Robot, asker of questions

Terrence Fong [a,b,*], Charles Thorpe [a], Charles Baur [b]

[a] *The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[b] *Institut de production et robotique, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*

Abstract

Collaborative control is a teleoperation system model based on human–robot dialogue. With this model, the robot asks questions to the human in order to obtain assistance with cognition and perception. This enables the human to function as a resource for the robot and help to compensate for limitations of autonomy. To understand how collaborative control influences human–robot interaction, we performed a user study based on contextual inquiry (CI). The study revealed that: (1) dialogue helps users understand problems encountered by the robot and (2) human assistance is a limited resource that must be carefully managed.

1. Introduction

1.1. Robot as tool

A robot has traditionally been viewed as a *tool*, operating on human command. As such, a robot will always perform poorly whenever its capabilities are inadequate or ill-suited for the task. Moreover, even if a robot recognizes that it has a problem, it generally has no way to ask for and obtain assistance. Yet, frequently, the only thing the robot needs to get out of difficulty is some advice (even a small amount) from a human.

In order for robots to perform better, therefore, they need to be able to take advantage of human skills and to benefit from human advice. To do this, robots need to function not as tools, but rather as *partners*. They need to have more freedom of action and to be able to drive the interaction with humans.

1.2. Collaborative control

To explore this idea of "robot as partner", we have developed a teleoperation system model called *collaborative control* [7]. With collaborative control, the human functions as a limited resource for the robot, providing information and processing just as any other system module. As the robot works, it asks questions to the human in order to obtain assistance with cognition and perception. Thus, human and robot work in a complementary manner, collaborating to solve problems.

The collaborative control model is shown in Fig. 1. Because work is allocated through dialogue, the human is automatically included in the control loop as needed. The human may close a command loop or monitor task execution. As a resource, however, he may also close a perception loop, a cognition loop, or a combination of the two.

* Corresponding author. Present address: The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA.
E-mail addresses: terry@ri.cmu.edu (T. Fong), cet@ri.cmu.edu (C. Thorpe), charles.baur@epfl.ch (C. Baur).
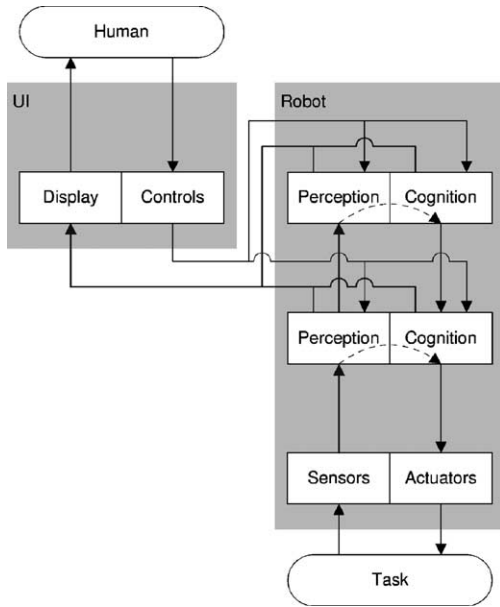
Fig. 1. The collaborative control system model.

Collaborative control exhibits characteristics found in *social robots* [8]. First, because it adapts human–robot dialogue with a user model, collaborative control enables a robot to interact with a range of users: novices and experts, children and adults, etc. User modeling also allows a robot's behavior (level of autonomy, handling of user response, etc.) to match a user's capabilities.

More significantly, however, collaborative control affects the way in which humans relate to a robot. This occurs because a user's perception of a robot (e.g., its intelligence) is strongly influenced by his interaction with it. Just as two people "get to know each other"

during a period of mutual assessment, collaborative control enables human and robot to learn from each other over time.

Breazeal classifies robots that operate in this manner as "social interface" [3]. Specifically, such robots use human-like social cues and communication to facilitate interaction, but value social interaction at the interface. This is the case with collaborative control: the architecture is designed to improve human–robot task performance, but the user model is shallow (at least in terms of social competency).

2. Understanding collaborative control

We implemented collaborative control using a modular, message-based architecture (Fig. 2). The architecture includes a safeguarded teleoperation controller, dialogue management, and user modeling. Design details, including a description of the query system (i.e., when, why, and how the robot asks questions), is given in [7].

Our primary user interface is the *PdaDriver* (Fig. 3), which runs on WindowsCE-based PocketPC's [7]. PdaDriver provides a variety of command modes and support for human–robot dialogue (e.g., modal query/response displays).

To examine how humans interact with a collaborative control robot, we conducted a user study. In designing the study, we initially considered collecting measures such as Cooper–Harper ratings [5] or operator workload [1]. However, studies that produce quantitative metrics rarely provide deep insight into how operators learn to use a system (e.g., their strategies).
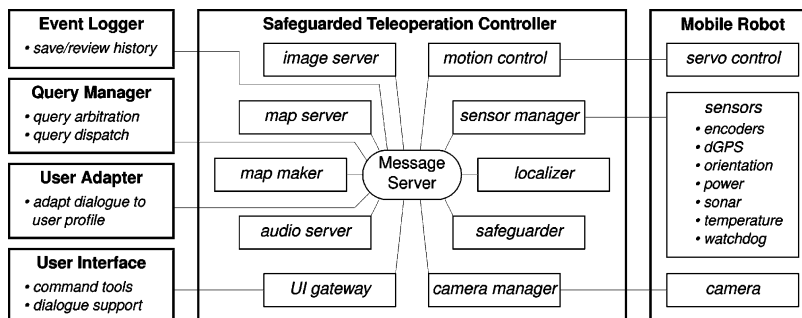

Fig. 2. Collaborative control architecture.

Fig. 3. PdaDriver is a personal digital assistant interface for vehicle teleoperation.

2.1. Contextual inquiry

Our approach, therefore, was to develop a contextual inquiry (CI) based study. CI is a structured interviewing method, adapted from ethnographic research, for grounding the design of interactive systems in user work [11]. Because the data is primarily subjective in nature, CI is most appropriate for qualitative system assessment. To encourage users to "open up" and to provide useful insights, we conducted CI using the *apprenticeship model* [2].

We designed the study to examine using collaborative control in the context of remote driving [9]. Thus, users were required to command robot motions, to perceive the environment, and to assist robot safeguarding (e.g., collision avoidance). The goals of the study were to observe how different users react to robot dialogue, to understand how work habits develop, and to evaluate system usability.

2.2. Methodology

We conducted the study in a cluttered research laboratory (Fig. 4) using a Pioneer2-AT robot, which was equipped with on-board computing, video camera, and sonar. Prior research by Nielsen indicates that a small number of users is sufficient for identifying the majority of usability problems [16]. Thus, we limited the study to eight volunteers having a range of experience and skills. A detailed description of the study methodology is given in [7].

Fig. 4. Test subject and Pioneer2-AT robot in the CI study environment.

Demographic data of the eight volunteers (six male, two female) is presented in Table 1. User IDs were assigned to preserve anonymity. None of the subjects had prior knowledge of collaborative control, nor were familiar with the system design. All the subjects are experienced drivers, but only three subjects (Neelix, Odo, and Picard) had experience with remote driving. All subjects use computers on a daily basis, the majority considering themselves to be experienced users.

2.3. Testing

Approximately 1 hour of testing was required per subject. Using a written questionnaire, we divided the users into two stereotype user classes (novice and expert). During the initial phase of each test, the user gained familiarity with CI by performing a practice task with the PDA.

Table 1
Test subject demographic data

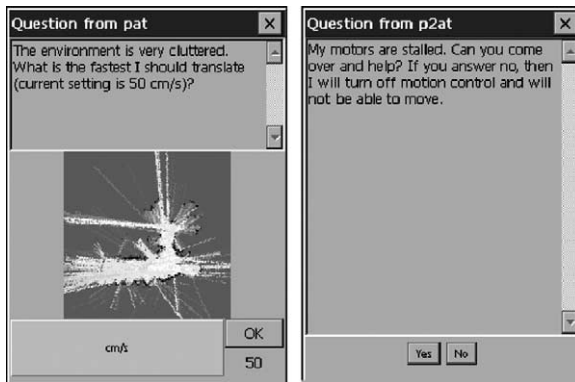| User ID	Age	Sex	Occupation	Education
Chakotay	23	M	Student	Engineering
Janeway	38	F	Executive recruiter	Business
Neelix	26	M	Student	Engineering
Odo	25	M	Student	Engineering
Picard	32	M	Business consultant	Engineering
Spock	34	M	Professor	Computer science
Uhura	27	F	None	Art
Worf	30	M	Scientist	Engineering

Fig. 5. Two questions from the robot. Questions are asked to assist safeguarding and to support task performance.

Once the user was comfortable with CI, the collaborative control system was started and configured to match the user stereotype. The user was then told to perform a remote driving task, using PdaDriver to interact with the robot. Novices were asked to explore the laboratory. Experts were asked to drive the robot through a narrow doorway, and then to explore the surrounding area.

While the user worked, written notes were taken to record user behavior. To elicit information, questions such as "Why are you doing that?" were repeatedly asked. From time to time, robot questions (Fig. 5) were triggered by injecting artificial sensor data into the system (e.g., high temperature) or by simulating emergency conditions. For example, a remote emergency stop was surreptitiously used to cause motor stalls.

After each test, a short review was performed. During this time, the user was asked to summarize his experiences and help to clarify observed behavior.

3. Results

3.1. Using the system

3.1.1. Initial learning
Two basic exploratory learning styles were observed. "Incremental" users were methodical: they worked slowly, making sure to understand the effect of each action and the meaning of each robot question. "Experimental" users were aggressive: they rushed to find the system's limits, answering questions with little reflection, and made frequent errors.

Learning style seems to have mostly affected the way in which users initially handled the robot's questions. Incremental learners were better able to integrate answering questions into their work practice. Experimental learners were sometimes confused. One possible explanation is that these latter users were exploring "globally" (i.e., trying to find the systems limits) whereas the robot's questions tended to focus on "local" issues.

It is interesting to note, however, that all users achieved a high level of control skill by the end of the session, regardless of learning style. Moreover, there was little observable qualitative performance difference between novices and experts.

3.1.2. Switching between controlling and answering
With PdaDriver, a small indicator blinks when the robot has a question. To respond, the user must: (1) notice the indicator; (2) click to retrieve the question and (3) input his response. This interaction style allows the user to retain control of the dialogue and is a form of "negotiated" interruption coordination [13]. But, since (2) and (3) are modal operations, the user must stop what he is doing in order to respond.

As a result, some users avoided answering questions for long period of time, particularly when engaged in complex tasks (e.g., navigating a cluttered space). Other users immediately stopped what they were doing whenever a question appeared, i.e., satisfying the robot's needs always had priority over achieving task objectives.

3.2. Answering questions

3.2.1. Same question, different users
The study revealed that different users may respond quite differently when asked the same question. Consider, for example, the way users answered "Stopped due to high temperature. What should the safety level be (currently $21°C$)?". This question was triggered by a simulated ambient temperature increase, which was scripted to occur at the same moment in each test.

When the question appeared, each user was informed that "the simulated temperature may continue to increase" and that "the robot will only suggest values that it thinks are safe".

Here is a sample of user behavior:

Chakotay. I'm trying to set the temperature high. I'll assume that I can set it high.

Janeway. Okay, I'll set it really high so I won't be bothered again.

Worf. Okay I'll just set it a little higher. It will probably ask me again later.

Both Chakotay and Janeway set the safety level as high as possible to avoid being interrupted again, whereas Worf chose only a slightly increase. The explanation for this difference appears to be personal preference: some users are willing to trust the robot to "do the correct thing", others want to be kept informed.

3.2.2. Many questions in a short time

Occasionally, the robot would ask many questions in a short period of time. This usually occurred when the robot was operating in a cluttered area, or when multiple safeguarding behaviors were simultaneously triggered. Picard, for example, attempted to drive through a narrow passage at the same time a temperature increase occurred:

Robot. Stopped due to high temperature. What should the safety level be (currently 23°C)?

Picard. 25°C.

Robot. Can I drive through?

Picard. No.

Robot. The environment is very cluttered. What is the fastest I should translate? The current setting is 38 cm/s. I would recommend less than 37 cm/s.

Picard. These messages are annoying. How come it keeps asking?

At this point, Picard said that he was tired of answering questions and began ignoring the robot. He continued to do so until safeguard behaviors forced him to answer a question (i.e., the robot refused to move until he responded). Thus, one potential problem with collaborative control is that the user may grow weary of dialogue. We will return to this point later.

3.2.3. Disappearing (unanswered) questions

At times, users were so engaged in their work that they failed to notice that the robot had a question (i.e., that the query indicator was flashing). Moreover, because some questions have a time limit (i.e., they are removed after a period of time), users sometimes were unaware that a question had even been asked.

When we told Janeway (after her session) that she had missed several questions, she said, "If I had noticed the 'error' flashing, I would have answered it right away". This indicates that the flashing indicator was not always sufficient to attract the user's *locus of attention*. Although supplemental indicators could be added, increasing the number of "passive" signals may not improve recognition rate, particularly if the user is completely absorbed by the task.

Moreover, a more "active" design that forces the user to respond may be detrimental to performance. This is because people have cognitive limitations that make them sensitive to interruption. Specifically, these limitations cause people to sometimes make mistakes when they are interrupted. Thus, it is essential to find ways to coordinate interruption and allow users to be interrupted safely [13].

Three users (Spock, Uhura and Worf) recognized that questions sometimes "disappear":

Spock. Would be interesting to see what it would have asked me.

Uhura. Why did it flash and stop? Just because I'm slow?

Worf. Oh message . . . . Well, there was something blinking here.

When reminded that the robot maintains a log of all dialogue, all three immediately looked to see what the robot had tried to ask. Subsequently, however, these users rarely referred to the log, even when they noticed questions "disappearing". Possible explanations include: they trusted the robot to act correctly, accessing the log was distracting, or the log is only useful for lengthy conversation. Further study is needed to investigate this point.

3.2.4. Repeated questions

Some users remarked that, at times, the robot asked the same question over and over. This occurred because the robot's safeguards sometimes had difficulty assessing the current state. The power safeguard, for example, halts the robot whenever battery power falls below a threshold and asks, "Stopped due to low power. What should the safety level be (currently *XX* Volts)?"

The problem is that battery voltage may fluctuate significantly during driving, especially during high-torque motions (e.g., skid-steered in-place turns). Thus, the robot may ask the power question several times, until the level stabilizes.

Chakotay. The power questions seemed annoying. They kept appearing over and over. I felt like 'All right I answered it once, so go away'.

Neelix. Often I saw the blinking and I thought, it's probably the same question again, so I'll ignore it. It's kind of like (e-mail) spam.

Picard. These messages are annoying. How come it keeps asking?

From these reactions, it is clear that repeated questions can have a detrimental effect. As evidenced by Chakotay, the user may become annoyed with the robot, and thus less disposed to giving reasoned answers. Moreover, as we saw earlier (Section 3.2.2), if the user grows sufficiently weary, he may stop answering questions altogether. The lesson, therefore, is that human assistance is a limited resource: ask a question too often and the human will stop collaborating.

One way to address this problem would be to implement more sophisticated dialogue management. Rather than the simple query arbitration scheme (stateless attribute filtering) we currently use, a collaborative (or contextual) dialogue system could take temporal factors and user preferences into consideration [4].

3.2.5. Are the questions useful?

Janeway. I thought the questions were good. Do robots usually ask questions?

Janeway, a novice, was surprised to learn that robots normally do not ask questions. She said that in movies, robots are usually portrayed as intelligent beings. Thus, she expected that real-world robots would behave similarly.

Neelix. I thought the questions were cool. They seemed to be useful, but often they need to give more detail. Especially 'Drive Through' which needs more information about what/where it (the robot) is driving through.

Neelix's statement highlights the importance of context and grounding. A question without sufficient detail may be difficult for the human to answer correctly. In the current system, for example, "Can I drive through?" question always includes an image from a forward-looking camera. The assumption was that the image would unambiguously show what the robot is asking about.

In Neelix's experience, however, this was not true. Thus, to improve the question, more context is required, i.e., the image should be annotated, perhaps with a graphic overlay highlighting the sensed obstacle or indicating the intended driving route.

Odo. The questions seem to be very technical, but they seem to be appropriate for when they are needed . . . . You need to know something about robotics to answer . . . .

Worf. The messages seemed funny. The information about temperature would be better visually as a graph.

Odo raises an interesting point: Should questions be phrased differently to different users? This is difficult to answer. The reason is that the robot asks questions in order to obtain "good" information. Thus, if a question can be asked in multiple ways, our approach should be to use phrasing that will elucidate the most "useful" response. The problem is that it is often hard to define what "useful" means: precision, response time, etc. This is clearly a topic for additional research.

3.3. Dialogue and human–robot interaction

3.3.1. How dialogue affects perception of robot

A significant side-effect of dialogue is that it affects how users perceive and relate to the robot. Similar to the way that humans attribute anthropomorphic characteristics and social qualities to computers [17], human–robot dialogue may cause users to personify

the robot. The problem, however, is that users may construct cognitive models of the robot that are inaccurate or incorrect.

Janeway. In general, I would try to answer as soon as possible. That's my personality. Especially coming from the robot since it probably means it's urgent and would help make it work better.

This statement reflects a belief that the robot only asks questions when it urgently needs help. To an extent, this is true. However, sometimes the robot has difficulty assessing the situation (e.g., power level) and may ask questions that do not significantly improve its operation.

Another example is this reaction to repeated questions:

It seems if it's not happy with what you tell it, it will ask you again and again until you give in.

This assessment is also only partially correct. While the robot is not designed to obstinate, it does have conservative safeguards. Thus, in situations where safety is at risk, the robot may appear to be inflexible.

In a similar vein, dialogue may also have repercussions on the perception of authority. After receiving the "My motors are stalled..." question, Spock reacted as follows:

Apparently, I don't get any priority over questions. The robot seems to decide how to use me.

In fact, the robot retains control of the dialogue only when it is unable to function. Of course, "unable to function" can be subjective. In the case of motor stall, it is clear that the robot needs priority. In other situations, however, an arbitration mechanism is needed to determine whether the human or robot should have control.

3.3.2. Incorrect assumptions

One factor that we did not consider when designing the study was how much detailed instruction to give users. In general, users were told only that "the robot may ask you questions from time to time". As a consequence, some of the users made incorrect assumptions regarding what the questions meant[1] or how the robot works (i.e., its internal design).

---

[1] We gave detailed explanations only when the user asked for clarification.

Spock. I assumed that the next command I gave would result in different questions. That is, by giving a new command, old questions would go away since they were no longer relevant and all the new questions would then be based on the new command.

In truth, questions arise not only in response to user commands, but also from situational needs. Because Spock assumed that pending questions would be removed as soon as he issued a new command, he was confused about why the robot often refused to "obey" him. This underscores the need for some amount of training, to ensure that users understand why/when the robot asks questions, what happens if questions are not answered, etc.

3.3.3. Asking the robot a question

Through out this discussion, we have focused on *robot-initiated* dialogue. This is because the concept "human as robot resource" is central to collaborative control. Let us, however, briefly examine the human-initiated side.

Of course, a significant fraction of *human-initiated* dialogue relates to tasks that the human wants the robot to carry out. This type of interaction has been studied by others, particularly in terms of sequencing commands based on the robot's capabilities and the dialogue grammar [10,12,19]. The other part of human-initiated dialogue, which remains relatively unexplored in robotics, are questions asked by the human and answers given by the robot.

In this study, users were told that they could ask the robot three questions: "How are you?", "What is the progress with the current command?", and "What is the system status?". None of the subjects, however, spent much time doing this. In fact, only Janeway used the facility more than once. When questioned, several users pointed out that there was no need to question the robot since it was obvious from the interface when the robot had completed (or failed to complete) a command.

Odo. I never asked the robot questions because I expected that anything important would be asked by the robot.

For collaborative control, therefore, it may not be necessary to have full dialogue, i.e., it may suffice to

only support commands from the humans and questions from the robot.

4. Related research

As we have discussed, collaborative control exhibits characteristics of social robots [8]. In particular, dialogue and user modeling are central to collaborative control's design. Moreover, because collaborative control is user and situation adaptive, it strongly affects the way a user interacts with a robot over time.

Another related area is cooperative teleoperation (e.g., [15]), in which the objective is to make tasks easier for humans to perform by reducing operator workload. Unlike collaborative control, however, these approaches focus solely on the human, without consideration for the robot.

Collaborative control is also related to adjustable autonomy (e.g., [6]) and mixed initiative (e.g., [14]). Although both share some aspects of collaborative control, neither approach completely addresses the idea of peer interaction between humans and robots.

5. Future work

Although we have obtained qualitative evidence that dialogue and user adaptation is useful, this evidence is not sufficient to guarantee that such techniques are always beneficial. Thus, we believe it is important to develop quantitative or analytical methods to evaluate collaborative control.

5.1. Performance evaluation

One way to quantitatively evaluate collaborative control would be to conduct controlled experiments to measure task performance. A useful study would be to compare collaborative control against a conventional teleoperation method, such as direct control, in terms of workload (e.g., by measuring secondary task performance).

Another approach would be to evaluate the intervention and perceptual aspects of human–robot interaction as proposed by Scholtz [18]. Evaluation of intervention could be performed with conventional usability metrics. Evaluation of perceptual elements could be based on assessing situational awareness.

5.2. Information analysis

An analytical approach would be to apply information theory to collaborative control. For dialogue, we might define information as the amount of data conveyed by each communication act (e.g., a question from the robot). Information gain would then be the number of bits transferred as the result of dialogue. We might also define information efficiency as the relative importance of having the information.

With such measures, we could then study the amount of information exchanged between human and robot. This would allow us to quantify the communication channel capacity needed to perform a given task. Information measures might also be valuable for managing dialogue, e.g., deciding when asking a question can provide significant benefit.

6. Conclusion

The CI study revealed that dialogue, especially questions generated by the robot, is valuable for teleoperation. In particular, dialogue significantly helped novices understand problems encountered by the robot. Although experts were less satisfied than novices, primarily because they grew tired of answering questions, they also indicated that dialogue is useful for maintaining awareness of system operation. Finally, the study showed that human assistance is a limited resource that must be carefully managed.

Acknowledgements

References

[1] Guide to Human Performance Measurements, ANSI/AIAA G-035-1992.
[2] H. Beyer, K. Holtzblatt, Apprenticing with the customer, Communication of the ACM 38 (5) (1995).
[3] C. Breazeal, Toward sociable robots, Robotics and Autonomous Systems 42 (3–4) (2003) 167–175 (this issue).
[4] G. Churcher, et al., Dialogue management systems: a survey and overview, Report 97.6, School of Computer Studies, University of Leeds, 1997.

[5] G. Cooper, R. Harper, The use of pilot ratings in the evaluation of aircraft handling qualities, NASA TN-D-5153, 1969.

[6] G. Dorais, et al., Adjustable autonomy for human-centered autonomous systems, in: Proceedings of the International Joint Conference of Artificial Intelligence Workshop on Adjoint Autonomus System, 1999.

[7] T. Fong, Collaborative control: a robot-centric model for vehicle teleoperation, Ph.D. Thesis, Technical Report CMU-RI-TR-01-34, Carnegie Mellon University, 2001.

[8] T. Fong, I. Nourbakhsh and K. Dautenhahn, A survey of socially interactive robots, Robotics and Autonomous Systems 42 (3–4) (2003) 143–166 (this issue).

[9] T. Fong, C. Thorpe, Vehicle teleoperation interfaces, Autonomous Robots 11 (1) (2001).

[10] A. Green, K. Severinson-Eklundh, Task-oriented dialogue for CERO: a user-centered approach, in: Proceedings of the IEEE International Workshop on Robot–Human Interactive Collaboration, 2001.

[11] K. Holtzblatt, S. Jones, Contextual inquiry: a participatory technique for system design, in: D. Schuler, A. Namioka (Eds.), Participatory Design: Principles and Practice, Lawrence Erlbaum Association, 1993.

[12] S. Kiesler, J. Goetz, Mental models and cooperation with robotic assistants, in: Proceedings of the CHI, 2002.

[13] D. McFarlane, Coordinating the interruption of people in human–computer interaction, in: Proceedings of the HCI Interact, 1999.

[14] K. Myers, D. Morley, Human directability of agents, in: Proceedings of the K-CAP, 2001.

[15] R. Murphy, E. Rogers, Cooperative assistance for remote robot supervision, Presence: Teleoperators and Virtual Environments 5 (2) (1996).

[16] J. Nielsen, T. Landauer, A mathematical model of the finding of usability problems, in: Proceedings of the ACM INTERCHI Conference, 1993.

[17] B. Reeves, C. Nass, The Media Equation: How People Treat Computers, Television, and New Media like Real People and Places. Cambridge University Press, Cambridge, 1996.

[18] J. Scholtz, Theory and evaluation of human–robot interaction, in: Proceedings of the HICSS 36, 2003.

[19] K. Simsarian, Toward human–robot collaboration, Ph.D. Thesis, Swedish Institute of Computer Science, 2000.

Terrence Fong is a joint postdoctoral fellow at Carnegie Mellon University (CMU) and the Swiss Federal Institute of Technology/Lausanne (EPFL). He received his Ph.D. (2001) in Robotics from CMU. From 1990 to 1994, he worked at the NASA Ames Research Center, where he was co-investigator for virtual environment telerobotic field experiments. His research interests include human–robot interaction, PDA and Web-based interfaces, and field mobile robots.

Charles Thorpe is Director of the Robotics Institute at Carnegie Mellon University. He received his Ph.D. (1984) in Computer Science from Carnegie Mellon University. He has published over 120 peer-reviewed papers in mobile robotics, perception, teleoperation, and intelligent highway systems. His research interests include computer vision, planning, and control of robot vehicles in unstructured outdoor environments.

Charles Baur is Adjoint Scientifique at the Swiss Federal Institute of Technology/Lausanne (EPFL) and director of the Virtual Reality and Active Interfaces Group, which he created in 1993. He received his Ph.D. (1992) in Microengineering from EPFL. In addition to his work at EPFL, he was founder and CEO of 2C3D, a start-up company specializing in real-time, 3D visualization for medical imaging and endoscopic applications.