Essential Features of Telepresence Robots

Munjal Desai, Katherine M. Tsui, and Holly A. Yanco
University of Massachusetts Lowell
One University Avenue, Lowell, MA 01854
Email: {mdesai, ktsui, holly}@cs.uml.edu

Chris Uhlik
Google Inc.
1600 Amphitheatre Parkway
Mountain View, CA 94043
Email: cuhlik@google.com

***Abstract*—Telepresence robots are mobile robot platforms capable of providing two way audio and video communication. Recently there has been a surge in companies designing telepresence robots. We conducted a series of user studies at Google in Mountain View with two different commercially available telepresence robots. Based on the data collected from these user studies, we present a set of guidelines for designing telepresence robots. These essential guidelines pertain to video, audio, user interface, physical features, and autonomous behaviors.**

I. Introduction

Contemporary telepresence robots can be described as embodied video conferencing on wheels. Over the last decade, companies such as Anybots, HeadThere, InTouch Health, iRobot, RoboDynamics, VGo Communications, and Willow Garage have produced these robots with the intention for them to be used in a wide variety of situations ranging from ad-hoc conversations at the office to patient rounds at medical facilities. When comparing these first generation telepresence robots, there are few similarities between them, if any. Each company designed their own robot for the application of telepresence, and all arrived at very different robots. Our goal was to understand what the essential features should be for the next generation of telepresence robots. In July and August 2010, we conducted a series of studies for several office-related use cases [1] at Google in Mountain View, California, with two telepresence robots: Anybots' QB and VGo Communications' VGo (in beta and alpha testing respectively).

In this paper, we provide a set of guidelines that we believe are essential for all telepresence robots based on the results of our studies and our observations during them. These guidelines are not merely a list of desirable features, but, rather, they constitute an essential set of features that must be incorporated in telepresence robots. In coming up with these guidelines, we have constrained ourselves to only those features that are technically feasible given the current technology.

II. User Tests of Telepresence Systems

We conducted five user studies over a period of two months. Through these user studies, we investigated the different aspects of the two telepresence robots from the robot driver's perspective, the perspective of the person physically present with the robot, and the bystander's perspective.

Study 1: Initial user impressions. The aim of this study was to evaluate how easy it was to use the robots and also to gauge the participants' impressions of the robots after initial interactions. We conducted a between-subjects study in which 31 participants used either the QB or the VGo robots. The participants were asked to perform three tasks. The first task involved driving the robot from an unknown starting location, then moving through a cube area to a specified conference room. The second task involved having a short one-on-one conversation with one of the experimenters, and the third task involved a whiteboard interaction. All three tasks occurred in one 60 minute session. In a post-experiment questionnaire, the participants described what features they liked about the robots, which they disliked, and what features, if any, were missing. We also asked the participants for their thoughts about the robot's height and the camera.

Study 2: Remote user interactions. The aim of this study was to evaluate if telepresence robots could be used in formal, conference room meetings. We selected 6 remote participants from the United States and Europe who had recurring meetings with teammates in Mountain View. Participants used the robots to attend their meetings in place of their normal video conferencing setup. Meetings ranged in length from 15 to 60 minutes; participants used the robots for one to eight meetings. (Details can be found in [1].)

Study 3: Video comparisons. The aim of this study was to compare the video streams from the QB and VGo robots against a Sprint EVO Android phone. The EVO phone streamed video through the Qik [2] application using its wireless internet connection. We conducted a within-subjects study in which 24 participants used both the QB and the VGo robots with the EVO phone mounted on each. A session took 60 minutes. We asked the participants to rate the video from the robot and EVO with respect to field of view, latency, ability to perceive scale, contrast, resolution, color depth, and quality of degradation while driving the robot. Before the start of each run, participants were asked to read the letters on an eye chart [3] that was four feet from the robot. The participants were asked to read the letters from both the phone and the robot's video display.

Study 4: Walking and talking conversations. The aim of this study was to investigate an ad-hoc scenario involving movement while simultaneously having a conversation. We felt this was an important scenario that needed to be explored, as mobility is the characteristic that differentiates telepresence robots from video conferencing technologies. We conducted a between-subjects study with 24 participants in which one person operated the robot (robot driver) and had a walking conversation with another person who was physically present with the robot (walker). Both participants walked with and

TABLE I
KEY FEATURES OF THE ANYBOTS' QB (LEFT) AND THE VGO (RIGHT) ROBOTS USED IN JULY AND AUGUST 2010 STUDIES.

	QB	**VGo**
Unit cost	$15K	$5K
Drive	2 wheels (dynamically balancing)	2 wheels and 2 casters
Top speed	3.5 mph	2.75 mph
Height	3'2" to 6'3" (manually adjusted)	4'
Weight	35 lbs	18 lbs
Battery life	4-6 hours	6 or 12 hour battery option
Microphones	3 on top of head (equally spaced)	4 around video screen (2 front, 2 back)
Speakers	1 on top of head	2 (woofer in base, tweeter in head)
Screen size	3.5" diagonal	6" diagonal
Number of cameras	1 front facing and 1 facing down	1 front facing
Resolution of cameras	5 mega pixel for front facing	3 mega pixel
Camera tilt	no (fixed)	180 degrees
Deictic reference	yes (laser pointer)	no
Operating systems	MacOS with Firefox 3.6	Windows 7/Vista/XP
Navigation control	keyboard (arrow keys or WASD)	mouse "Click and Go" or arrows keys
2-way audio	yes	yes
2-way video	no (planned feature)	yes
WiFi AP switching	no (planned feature)	yes

drove the robot. There were 13 60-minute sessions (6 VGo and 7 QB); for two sessions, one participant did not show up and an experimenter substituted for the missing role. We asked the participants to rate their ability to operate the robot while talking, how useful they thought autonomous robot behaviors would be, and the height of the robot.

Study 5: Bystander impressions. At the end of August, we asked the general office population who had seen the robots in the hallways about their experiences. We created an anonymous online survey and placed comment boxes in two locations. We received a total of 10 responses. Throughout the summer, we also noted people's comments regarding the robots which were either directed at us as the robots' handlers or overheard.

III. DESIGN GUIDELINES

The recommendations provided in this section are based on objective and subjective data from the user studies as well as the experimenters' observations. We conservatively estimate spending over 320 hours in total using, maintaining, or running experiments with the QB and VGo telepresence robots. For ease of comprehension, the recommendations have been grouped together based on different aspects of telepresence robots.

The term "telepresence robot" encompasses a wide range of robots, which vary with respect to mobility, size, capabilities, etc. For the purposes of this paper, we restrict the definition to mobile robots with two-way video and audio with the capability to transfer information over wireless networks.

A. Video

Video information is critical in telepresence robots for conversation and navigation. Due to the mobility afforded by these robots, the information must be transferred wirelessly. Video streams constitute a significant portion of the data being transferred and can be adversely affected by the network connection. The quality of a wireless connection is influenced by several factors including bandwidth, latency, and packet loss. Working within these limitations requires tradeoffs between the different characteristics of a video feed. Based on the results and our observations, the guiding principle for video streams for telepresence robots is to have two video profiles: one while the robot is mobile (dynamic video profile), and another profile for when the robot is not moving (stationary video profile). Two profiles are needed because the required video characteristics are mutually exclusive at times.

TABLE II
PARTICIPANTS' PREFERENCES FROM STUDY 3 (VIDEO). THE HIGHLIGHTED CELLS HAVE SIGNIFICANT RESULTS ($p<0.05$). THE RATING QUESTIONS WERE ON A 7 POINT SEMANTIC DIFFERENTIAL SCALE.

Video characteristics	QB *vs* EVO	VGo *vs* EVO	QB *vs* VGo
Overall quality	QB (4.17 *vs* 4.08)	Same (4 *vs* 4)	QB (4.17 *vs* 4)
Field of view	QB (3 *vs* 3.5)	Same (3 *vs* 3)	QB (3.5 *vs* 3)
Scale perception	QB (5.25 *vs* 4.42)	VGo (4.67 *vs* 4.08)	QB (5.25 *vs* 4.67)
Pauses in video	QB (4.17 *vs* 3.25)	EVO (5.92 *vs* 3.82)	QB (4.17 *vs* 5.91)
Latency	QB (3 *vs* 5.17)	EVO (4.42 *vs* 4.33)	QB (3 *vs* 4.42)
Contrast	EVO (3.91 *vs* 5)	VGo (5.08 *vs* 4.08)	VGo (3.92 *vs* 5.08)
Resolution	EVO (3.17 *vs* 3.67)	VGo (3.75 *vs* 2.67)	VGo (3.17 *vs* 3.75)
Color depth	EVO (3.5 *vs* 4.8)	VGo (4.75 *vs* 3.83)	VGo (3.5 *vs* 4.75)
Degration in quality	EVO (4.16 *vs* 4.08)	EVO (4.58 *vs* 4.67)	VGo (3.58 *vs* 4.58)

Dynamic video profile: Video is the most important sensor information while controlling a telepresence robot. Study 3 (video) examined the video characteristics of the QB and VGo robots with the EVO phone while the robot was mobile. The results shown in Table II indicate that the participants rated the QB robot as having better dynamic characteristics such as low latency, fewer pauses in the video stream, wider field of view, and better scale perception as compared to the VGo robot and the EVO phone. The EVO phone was rated better than the VGo robot for its low latency and few pauses in video. The importance of these dynamic characteristics is highlighted by the response of the participants when they rated the devices for the best overall video; six participants selected QB, four selected the EVO phone, and only two selected VGo. These results suggest that while driving the robots, participants were willing to sacrifice the resolution, color depth, and contrast of the incoming video feed for dynamic characteristics like low latency and higher frame rates.

Stationary video profile: While the QB's video feed was better while the robot was moving, VGo's video was better for a non-mobile situation. Participants rated VGo as having better contrast/white balance, resolution, and color depth than QB and EVO. In Study 3 (video), participants were able to read more characters on an eye chart [3] from 4 feet away via the VGo robot ($\bar{x}$=21.5, SD=5.8) than QB ($\bar{x}$=11.3, SD=3.9) ($t(11)$=5.96, p <0.0001, using a two-tailed paired t-test). In the whiteboard interaction task in Study 1, it was easier to view the large map drawn on a whiteboard using the VGo robot than the QB robot.

Video resolution: High video resolution is necessary at times. Drivers may need to read maps on the wall and office numbers, for example. While it is not feasible to expect high enough video resolutions to read all signs, the drivers should be able to read signs less than a few feet away with sans serif fonts with character height at least 5/8 inch [4] while not in motion. The real constraint is not the availability of high resolution cameras but the availability of bandwidth to transmit the video data.[1] While reducing the dynamic video compression rates can be useful in situations like these, other technologies can also be utilized. For example, the VGo robots are equipped with the ability to take high resolution screenshots and transfer them to the driver's desktop. However, this feature was utilized by only one participant in Study 1. We speculate the main reason for this lack of use is the time required to take an image, transfer it, and then view it. A physical or digital zoom could also provide a close up for detailed information. Another viable alternative would be to run an optical character recognition engine (similar to Google Goggles [5]) on the robot and present the text information to the robot drivers overlaid on the video like augmented reality.

Graceful degradation: There will be fluctuations in the quality of network connection, and the video stream must self-adjust to those changes. For example, if there is packet loss, the video codec should update the video with the incoming packets rather than waiting for an entire frame. As the bandwidth decreases, the compression on the video can be increased, thereby decreasing the video quality. This gradual degradation serves two purposes. First, it provides immediate feedback about the network connection. Second, it helps to maintain a low latency connection. This degraded but persistent video is especially important when the robot is in motion.

If at some point the connection deteriorates to an extent that video can not be transmitted, robot commands should not be transmitted either. There were instances in which the video was no longer being updated, but the drive commands were being sent. We observed both the QB and VGo collide into some object or a wall when this loss of video happened.

B. Audio

Audio quality: The most important component of communicating through a telepresence robot is the conversation itself. The audio quality must be comparable to that of a landline phone conversation. In Study 4 (walk and talk), we found that the robots had several audio issues including echo, feedback, and cut outs through analysis of the robot driver's screen captured video. The robot drivers could hear themselves talking through the system (QB: 3 of 13 runs, VGo: 8 of 11). There was also feedback through the laptop (QB: 7 of 13 runs, VGo: 2 of 11). In 4 of 13 runs using the QB robot, the audio stream that the driver received was missing parts of sentences. There were no runs in which the VGo robot had choppy audio. These audio issues make it difficult to have any conversation, let alone a natural conversation.

Volume control: It is important both for the person operating the telepresence robot and for the person physically present with the robot. In Study 4 (walk and talk), the robot drivers also had difficulty hearing the walkers due to the audio levels being excessively low in 11 of 14 QB runs; there were no runs in which the VGo robot had low audio levels. For all of the QB runs, a 15-inch MacBook Pro running OS X was used. The same MacBook running Windows Vista through Bootcamp was used for the VGo runs. It should be noted that the audio levels on the user interface were set to the maximum level. The audio was output to the laptop's internal speakers, and the participants did not have access to headphones.

The person physically present with the robot may need to adjust the robot's volume when moving from one space to another. In Study 4 (walk and talk), we asked the participants to move down a hallway from a noisy kitchen area to a quiet area near offices and cube spaces. The appropriate volume for the kitchen area was much too loud for the office area. While conducting Study 4, a bystander who worked in an office area along the path directly requested the experimenters to lower the robot's volume. We received three comments in Study 5 (bystander impressions) which noted the robot's excessive volume [6]. We were able to reduce the VGo's volume to 70%;[2] however, volume control did not exist on the QB robot.

C. User Interface

The user interface (UI) is a critical component of the telepresence system. It is the driver's portal to the remote world. The UI must be simple, easy to use, not distracting, and provide the necessary functionality without overwhelming the driver. User interfaces for controlling remote robots have been well researched (e.g., [7]). Important lessons for creating a video centric UI and fusing sensor information to create an uncluttered UI can be learned from them.

UI type: A UI that is platform independent ensures that the users can access the robots from any computer available to them. Platform independence is particularly important for commercial environments in which the company policy might only allow certain operating systems. Creating a platform independent web based UI over a platform independent application based UI is also recommended. A web based UI allows the user to start using the robot without any installation procedure and hence without requiring administrative permission on the system.

Sensor information: Providing relevant and accurate sensor information to the robot driver is necessary. However, too much information can overwhelm the driver and can be counterproductive. For example, the VGo robots had distance sensors that would inform the driver if an object was close on the front, left, or right. Unfortunately, this information was not found to be useful by the participants. The robots were operating in an office environment with narrow hallways

[1] QB and VGo had cameras with 3 megapixel resolution or higher.

[2] VGo drivers can now change the robot's output volume.

that frequently triggered the sensors. We believe the drivers quickly ignored the distance warnings. In Study 1 (initial user impressions), we had 7 participants drive the VGo robot with the distance information not displayed and 12 participants with the distance information displayed. There was no significant difference between the number of the robot's collisions with the environment when the sensor information was displayed (n=12, $\bar{x}$=1.67 hits, SD=1.37) and when the sensor information was not displayed (n=7, $\bar{x}$=1.86, SD=1.21) (t(17)=0.3039, p=0.7).

Feedback: Network latency as well as physical inertia make providing appropriate feedback crucial in controlling remote robots. Feedback provided to the driver is useful only if it is correct and provided in a timely fashion. For example, the later versions of the UI for QB used a feature that reduced the apparent latency in the video stream when the robot was moving. When the driver instructed the robot to turn right or left via the UI the video feed would immediately turn right or left indicating that the robot responded to the commands. This gave the appreance of there being minimal lag in controlling the robot and the participants reacted positively to this feature. Since the UI responded based on input from the driver rather than waiting for the actual data from the robot, this feature worked even when the network connection to the robot was lost, thus leading the driver to believe that they still had control over the robot. Participants often found this confusing since they could apparently turn the robot but not make it move forward or back.[3]

The VGo robot provided wireless status information. However, while switching access points, the wireless status would not update until the robot had already switched to the new access point; this process took upwards of 20 seconds. During this time, the participants were left with inaccurate and outdated information. The robot would show good wifi status, show no connectivity for a moment, and then return to show good wifi status.[4]

Integrated map: In Study 4 (walk and talk), 20 participants out of 24 reported wanting a map of the environment. Incorporating a map in the UI would be analogous to a person looking at a map on the wall when in an unfamiliar location. The "you are here" dot typically found on building maps can be provided through the robot's sensor information through localization of distance information or other means. A map can also be used to provide a goal destination for where the robot should go as further described in the next section.

D. Physical features

Robot height: Ideally, the driver should be able to change the robot's height to any desired length remotely. If such a mechanism is too complicated and expensive, the robot should at least be able to switch between two preset heights so that the driver can be eye-level for sitting and standing conversations. In Study 4 (walk and talk), the drivers of the robots wanted the height to be between 5 and 6 feet (5'-5.5' = 10 and 5.5'-6' = 12). The person interacting with the robot (i.e., walking beside the robot and having a conversation with the driver) wanted the robot's height to be in the same range (5'-5.5' = 14 and 5.5'-6' = 9). We believe that more participants wanted the height of the robot to be in 5' to 5.5' range as walkers because they did not want to look up to a robot. The QB robot's recommended height was in the 5' to 6' range and was comfortable for talking to people that were standing up. For sitting conversations, participants preferred that the robot's height be the same as them sitting down. The VGo robot's height was in a comfortable range for the sitting situation.

Robot speed: The robots should be able to move at average human walking speeds of about 3 miles per hour. This speed is especially important for situations in which the driver is walking with a person or a group of people while talking to them. Both the QB and the VGo robots were capable of achieving speeds around 3 mph (QB: 3.5mph and VGo: 2.75mph).

Multiple cameras: The robots must have at least two cameras. One of them must be a forward facing camera that can be used during conversations and while driving to view the path ahead. The conversation camera must be high enough off the floor to show the face of the person physically with the robot. Unless the field of view for the conversation camera is vertically wide enough, it is not possible to see both the person's face and see the area immediately surrounding the robot. A second camera showing the base of the robot can be used as a reference point by the driver for navigation. The QB robot had one camera facing forward and one camera looking at the base of the robot. The VGo robot had one camera, but the participants could tilt the camera up and down as needed.[5] In Study 1 (initial user impressions), participants had significantly fewer hits (p=0.04, t(28)=2.12) with the QB robot (n=11, $\bar{x}$=0.73, SD=1.19) than the VGo robot (n=19, $\bar{x}$=1.67, SD=1.28).[6] In general, QB's down facing camera was found to be useful by the participants.

Wide field of view: The front facing camera must have a wide field of view (FOV). A wide FOV is essential during navigation because it provides the driver with better situation awareness (SA) [8]. In the post experiment questionnaire in Study 4 (walk and talk), 91% of the participants (22 of 24) indicated that they wanted a a wider FOV.

Independent head/torso: Participants in Study 1 (initial user impressions), Study 2 (remote user interactions), and Study 4 (walk and talk) often asked if the robot's 'head' could pan and tilt. In the post experiment questionnaire for Study 4 (walk and talk), 21 of 24 participants wanted a head that could pan, and 18 participants wanted a head that could tilt.

Having a head that can pan and tilt is useful in several situations. While having a walking conversation, having head that allows the screen and conversation camera(s) to turn and face the participant is useful. It allows the person walking next to the robot to walk in a comfortable pose and posture. In Study 4 (walk and talk), the walkers were diagonally in

[3]Anybots was informed about this, and they are addressing this issue.

[4]VGo Communications has since reduced the AP switch times and addressed the wireless status issue.

[5]For Study 1 (initial user impressions) in the mouse navigation mode, the VGo's camera was set to auto-tilt based on the robot's speed. In the keyboard navigation mode, the camera had to be manually tilted.

[6]One participant who used a QB robot was excluded as the number of hits was beyond two standard deviations (8 hits total).

front of the robot and looking back at the robot while walking ahead [1]. The walker's face was visible over 50% of the time in only 2 runs out of 11 for VGo and 6 out of 12 for QB. In the whiteboard interaction portion of Study 1 (initial user impressions), participants who used a QB robot had to move their robot backward away from the whiteboard in order to view the top portion. However, we observed that moving the robot backward made it difficult to read the contents of the whiteboard due to low video resolution.

Panning and tilting the robot's head can also be used as social cues. In the meeting room, participants in Study 2 (remote user interactions) physically panned the robot toward the person speaking to show their attention. A head that can tilt can also be used to provide feedback via nodding, much as people do.

Access point switching: As robots move around, they have to switch access points. Depending on the environment, there can be multiple access points that the robot might have to connect to while moving. While switching access points, the flow of information to and from the robot might get interrupted, which can be a potential problem, especially when the robot is being teleoperated. While autonomous behaviors can help mitigate the issue of safely controlling the robot, the video and audio streams will still be affected. To avoid these issues, alternatives must be explored. One approach would be to have two wireless interfaces, ensuring that both do not switch at the same time. Another alternative may be to support connecting to 4G networks, which are expected to have bandwidth of up to 1 Gb per second in low mobility and stationary situations.

E. Autonomous navigation

Autonomous navigation behaviors are desirable because of safety reasons and for ease of use. For example, a remote driver may see in their video feed a person exiting a conference room and stop so the person can pass in front of the robot. Under teleoperation, the robot may or may not stop in time, depending on the delay given the robot's video feed to the driver and then the navigation command back to the robot. Processing the sensor data locally allows the robot to take immediate action, thereby providing a tighter closed loop control of the robot. Hence, autonomous behaviors allow for better control of the robot under varying network conditions.

Driving a remote robot is also cognitively demanding. In Study 1 (initial user impressions), we asked participants to think aloud while driving the robot. Only 4 of 30 participants talked while driving the robots.[7] In Study 4 (walk and talk), we asked the robot drivers to rate their distribution of attention between the talking task and navigation task tasks (1 = talking task, 5 = driving task). The robot drivers reported giving significantly more attention to the driving task ($p<0.001$, $\chi^2(4)=20.583$). Autonomous behaviors such as guarded motion, obstacle avoidance, and navigation planners have been designed and implemented for robots (e.g., [9]). They have the potential to ease cognitive demands associated with teleoperating a remote robot. The sections below highlight the different aspects of autonomous behaviors that must be considered for telepresence robots.

Assisted navigation: It is imperative that telepresence robots move safely. In Study 1 (initial user impressions), two thirds of the participants (21 of 31) collided with the environment while driving the robot through an office space. Collisions generally occur when a driver does not have good SA of the robot's immediate surroundings. While adding more sensors can help with better SA, sending the sensor data to the driver is not always feasible due to bandwidth restrictions or desirable due to cognitive overload. However, the same data can be used by autonomous behaviors to safely navigate the space. For example, the QB robots used a Hokuyo URG laser for assisted navigation for passing through doorways. This assistance could also have been used while driving down hallways; e.g., if the robot drove at an angle towards a wall, then QB could have autonomously corrected its direction.

Human-speed navigation behaviors: While both the QB and VGo robots were capable of moving close to walking speed, we observed that the average speed of 1.4 mph for participants in Study 4 (walk and talk) was lower than a walking pace of 3 mph ($p<0.0001$, $t(25)=7.55$). It could be argued that these participants were novice robot drivers. However, the experimenters, who had over 4 years experience driving remote robots, drove the robots in a time test. The fastest run for the Study 1 180' path took 73 seconds (1.68 mph) using a QB.[8] It should be noted that physically walking the path took the experimenters an average 38 seconds or 3.23 mph. Thus, the speed at which the robots can be driven is not limited by the maximum speed of the robots, but rather the network latency, the driver's familiarity with the remote environment and situation awareness.

In order for telepresence robots to safely approach human walking speed, autonomous behaviors must be utilized. In Study 4 (walk and talk), from the perspective of both the robot driver and the walker, a "follow person" behavior and a "go to destination" mode were rated as potentially quite useful ($p<0.01$ using a χ^2 test) on a scale of 1 (not useful) to 5 (very useful). Of the 24 participants, 20 participants wanted a "follow the person" autonomy mode, and 19 participants wanted a "move to specified destination" mode. A "follow a person" behavior would allow the robot to automatically keep pace with the person physically present with it, thereby allowing the driver to dedicate his or her attention to the conversation and not the navigation.

One of the recommended guidelines listed above was to have an independent head or torso. However, while helpful for conversation, the additional degree of freedom would make controlling the robot more complex. Using autonomous behaviors for either navigation or to keep the robot's face oriented towards the walker (or both) could simplify the task and allow the driver to concentrate more on the conversation.

Depending on the size of the location in which the robots operate and where the robots are housed, there might be a significant amount of driving involved. Traversing a long

[7] Audio for one participant was not recorded.

[8] The VGo robot would switch access points during these trials, and therefore, we could not time it.

distance could potentially detract people from utilizing these robots if driving takes an order of magnitude more time than the alternatives (e.g., teleconferencing or video conferencing). A driver could instruct the robot to go to a specific meeting using a "go to destination" mode; while the robot is navigating there, the driver could use their time for last minute meeting preparations, e-mail, etc. If a scheduling system or calendar existed for sharing telepresence robots, people could schedule robots for specific time slots and locations. The robot would then navigate on its own to the location by the specified time, thereby saving the driver the effort to drive the robot from the charging station to the meeting area. When the driver has finished with using the robot, the robot could simply drive itself back to the charging station or to the location of the next scheduled appointment.

Adjustable autonomy: The mixture of low- and high-level autonomous behaviors for these robots will allow the drivers to switch to a supervisory role and perform other tasks while the robot autonomously navigates. There will be situations in which the driver might want to directly teleoperate the robot and the lower autonomy levels would allow the driver to do that. One participant suggested wanting to take control when they saw someone they wanted to talk to while the robot was autonomously navigating to the specified destination. In this scenario, the robot driver could teleoperate the robot into the person's office, and after the conversation, the robot could then resume its navigation. To allow for this flexibility, an adjustable autonomy or dynamic autonomy system must be implemented that allows the driver to select from a range of autonomous behaviors or levels.

F. Social considerations

The previous guidelines provide the technical and functional competence of a telepresence robot. Social acceptance will also be required for long-term acceptance [10]. We note in [6] that it will be important for bystanders to know if a telepresence robot is being used or not, and if it is, what the identity of the robot driver is. Both the QB and VGo robot have LEDs to indicate when the robot is in use.

Appropriate occupancy awareness: The QB and VGo robots have audible announcements for when a person is using the robot. The QB robot plays a jazzy tune when it is first driven after log in; the VGo robot uses text to speech to announce the name of the person who has started to use the robot and when the person has left the robot. These announcements were designed to provide awareness to the people around the robot that the robot was occupied by a person.

In a meeting scenario, we observed that these same announcements caused disruptions. The robot drivers of Teams 5 and 6 in Study 2 (remote user interactions) experienced disconnections from their robots during their meetings. For Team 5, the robot driver then logged back into the robot and put it into drive mode which caused the announcement song to play. The conversation had continued between the teammates in Mountain View, and, when the robot's song played, it startled the teammates and caused an awkward pause in the conversation. The scenario was similar for Team 6. However, there was an additional awkwardness when the robot (a VGo) lost connection because it made an announcement that the robot driver had left the meeting when the meeting had not yet ended. A few minutes later, the robot driver logged back in and his entrance was again announced.

IV. The Next Generation

The guidelines presented in this paper were derived from use cases for robots in a corporate office environment; however, we believe they apply to the majority of telepresence applications. Implementing the suggested guidelines to address the issues listed in the paper will help improve telepresence robots. The set of guidelines presented constitute the important features, but is not exhaustive.[9] The suggested guidelines must be implemented carefully so as to avoid creating additional problems. This particularly holds true while designing user interfaces and autonomous behaviors.

While the current version of telepresence robots are designed with very specific and limited scope, we firmly believe that telepresence robots are capable of providing true remote presence. We also believe that telepresence robots will also need to go through cycles of evaluation and redesign until they become a mature and stable technology. However, before telepresence robots are widely available, the specifications for these robots will need to be standardized. Standards pertaining to privacy and security must be established. Also, standardizing robot operating systems and communication protocols will help improve interoperability between robots. For large scale adoption, particularly in the corporate world, telepresence robots must also allow for integration into the existing infrastructure.

V. Acknowledgments

This research has been funded in part by NSF (IIS-0905228, IIS-0546309). We would like to thank Anybots and VGo Communications for loaning us prototype robots. The opinions expressed in this paper are that of the UMass Lowell researchers and not of Google.

References

[1] K. M. Tsui, M. Desai, H. A. Yanco, and C. Uhlik, "Exploring Use Cases for Telepresence Robots," in *Proceedings of the 6th ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2011.

[2] Qik, "Record and Share Video Live from Your Mobile Phone," Webpage, 2010, http://qik.com/ (accessed Oct. 2010).

[3] Joel Schneider, "Snellen Chart," Webpage, http://www.i-see.org/block_letter_eye_chart.pdf, accessed July. 2010.

[4] Department of Justice, "Americans with Disabilities Act," 1991, http://www.ada.gov/reg3a.html (accessed Oct. 2010).

[5] Google, "Google Goggles," Webpage, http://www.google.com/mobile/goggles/, accessed Jan. 2011.

[6] K. M. Tsui, M. Desai, H. A. Yanco, and C. Uhlik, "Telepresence Robots Roam the Halls of My Office," in *Human-Robot Interaction Workshop on Social Robotic Telepresence*.

[7] B. Keyes, M. Micire, J. Drury, and H. Yanco, "Improving Human-Robot Interaction through Interface Evolution," *Human-Robot Interaction, Daisuke Chugo (Ed.)*, pp. 183–202, 2010.

[8] M. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 37, no. 1, pp. 32–64, 1995.

[9] B. Gerkey, R. Vaughan, and A. Howard, "The Player/Stage Project: Tools for Multi-Robot and Distributed Sensor Systems," in *Proc. of Intl. Conf. on Advanced Robotics*, 2003, pp. 317–323.

[10] P. Salvini, C. Laschi, and P. Dario, "Design for Acceptability: Improving Robots Coexistence in Human Society," *International Journal of Social Robotics*, pp. 1–10, 2010.

[9] We expect the entire set of guidelines to be published in the near future.