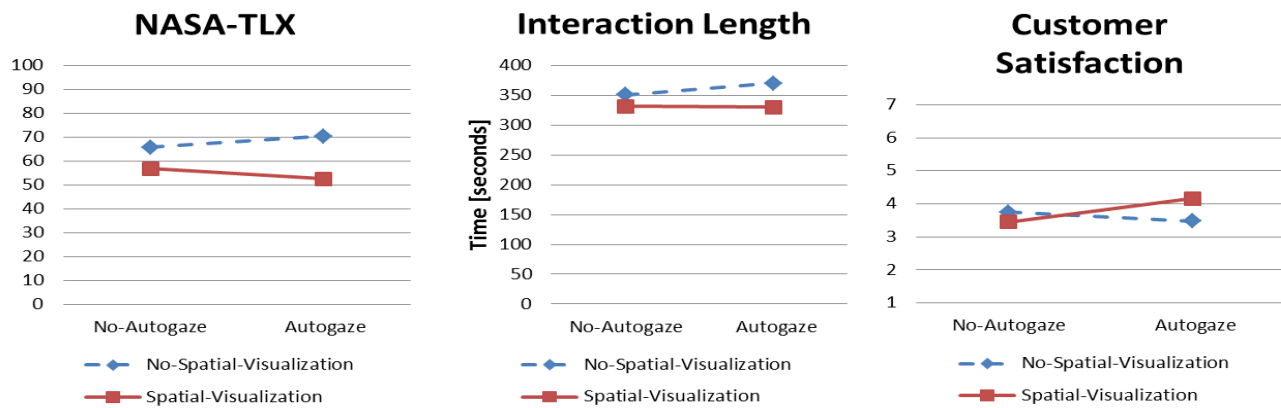


Figure 6. NASA-TLX (left), Interaction Length (center), and Customer Satisfaction (right).

1) Manipulation check

First, we confirmed that operators took advantage of the prepared functionality as designed. About automatic gaze, we analyzed “gaze time”. There was a significant main effect revealed with the automatic gaze control factor ($F(1,21)=26.384$, $p<.001$, partial $\eta^2=.557$), but no significant effect in the visualization of spatial relationships factor ($F(1,21)=.182$, $p=.674$, partial $\eta^2=.009$) or in the interaction of these factors ($F(1,21)=1.146$, $p=.297$, partial $\eta^2=.052$).

That is, as expected from hypothesis **A1**, gaze was actuated more often in the automatic gaze condition either with spatial visualization (avg. 233.4 [sec], s.d. 196.9) or without spatial visualization (avg. 197.5 [sec], s.d. 188.7) than under the manual gaze conditions either with spatial visualization (avg. 33.0 [sec], s.d. 29.9) or without spatial visualization (avg. 46.5 [sec], s.d. 40.0).

For spatial visualization, we analyzed “the operator’s awareness of surroundings”. There was a significant main effect revealed with both the visualization of spatial relationships factor ($F(1,21)=135.746$, $p<.001$, partial $\eta^2=.866$) and automatic gaze control factor ($F(1,21)=4.416$, $p=.048$, partial $\eta^2=.174$). Interaction between the two was not significant ($F(1,21)=1.004$, $p=.328$, partial $\eta^2=.046$). As expected, subjective evaluations of the operator’s awareness of surroundings were better with spatial visualization (in automatic gaze condition: avg. 5.98, s.d. .745, in manual gaze conditions: avg. 5.44, s.d. 1.39) than without spatial visualization (in automatic gaze condition: avg. 3.38, s.d. 1.03, in manual gaze condition avg. 3.17 [sec], s.d. .847).

These results confirmed our predictions as presented in **A1** and **V1**. As designed, automatic gaze control enabled more frequent actuation of gaze, and spatial visualization provided better awareness of surrounding spatial relationships.

2) NASA-TLX

The results for perceived workload measured using the NASA-TLX test are depicted in Figure 6 (left). A significant main effect was revealed with the visualization of spatial relationships factor ($F(1,21)=14.693$, $p=.001$, partial $\eta^2=.412$) but did not show significance with the automatic gaze control factor ($F(1,21)=.006$, $p=.939$ partial $\eta^2=.000$). Interaction within these factors was significant ($F(1,21)=4.984$, $p=.037$, partial $\eta^2=.192$).

The simple main effects in the interaction were further investigated. Regarding visualization of spatial relationships, there were significant simple main effects with both automatic gaze ($p<.001$) and with manual gaze ($p=.041$).

Regarding the simple main effect of automatic gaze control, there was a significant trend when there was no spatial visualization ($p=.066$), but the comparison was not significant when there was spatial visualization ($p=.206$).

Overall, these results partially confirm our hypothesis. Spatial visualization affected perceived workload in the way we stated in hypothesis **V2**; however, contrary to our hypothesis **A2**, automatic gaze alone did not reduce the perceived workload.

3) Interaction Length

The results for interaction length are shown in Figure 6 (center). A significant main effect was revealed in the visualization of spatial relationships factor ($F(1,21)=8.747$, $p=.008$, partial $\eta^2=.294$). No significant effect was shown by the automatic gaze control factor ($F(1,21)=1.190$, $p=.288$, partial $\eta^2=.054$), and the interaction between these two factors did not present a significant effect ($F(1,21)=.798$, $p=.382$, partial $\eta^2=.037$). From these results it can be seen that when the operator was provided with the visualization of spatial relationships, interactions were shorter.

These results support our hypothesis **V3** with respect to the effect from the visualization of spatial relationships. However, our prediction in hypothesis **A3** regarding the effect of automatic gaze was not confirmed.

4) Customer Satisfaction

Figure 6 (right) shows the results for customer satisfaction. No significant main effect was revealed for either the automatic gaze control factor ($F(1,21)=2.094$, $p=.163$, partial $\eta^2=.091$) or the visualization of spatial relationships factor ($F(1,21)=1.817$, $p=.192$, partial $\eta^2=.080$). However, the interaction between the visualization of spatial relationships factor and automatic gaze control factor was significant ($F(1,21)=5.431$, $p=.030$, partial $\eta^2=.205$). Simple main effects in the interaction were further investigated.

Regarding visualization of spatial relationships, the simple main effect was only significant when automatic gaze control was used ($p=.015$), but not significant under manual gaze control ($p=.250$).

Regarding the automatic gaze control, the simple main

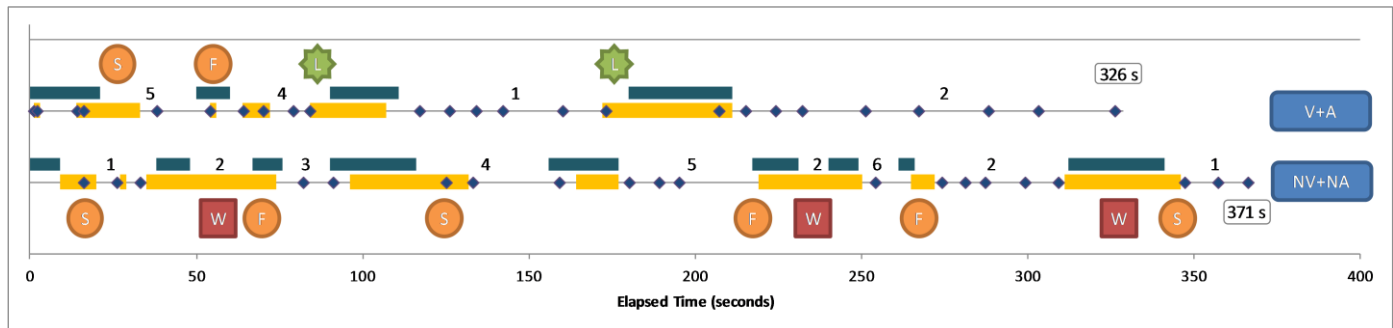


Figure 7. Timing diagram comparing events in two interactions. The top interaction included both object visualization and autogaze, and the bottom interaction used no object visualization or autogaze. Diamonds represent robot utterances, light-colored bars indicate robot locomotion, dark bars indicate customer movement, and numbers above each line show the number of the watch where the customer is currently stopped.

effect was only significant when spatial visualization was provided ($p=.003$), but not significant when there was no spatial visualization ($p=.367$).

Overall, these results partially confirm our hypotheses **A4** and **V4**. While these results do not confirm our hypotheses about either automatic gaze (**A4**) or spatial visualization (**V4**) alone, significant effects were observed when the two were used in combination. This result indicates that both automatic gaze and spatial visualization should be simultaneously used to gain a positive effect on customer satisfaction.

G. Case Studies

Our experimental results showed that the visualization of spatial relationships resulted in shorter interaction times. A “short” interaction refers in this paper, to an interaction where the robot will introduce the targeted watch after having shown the customer just one or two watches, being this the result of correctly observing the facial and body gestures of the customer. In such interaction, the robot will lead the customer in a proactive manner and will not have navigational issues represented by behaviors perceived by the customer such as “getting lost” in the environment which will consume time wastefully.

Figure 7 presents a case study representing a successful interaction with a length of 326 seconds. Successful interactions in our experiment had an approximate length of 330 seconds or less. It should be noted that whereas overall duration of the interaction does not directly affect or reflect its quality, unnecessary and awkward delays in utterances and motion contributed substantially to dissatisfaction in the interactions, and the effect of these delays can be seen to some degree in overall interaction length. In this section we will analyze two specific examples to illustrate how such delays can influence the quality and duration of interactions.

These trials show the robot being controlled by the same operator, in two interactions: one with both visualization and autogaze (V+A), and one without either (NV+NA). These trials were chosen to be as close as possible to the average lengths of trials in their respective conditions, and in addition they were the third and fourth trials, respectively, for that particular operator, which should minimize any difference due to learning effects. They also contain similar numbers of robot utterances (24 for the V+A case, and 20 for the NV+NA case).

It should be emphasized that there were a number of

different behavior patterns observed, and this analysis represents only two specific interactions which occurred. The intention here is to show a few concrete examples of how the experimental conditions affected robot performance, and also to give a sense of how differences in wait time contributed to differences in the customer experience, which may not be intuitively clear from the numerical results alone. Qualitative descriptions such as “too slow,” or “too long,” are casual observations of the authors and do not represent results from the formal coding process.

Figure 7 shows timing diagrams for these two interactions. The diamonds along each timeline indicate the start times of each utterance by the robot, and the light and dark bars on those lines indicate robot and customer movement inside the shop. The diagram is annotated with several letters, which represent events defined as follows:

- "L": Robot **leads** customer. Robot suggests a destination, moves there first, and the customer follows. This is a desirable behavior.
- "F": Robot **follows** customer. This is less desirable as it could mean that the robot is not effective at proactively introducing watches.
- "S": Robot is a bit **slow** or makes the customer wait a short time. Somewhat undesirable.
- "W": Robot makes customer **wait** a long time, and customer seems bored. This is very undesirable.

The sequence of events in these two examples proceeded as follows (C stands for customer, R stands for robot, O stands for operator).

1) With visualization and auto-gaze

In this case, the operator was able to navigate the robot adeptly, moving quickly from one watch to the next and sometimes controlling the robot to speak and drive at the same time. Twice during the interaction, the operator was able to lead the customer quite smoothly, offering to show a new watch then immediately turning and driving to that watch.

The customer satisfaction score for this trial was a 4 (average for this condition was 4.2). The customer never seemed to be waiting for a long time during this interaction, and only once, while following the customer to the first watch, did the robot’s operation seem to be a bit too slow.

- C enters, O turns on auto-gaze. R drives towards C after 1 second and greets C after 2 seconds, and then C describes desired features in the watch. R says "please look around freely".
(*Robot approached customer promptly*)
- C moves to Watch 5, R moves to that watch, introduces price, and waits.
(*"S" = short wait because C arrived first*)
- C is uninterested and moves to watch 4. R follows.
(*"F" = robot follows customer*)
- After R talks about Watch 4 for a bit, C does not seem interested in that watch, so R offers to see another watch and immediately leads the way. C follows.
(*"L" = robot leads customer*)
- R moves to Watch 1, and C follows. After C arrives, robot introduces Watch 1.
- After talking about Watch 1 and observing C's reaction, R offers to introduce Watch 2. C accepts, and R moves directly to Watch 2. C follows.
(*"L" = robot leads customer*)
- R introduces Watch 2 and talks about many features of the watch. There are some long silences, but robot continues to introduce features. C finally thanks the robot and buys the watch. C leaves and O turns off auto-gaze.

2) No visualization and no auto-gaze

This interaction took place 15 minutes after the interaction described above. This time, the operator notably had difficulty with robot navigation, which we ascribe to the absence of object visualization. Because of this, the operator often stopped the robot and turned its body or head several times to determine its location. Other times, the operator drove the robot along a wandering, inefficient path from one watch to the next. Several times, the customer was made to wait excessively due to this slow navigation.

Also, the robot did not greet the customer for some time at the beginning of the trial. It appears that the operator did not notice the customer entering the room, which we also ascribe to the no-visualization condition.

The customer satisfaction score for this interaction was a 3 (average for this condition was 3.8). In total, this interaction was 45 seconds longer than the previous one, which in this case was mostly due to inefficient locomotion, and the effect of this can be seen in the number of times the customer was made to wait. This interaction proceeded as follows.

- C enters but robot does not react. C goes to watch 1.
- R begins approaching after C has waited for 9 seconds. R greets C. C explains the desired features of the watch.

(*"S" = Robot very slow in approaching customer*)

- R says "look around freely" and begins driving to Watch 2. C is not visible to operator while R is driving. C quickly moves to Watch 2.
- R drives far out of the way, stopping and turning left and right while O adjusts the gaze direction several times to see obstacles and find its way to the destination. Meanwhile, C has waited at Watch 2 for 19 seconds. C gets bored and moves to Watch 3.
(*"W" = customer made to wait for a very long time*)
- R continues driving its long path, sees that C has moved, changes course to go to Watch 3, reaches C after 39 seconds, and introduces features of Watch 3.
(*"F" = robot follows customer*)
- C seems uninterested, so R offers to see Watch 4. The operator is attempting to "lead" C, but first R needs to look around so the operator can decide its path, so C walks towards Watch 4 first.
- Again, R drives a long, inefficient path, and C arrives much earlier. R arrives after 26 seconds and introduces Watch 4.
(*"S" = R's reaction is a bit late because C arrived first*)
- Robot then offers to introduce Watch 5, very close by. C walks directly there. R drives away from watch and back to it (again, the operator appears lost), while C waits. R talks about Watch 5, C seems bored and returns to Watch 2. R follows.
(*"F" = robot follows customer*)
- Robot looks around for C and slowly navigates back, looking for Watch 2. O adjusts gaze direction several times, looking at watches and tables. After waiting for 9 seconds and seeing that R is still far away, C moves to Watch 6.
(*"W" = customer made to wait for a long time*)
- R offers to introduce Watch 6 but C moves back to Watch 2. R follows.
(*"F" = robot follows customer*)
- R arrives at Watch 2 and explains several features. R then offers Watch 1 and begins to drive in the wrong direction. R stops, turns body and head a few times, and starts driving in the correct direction toward Watch 1. During this time C wanders around, waiting for R.
(*"W" = customer made to wait for a long time*)
- C goes to Watch 1 as R approaches. R arrives, O adjusts gaze direction manually to see customer's face, and R explains Watch 1 again. Customer listens to explanation and finally thanks robot and leaves.
(*"S" = R's reaction is a bit late because C arrived first*)

H. Qualitative Analysis of Interactions

These case studies illustrate a few typical behavior patterns that we believe contributed to the differences in customer satisfaction scores. To quantify these effects, an analysis was performed using video and audio data recorded during the experiment.

1) Analysis

For this analysis, we first identified four behavior patterns across all trials that seemed to result from poor operation, and which appeared to negatively influence the quality of the interaction. Generally speaking, these behaviors exhibited poor responsiveness or lack of initiative in the interaction. The four behavior patterns were defined as follows:

1. Only reactive

Behavior: The robot fails to lead the conversational interaction, and only passively responds to the customer. Operators were instructed to actively lead the dialogue, and to proactively suggest products and features, but sometimes the robot behaved reactively rather than proactively, leading to a poor interaction and low customer satisfaction. We believe this was often due to an operator having high workload or low situation awareness.

Coding: Coders judged the robot to be behaving reactively if long periods of silence occurred without the robot taking action, or if the robot uttered very few spontaneous utterances that were not in response to a customer's question.

2. Does not approach

Behavior: When operators had poor awareness of the robot's environment, they often did not notice a customer entering the shop. At these times the robot did not approach to greet the customers for some time, resulting in poor service and low customer satisfaction.

Coding: If the coders judged that the robot should have approached the customer but did not, or if the robot's approach was very late, they considered this behavior to have occurred. If the customer approached the robot immediately and the robot did not need to move to meet the customer, they did not consider this behavior to have occurred.

3. Customer initiative when moving

Behavior: Operators were instructed to lead the customer to different watches. However, when operators had poor awareness of the customer, they often failed to do this, and after some time the customers disregarded the robot and walked to the next watch on their own.

Coding: If the customer moved to another watch before the robot mentioned that watch or began moving towards the watch, this behavior was considered to have occurred.

4. Customer seems bored

Behavior: Overt signs of boredom were another indication that the operator was not controlling the robot in a responsive manner. Sometimes customers looked away, yawned, and expressed impatience or boredom with the robot.

Coding: This was judged by the coders subjectively based on the customer's expressions and actions.

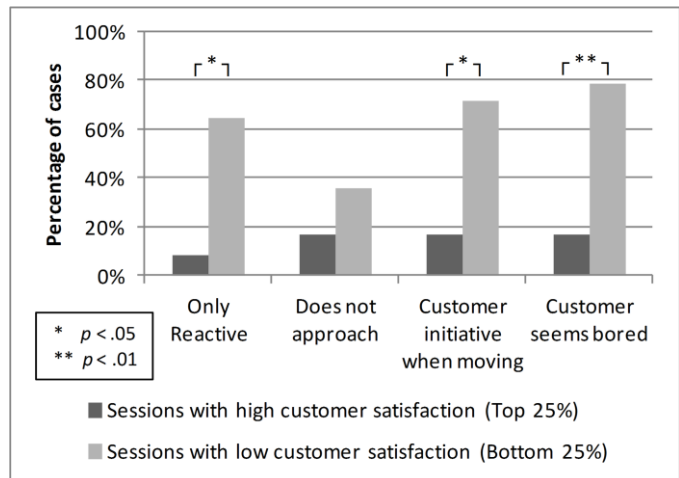


Figure 8. Frequency of occurrence for four undesirable event categories, compared between high and low customer satisfaction cases.

Next, two sets of trials were designated. Trials were sorted in order of customer satisfaction, then the top quartile of trials was designated as the "good" set and the bottom quartile as the "bad" set. Customer satisfaction scores ranged from 4.5 to 6 for the good set (mean 5.1) and from 1.5 to 3.0 for the bad set (mean 2.3).

Two independent coders analyzed recorded video and audio from these trials to identify whether any of these four behavior patterns occurred. Trials were presented to the coders in an arbitrary order, and the coders were not informed as to whether the trials belonged to the "good" or "bad" set. Coding for each category was binary for each trial – coders marked a "1" if the behavior occurred during that trial, or "0" if it did not.

2) Results

Between the two coders, Cohen's Kappa was computed to be 0.63, indicating reasonably good agreement. The coders then discussed the results to resolve their disagreements and produced a final set of ratings for the trials. The final results presented in Figure 8, show that all four of these behavior patterns occurred much more often in the "bad" set of trials than in the "good" set.

A Chi-squared analysis of the results revealed significant differences in three of the categories rated: "Only Reactive" ($\chi^2(1)=6.346, p<.05$), "Customer initiative when moving" ($\chi^2(1)=5.749, p<.05$), and "Customer seems bored" ($\chi^2(1)=7.583, p<.01$). No significance was found in "Does not approach" ($\chi^2(1)=0.420, p>.05$).

3) Discussion

Whereas our first results showed that the combination of spatial visualization and automatic gaze control can affect customer satisfaction, this analysis provides some insight into how those factors may translate into social interaction quality.

These four behavior patterns arise from two main factors: lack of responsiveness of the operator, and lack of initiative in conducting the interaction. Both of these factors are directly affected by the operator's situation awareness and workload.

The proposed technique of spatial visualization helps to increase the operator's awareness of the environment, and

automatic gaze control helps to increase the operator's awareness of the customer. Both techniques reduce the operator's workload by removing the need to actively actuate the robot's gaze to gain information. The operator thus attains higher awareness and lower workload, helping to avoid negative behavior patterns such as those studied in this analysis, finally resulting in higher customer satisfaction.

TABLE I. SUMMARY OF HYPOTHESIS RESULTS

Hypothesis	Hypothesis Statement	Confirmed by experimental data?
A1	When automatic gaze control is available, operators will more often direct the robot's gaze to the customer.	Yes
A2	The use of automatic gaze control will reduce the operator's perceived workload.	No
A3	The use of automatic gaze control will result in shorter interaction length.	No
A4	When automatic gaze control is available, customers will report higher satisfaction with the interactions.	Partially by itself, Yes in conjunction with V4
V1	Visualization of spatial relationships will improve the operator's awareness of the robot's surroundings	Yes
V2	Visualization of spatial relationships will reduce the operator's perceived workload.	Yes
V3	Visualization of spatial relationships will result in shorter interactions.	Yes
V4	Visualization of spatial relationships will result in higher satisfaction from the customer.	Partially by itself, Yes in conjunction with A4

VI. DISCUSSION

A. Summary

The results of our study indicate that when: a) an operator has an understanding of the spatial relationships, and b) the level of actuation the operator has to perform is decreased through automation of necessary and/or routine tasks, the operator can more effectively control the robot in social interactions. These results are summarized in Table I, and expressed in terms of the hypothesis proposed in this paper and how they were supported by our experimental data.

In our setting, the visualization of where the persons and the objects are, combined with automatic gaze control that frees the operator from tracking the person in order to observe them and thus determine their intentions, has resulted in improved customer satisfaction, that could be related to the reduced operator workload.

However, it was observed that the automation of the gaze, by itself, did not enhance the customer satisfaction.

The automatic gaze control enabled the operator to effectively observe the facial gestures of the customer while being aware of the surroundings of the environment. An appropriate visualization of the spatial relationships of the environment, as the proposed in this paper, allows the operator to have such intuitive understanding. If this visualization is not available while the automatic gaze control is, the operator may incur in continuous socially awkward movements of the robots head and body which in turn may convey an erroneous message to the customer.

Therefore, the authors would argue towards an approach in teleoperation architecture design that incorporates both the visualization of spatial relationships and the automation of processes that are necessary within an HRI context to aid the operator in improving their understanding of human non-verbal communication and which are crucial for social interactions.

This approach has applications both for teleoperated systems (for improving the operator performance), and also for research towards fully-automated systems, as first steps towards understanding the requirements necessary to implement the social processes to be automated (such as the automatic gaze control in our current work).

B. Limitations

In our current work, the robot can keep track of a single person within its field of view. However, it is conceivable that in a different social context, the robot would have to interact with multiple people at the same location (e.g. guiding a crowd at a museum). In the future, this could be augmented by additional mechanisms that e.g. automatically determine the gaze of the person or any pointing gestures. The visualization of spatial relationships currently relies on *a priori* knowledge of a static environment, as well as the existence of environmental sensors. Both of these limitations may be addressed by using traditional robot navigational and localization techniques and also by relying on on-board sensors.

The interaction addressed in the study was rather limited for simplicity. This was because our aim was to study the phenomena at the operators' side. Also, the assumptions of what constitutes good interactions and how they can be achieved were not based on a theoretical or empirical understanding of shopping interactions. Instead, they were based on criteria defined by the authors as good or bad (e.g. proactive / non-responsive behavior by the robot). Overall customer satisfaction when robots will be used in a real field will be affected by various factors, e.g. context role, expectation, and interaction design, thus we consider that the obtained results about customer satisfaction should be carefully interpreted.

VII. ACKNOWLEDGMENTS

The authors would like to acknowledge Florent Ferreri and Kyle Sama for their invaluable help on the development of the proposed system and the experiment carried out and presented in this paper. This research was supported by the Ministry of

Internal Affairs and Communications of Japan.

REFERENCES

- [1] D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita, "Simultaneous teleoperation of multiple social robots," HRI '08: Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction. ACM, 2008, pp. 311-318.
- [2] K. Zheng, D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita, "How many social robots can one operator control?" in Human-Robot Interaction (HRI), 2011 ACM/IEEE 6th Annual Conference on, March 2011, pp. 379-386.
- [3] A. Green, H. Huttenrauch, and K. Eklundh, "Applying the Wizard-of-Oz Framework to Cooperative Service Discovery and Configuration," in Robot and Human Interactive Communication, 2004. ROMAN 2004. 13th IEEE International Workshop on, 2004, pp. 575-580.
- [4] A. Steinfeld, O. C. Jenkins, and B. Scassellati, "The oz of wizard: Simulating the human for interaction research," in ACM/IEEE International Conference on Human-Robot Interaction (HRI), March 2009.
- [5] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of oz studies: why and how," in Proceedings of the 1st international conference on Intelligent user interfaces. ACM, 1993, pp. 193-200.
- [6] Álvarez, M., R. Galán, et al., "An Emotional Model for a Guide Robot." IEEE Trans. on Systems, Man, and Cybernetics. Part A: Systems and Humans, 2010, vol. 40, no. 5, pp. 982-992.
- [7] J. Scholtz, "Human-robot interactions: Creating Synergistic Cyberforces," in Proceedings of the Hawaii International Conference on Systems Sciences, January 6-9, 2003.
- [8] B. Pitzer, M. Styer, C. Bersch, C. DuHadway, and J. Becker, "Towards perceptual shared autonomy for robotic mobile manipulation," in Robotics and Automation (ICRA), 2011 IEEE International Conference on, May 2011, pp. 6245-6251.
- [9] J. L. Burke, R. R. Murphy, M. D. Covert, and D. L. Riddle, "Moonlight in Miami: A field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise," Human Computer Interaction, 2004, vol. 19, pp. 85-116.
- [10] P. Wells and D. Deguire, "Talon: A universal unmanned ground vehicle platform, enabling the mission to be the focus," Unmanned Ground Vehicle Technology VII, 2005, vol. 5804, no. 1, pp. 747-757.
- [11] B. Yamauchi, "Packbot: A versatile platform for military robotics," in Proceedings of SPIE 5422, 2004, pp. 228-237.
- [12] M. Buehler, K. Iagnemma, and S. Singh, The 2005 DARPA Grand Challenge: The Great Robot Race. Springer Publishing Company, Incorporated, 2007.
- [13] Chiang, H.-H., S.-J. Wu, et al., "The Human-in-the-Loop Design Approach to the Longitudinal Automation System for an Intelligent Vehicle." IEEE Trans. on Systems, Man, and Cybernetics. Part A: Systems and Humans, 2010, vol. 40, no. 4, pp. 708-720.
- [14] J. L. Drury, L. Riek, and N. Rackliffe, "A Decomposition of UAV-related situation awareness," in HRI '06: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction. ACM/IEEE, March 2006, pp. 88-94.
- [15] J. L. Drury, J. Scholtz, and H. A. Yanco, "Awareness in Human-Robot Interactions," in Proceedings of the IEEE Conference on Systems, Man and Cybernetics, October 2003, pp. 111-119.
- [16] J. Carff, M. Johnson, E. M. El-Sheikh, and J. E. Pratt, "Human-robot team navigation in visually complex environments," in Proceedings of 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, October 2009, pp. 3043-3050.
- [17] R. Meier, T. Fong, C. Thorpe, and C. Baur, "A sensor fusion based user interface for vehicle teleoperation," in International conference on field and service robotics (FSR), 1999, pp. 279-286.
- [18] C. W. Nielsen and M. A. Goodrich, "Comparing the usefulness of video and map information in navigation tasks," in HRI '06: Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction. ACM, 2006, pp. 95-101.
- [19] J. L. Drury, B. Keyes, and H. A. Yanco, "LASSOing HRI: analyzing situation awareness in map-centric and video-centric interfaces," in HRI '07: Proceedings of the ACM/IEEE international conference on Human Robot interaction. New York, NY, USA: ACM, 2007, pp. 279-286.
- [20] C. Nielsen, M. Goodrich, and R. Ricks, "Ecological interfaces for improving mobile robot teleoperation," Robotics, IEEE Transactions on, 2007, vol. 23, no. 5, pp. 927-941.
- [21] Crandall, J. W., M. L. Cummings, et al., "Computing the Effects of Operator Attention Allocation in Human Control of Multiple Robots." IEEE Trans. on Systems, Man, and Cybernetics. Part A: Systems and Humans, 2011, vol. 41, no. 3, pp. 385-397.
- [22] Crandall, J. W., M. A. Goodrich, et al., "Validating Human-Robot Interaction Schemes in Multi-Tasking Environments." IEEE Trans. on Systems, Man, and Cybernetics. Part A: Systems and Humans, 2005, vol. 35, no. 4, pp. 438-449.
- [23] H. Kuzuoka, K. Yamazaki, A. Yamazaki, J. Kosaka, Y. Suga, and C. Heath, "Dual ecologies of robot as communication media: thoughts on coordinating orientations and projectability," in CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, 2004, pp. 183-190.
- [24] Trafton, J. G., N. L. Cassimatis, et al., "Enabling Effective Human-Robot Interaction Using Perspective-Taking in Robots." IEEE Trans. On Systems, Man, and Cybernetics. Part A: Systems and Humans 2005, vol. 35, no. 4, pp. 460-470.
- [25] S. Woods, M. Walters, K. Koay, and K. Dautenhahn, "Comparing Human Robot Interaction Scenarios Using Live and Video Based Methods, Towards a Novel Methodological Approach," in Advanced Motion Control, 9th IEEE International Workshop on, 2006, pp. 750-755.
- [26] D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita, "Field trial for simultaneous teleoperation of mobile social robots," in HRI '09: Proceedings of the 4th ACM/IEEE international conference on Human robot interaction. ACM, 2009, pp. 149-156.
- [27] D. F. Glas, T. Miyashita, H. Ishiguro, and N. Hagita, "Laser-based tracking of human position and orientation using parametric shape modeling," Advanced Robotics, 2009, vol. 23, no. 4, pp. 405-428.
- [28] A. A. Nofi, "Defining and measuring shared situation awareness," in Center Naval Anal., Nov. 2000.
- [29] D. F. Glas, T. Miyashita, H. Ishiguro, and N. Hagita, "Automatic position calibration and sensor displacement detection for networks of laser rangefinders for human tracking," in Proc. IEEE/RSJ Int'l Conf. Intelligent Robots and Systems (IROS'10), 2010, pp. 2938-2945.
- [30] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in Human Mental Workload, 1988, pp. 139-183.



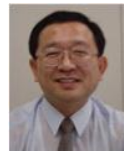
Andres Mora (B.Sc.'01, M.Sc.'06, Ph.D.'09) Andres Mora was born in San Jose, Costa Rica. He received his B.Sc. in Electronics Engineer from the Universidad Interamericana de Costa Rica and his M.Sc. and Ph.D. in Aerospace Engineering (Space Robotics) in 2006 and 2009 respectively at the Space Robotics Laboratory, Tohoku University, Japan. His research interests span path planning, teleoperation and control of mobile robots and the design of graphical user interfaces for mobile robots.



Dylan Glas received S.B. degrees in aerospace engineering and in earth, atmospheric, and planetary science from MIT in 1997, and he received his M.Eng. in aerospace engineering in 2000, also from MIT. He has been a Researcher at the Intelligent Robotics and Communication Laboratories (IRC) at the Advanced Telecommunications Research Institute International (ATR) in Kyoto, Japan since 2005. From 1998-2000 he worked in the Tangible Media Group at the MIT Media Lab. His research interests include networked robot systems, teleoperation for social robots, human-machine interaction, ubiquitous sensing, and artificial intelligence.



Takayuki Kanda (M04) received his B. Eng, M. Eng, and Ph. D. degrees in computer science from Kyoto University, Kyoto, Japan, in 1998, 2000, and 2003, respectively. This author became a Member (M) of IEEE in 2004. From 2000 to 2003, he was an Intern Researcher at ATR Media Information Science Laboratories, and he is currently a Senior Researcher at ATR Intelligent Robotics and Communication Laboratories, Kyoto, Japan. His current research interests include intelligent robotics, human-robot interaction, and vision based mobile robots.



Norihiro Hagita (M85 SM99) received his Ph.D. degree from Keio University (Japan) in 1986 in electrical engineering and joined Nippon Telegraph and Telephone Public Corporation (NTT) in 1978. He engaged specially in developing handwritten character recognition. He also stayed as a visiting researcher at Prof. Stephen Palmer's lab in University of California, Berkeley (Dep. of Psychology) during 1989-1990. He is currently the director of ATR Intelligent Robotics and Communication Laboratories (IRC) at the Advanced Telecommunications Research Institute International (ATR) in Kyoto, Japan.