

Multi-Robot Remote Driving With Collaborative Control

Terrence Fong, Charles Thorpe, and Charles Baur

Abstract—Multi-robot remote driving has traditionally been a difficult problem. Whenever an operator is forced to divide his limited resources (attention, cognition, etc.) among multiple robots, control becomes complicated and performance deteriorates as a result. Thus, we need to find ways to make command generation and coordination efficient, so that human–robot interaction becomes transparent and tasks are easy to perform. In this paper, we discuss how human–robot collaboration and dialogue provide an effective framework for achieving this.

Index Terms—Collaborative control, human–robot interaction, mobile robots, remote driving, vehicle teleoperation.

I. INTRODUCTION

A. Multi-Robot Remote Driving

The American military is developing mobile robots to support future combat systems. These robots will be used to perform reconnaissance, surveillance, and target acquisition. Because such tasks have traditionally required significant human resources and risk taking, a primary area of interest is determining how to reduce these factors by having a single operator control multiple robots.

Vehicle teleoperation, however, is not easy to perform. With manual control, performance is limited by the operator's motor skills and his ability to maintain situational awareness. Additionally, humans have difficulty building mental models of remote environments. Distance estimation and obstacle detection can also be difficult [1].

With multiple robots, the difficulty is further increased because the human must divide his limited cognitive and sensorimotor resources. In general, whenever a single operator controls multiple robots, his workload becomes higher and task performance decreases. Moreover, if the robots are deployed in a dynamic setting, the operator often has to spend significant time resituating himself each time he switches context.

Manuscript received December 25, 2001; revised September 5, 2002. Abstract published on the Internet May 26, 2003. This work was supported in part by grants from the DARPA ITO MARS program, the National Science Foundation, and SAIC. This paper was presented in part at the 2001 IEEE International Workshop on Robot and Human Interactive Communication, Bordeaux and Paris, France, September 18–21.

T. Fong was with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA. He is now with the Institut de production et robotique, Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland (e-mail: terrence.fong@epfl.ch).

C. Thorpe is with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: cet@cs.cmu.edu).

C. Baur is with the Institut de production et robotique, Swiss Federal Institute of Technology (EPFL), CH-1015 Lausanne, Switzerland (e-mail: Charles.Baur@epfl.ch).

Digital Object Identifier 10.1109/TIE.2003.814768

Thus, for multi-robot remote driving to be productive, we need to make it easy for the operator to understand the remote environment, to assess the situation, and to generate commands. Our approach is to create techniques and tools that improve human–robot interaction [2], [3]. In particular, we have developed a new system model for teleoperation called *collaborative control*.

B. Collaborative Control

We are convinced that there are clear benefits to be gained from humans and robots working together. Specifically, we believe that in order for robots to perform better, they need to be able to take advantage of human skills (perception, cognition, etc.) and to benefit from human advice and expertise. To do this, robots need to function not as passive tools, but rather as active partners.

To this end, we have developed *collaborative control*, a system model in which human and robot collaborate to perform tasks and to achieve common goals [3]. Instead of a supervisor dictating to a subordinate, the human and the robot engage in *dialogue* to exchange information, to ask questions, and to resolve differences. With this approach, the robot has more freedom in execution and is more likely to find good solutions when it has problems.

With collaborative control, the human is able to function as a resource for the robot, providing information and processing just like other system modules. Specifically, the robot can ask questions to the human as it works, to obtain assistance with cognition and perception during task execution. In other words, human and robot collaborate in order to compensate for limitations of autonomy.

Collaborative control is both a novel and a useful paradigm for teleoperation. Collaborative control is novel because it uses dialogue as a framework for coordination, to direct joint task performance and to focus attention where it is needed. Collaborative control is useful because it provides an efficient mechanism for adaptation, to adjust autonomy and human–robot interaction to fit situational needs and user capabilities.

II. RELATED WORK

A. Human–Robot Collaboration

Humans and robots have been working together since the 1940s. At first, human–robot interaction was primarily unidirectional: simple switches or controls for operating manipulator joints and remote vehicles. However, as robots have become more autonomous, this relationship has changed to be more like

the relationship between two human beings. As a result, humans and robots now communicate and collaborate in a multitude of ways [4].

Personal service robots, for example, directly assist people in daily living activities. Baltus *et al.* discuss the development of mobile robots that provide a range of caretaking services such as patient monitoring and medical data collection [5]. Green and Severinson-Eklundh present a fetch-and-carry robot which assists physically impaired office workers [6]. Nourbakhsh *et al.* describe Sage, an educational mobile robot that gives museum tours [7].

Additionally, some researchers have begun studying how humans and robots can function as a unit, jointly participating in planning and problem solving. Laengle, Hoenger, and Zhu discuss human and robot working in teams [8]. Bonasso addresses the use of mixed-initiative and adjustable autonomy between humans and robots [9].

B. Robot Control Architectures

Although most robot control architectures are designed for autonomy, some have addressed the problem of mixing humans with robots. One approach is to directly incorporate humans into the design as a system element. DAMN, for example, is a behavior-based architecture in which individual modules vote on possible actions [10]. Command arbitration allows modules as disparate as autonomous safety behaviors and teleoperation to coexist.

Another approach is the use of prioritized control, in which operator commands may be overridden by autonomous modules. The best-known example of this is NASREM, which explicitly incorporated an operator interface into a layered, hierarchical control system [11]. More recently, the concept of safeguarded teleoperation has been used to enable novices to teleoperate a planetary rover [12].

In all of these approaches, the assumption is that the human will only provide control (command) input to the system. Collaborative control, however, is not limited by this narrow assumption. Instead, it also allows the human to contribute high-level planning or perception input to robot modules.

III. APPROACH

During the past year, we have developed a collaborative control system that includes human-robot dialogue management and a personal user interface [3]. We are using this system to teleoperate multiple, mobile robots in unknown, unstructured terrain. At present, we are using a Pioneer-AT and a Pioneer2-AT, both of which are skid-steered vehicles equipped with microprocessor-based servo controller, on-board computing and a variety of sensors.

A. Dialogue

Dialogue is the process of communication between two or more parties. Dialogue is a joint process: it requires sharing of information (data, symbols, context) and of control. Depending on the situation (task, environment, etc.), the form or style of dialogue will vary. However, studies of human conversation have

TABLE I
VEHICLE TELEOPERATION DIALOGUE

	Category	Messages
user → robot	robot command (<i>command for the robot</i>)	position (<i>pose, path</i>) rate (<i>translate, rotate</i>) stop camera pose (<i>pan, tilt, zoom</i>) camera config (<i>exposure, iris</i>) sonar config (<i>polling sequence</i>)
	query-to-robot (<i>question from the user</i>)	How are you? Command progress?
	response-from-user (<i>query-to-user response</i>)	y/n value
robot → user	info statement (<i>information for the user</i>)	pose (<i>x, y, z, roll, pitch, yaw</i>) rates (<i>translate, rotate</i>) message (<i>event, status, query</i>) camera state (<i>pan, tilt, zoom</i>) get new image
	query-to-user (<i>question from the robot</i>)	Can I drive through (<i>image</i>)? Is this a rock (<i>image</i>)? If you answer 'y', I will stay here. [exploration] The environment is very cluttered (<i>map</i>). What is the fastest I should translate? My motors are stalled. Can you come help? Motion detected. Is this an intruder? If you answer 'y', I will follow him [surveillance] Motion control is currently turned off. Shall I enable it? Safeguards are currently turned off. Shall I enable it? Stopped due to collision danger. Disable safeguards? Stopped due to high temperature. What should the safety level be? Stopped due to low power. What should the safety level be? Stopped due to rollover danger. Can you come over and help?
	response-from-robot (<i>query-to-robot response</i>)	How are you? → bargraphs (<i>health, rollover, collision</i>) Command progress? → stripchart (<i>progress over time</i>)

revealed that many properties of dialogue (e.g., initiative taking) are always present.

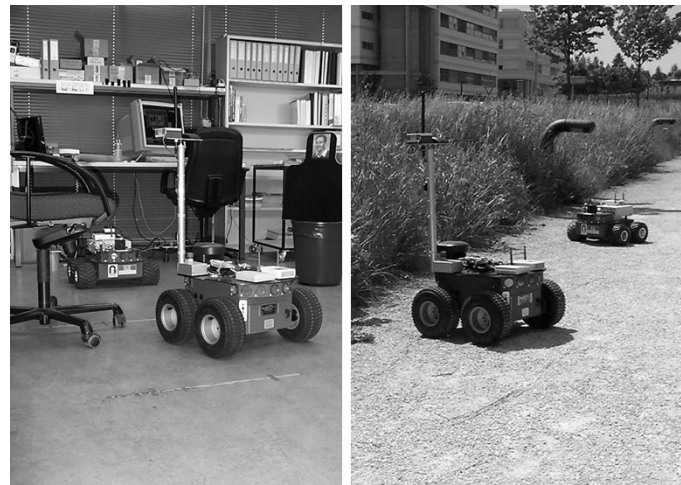
In our system, dialogue arises from an exchange of messages between human and robot. Effective human-robot dialogue does not require a full language, merely one which is pertinent to the task at hand and which efficiently conveys information. Thus, we do not use natural language and we limit message content to vehicle mobility (navigation, obstacle avoidance, etc) and task-specific issues.

At present, we are using approximately 30 messages to support vehicle teleoperation (Table I). Robot commands and information statements are unidirectional. A query (to either the human or the robot) is expected to elicit a response, though the response is not guaranteed and may be delayed.

To manage human-robot dialogue, we have developed a number of modules that perform user modeling, query arbitration, and message management. The *UserModeller* uses predefined, stereotype user profiles to adapt human-robot



Fig. 1. PdaDriver and control modes. (a) Rate/Position. (b) Image waypoint. (c) Map waypoint. (d) Sensor.



(a) (b)

Fig. 2. Multi-robot remote driving. (a) Indoor surveillance. (b) Outdoor reconnaissance.

dialogue to a specific user. The *QueryManager* performs query arbitration and dispatch using an attribute-filtering scheme. The *EventLogger* records system events as they occur and provides historical playback. The design of each of these modules is described in [3].

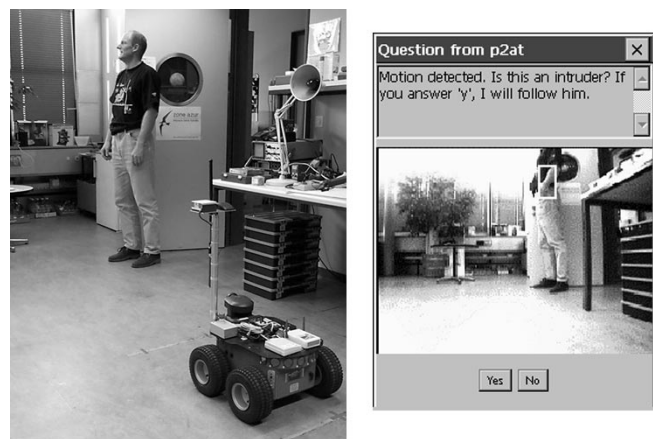
B. User Interface

Our current user interface (Fig. 1) is the PdaDriver [2]. We designed the PdaDriver to provide multiple command modes, to enable human-to-robot dialogue (i.e., the robot can query the user through and to support human-to-human interaction (audio and video). PdaDriver supports simultaneous (independent) control of multiple mobile robots and runs on WindowsCE Palm-size PCs such as the Casio Cassiopeia. The current version is implemented using Personal Java, a Java application environment designed for personal consumer devices.

Remote driving in unstructured, unknown environments requires flexible control. Because both the task and the environment may vary (depending on situation, over time, etc.), no single command-generation method is optimal for all conditions. For example, cross-country navigation and precision maneuvering have considerably different characteristics. Thus, PdaDriver provides a variety of command modes including direct rate/position control, image-based waypoint, and map-based waypoint [3].

IV. REMOTE DRIVING TESTS

We recently conducted two remote driving tests in which a single operator was responsible for controlling two mobile robots. These tests were designed to study basic surveillance and reconnaissance tasks that would normally be performed by humans, but which could fairly easily be performed by mobile robots. The first test examined an indoor surveillance mission and the second test looked at an outdoor reconnaissance mission (Fig. 2).



(a) (b)

Fig. 3. Intruder detection.

A. Indoor Surveillance

The first test required two primary tasks to be completed. Because the environment was unknown, the operator was instructed to begin by exploring the area. Once the operator had become sufficiently familiar with the area, he was then asked to use the robots to watch for intruders entering the area. To assist the human in detecting intruders, each robot was equipped with a motion detection module. This module detects motion by acquiring camera images and computing interframe differences whenever the robot is stationary. If the robot detects a moving object, it notifies the human and asks what to do.

Fig. 3(a) shows an intruder walking into the room. A robot, which had been positioned near the door, detected this motion and generated a question for the human: “Motion detected. Is this an intruder? If you answer ‘y’, I will follow him.” This question was presented to the user with an image of the motion area (marked by a bounding box), as shown in Fig. 3(b). At this point, the human became responsible for deciding if there really

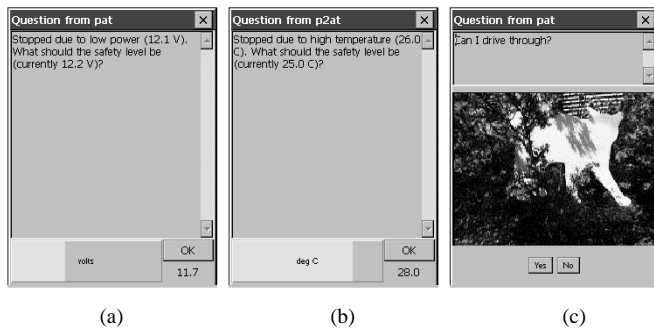


Fig. 4. Questions from the robot during outdoor reconnaissance.

was an intruder present in the scene, and if so, whether the robot should proceed to follow (track) him.

Although, in this case, the robot asked the human for confirmation before proceeding, it does not necessarily have to. If the human does not respond within some period of time, the intelligent decision for the robot might be to take the initiative and track the potential intruder until the human says to stop. If the robot had other capabilities, such as recognition (e.g., people identification) or intruder countermeasures, then further dialogue could be used to select the appropriate action to take.

B. Outdoor Reconnaissance

In the second test, the operator remotely drove the two robots through an unfamiliar outdoor environment. The objective for this test was to perform reconnaissance in the presence of moving and static hazards. Because the environment had not been previously explored, the operator was forced to rely on waypoint driving and on-robot safeguarding to conduct the task.

One of the key features of collaborative control is that it uses *query arbitration* to choose which questions to ask based on both immediate (local) needs and overall (global) strategy. This is particularly important when the robot is operating in an unknown or hazardous environment, i.e., to ensure that the most critical question (in terms of time, priority, etc.) is always asked first. In this test, for example, simultaneous questions from safety behaviors (on both robots) were arbitrated to direct the human's attention to where it was most urgently needed.

Fig. 4 shows some examples of the questions asked by the robots during the test. For example, both the Pioneer-AT ("pat") and the Pioneer2-AT ("p2at") determined that safety levels had been reached [Fig. 4(a) and (b)]. Although it is normal engineering practice to specify a safety level for normal operations, there are some occasions when a system must be used beyond its design specifications. This is particularly true for situations in which system loss may be acceptable if a goal is achieved (e.g., military combat missions).

V. USER STUDY

To better understand how collaborative control influences human-robot interaction, we performed a contextual-inquiry (CI)-based user study. CI is a structured interviewing method, adapted from ethnographic research, for grounding the design of interactive systems in the context of the work being performed [13]. Because the data it provides are primarily



Fig. 5. Test subject in the CI study environment.

subjective in nature, CI is most appropriate for qualitative system assessment, rather than for performance measurement.

We designed our study to evaluate collaborative control in the context of remote driving. Thus, users were required to perceive the environment, to generate commands, and to assist safeguarding autonomy. Our objectives were to observe system use (especially how users learn to use collaborative control) and to gain insight into how dialogue affects human-robot interaction (e.g., how do users relate to a robot that asks questions?).

We conducted the CI study in a cluttered, indoor environment (a research laboratory) at the Swiss Federal Institute of Technology, Lausanne, Switzerland (Fig. 5). A total of eight volunteers (six male, two female) participated. Based on data from a background questionnaire, we divided the users into two stereotype user classes (novice and expert) and asked each to explore the environment using a mobile robot and collaborative control.

A. Using the System

1) *Initial Learning*: Because none of the users had prior experience with collaborative control, they each spent several minutes learning how to use the system. Two basic exploratory learning styles were employed. "Incremental" users were very methodical in learning the system. They worked slowly, making sure to fully understand the effect of each control and carefully considering each robot question. "Experimental" users were very aggressive as they worked. These users rushed to find the system's limits, quickly answered questions, and made frequent errors.

Learning style appears to have mostly affected how users initially handled the robot's questions. Incremental learners, because they worked slowly and methodically, were better able to integrate answering questions into their work practice. Experimental learners, on the other hand, were sometimes confused by the robot's questions. One possible explanation is that these users were exploring "globally" (i.e., trying to find the systems limits) whereas the questions were focused on "local" issues.

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

2) *Switching Between Controlling and Answering:* With the current system, an indicator flashes on the PdaDriver display whenever the robot has a question to ask. To respond, the user must: 1) notice that the indicator is blinking; 2) press the indicator to retrieve the question; and 3) input and submit his response.

This style of interaction allows the user to retain control of the dialogue. However, since steps 2) and 3) are blocking operations (i.e., the interface prevents the user from doing anything else), it also means that the user must stop controlling the robot in order to respond to its questions. As a result, some users avoided answering questions for long periods of time, particularly when they were engaged in a complex task (e.g., precision driving).

An interesting consequence of switching between controlling and answering was that two users developed a precision driving technique. This occurred because they noticed that the robot asks “Can I drive through?” whenever it finds itself in a tight spot. Thus, their approach was to command the robot through a narrow opening (e.g., a door), wait for the question, and then answer yes. The other users, who did not discover this strategy, were unable to maneuver through tight passages (i.e., autonomy was inadequate for the robot to perform such maneuvers unassisted).

B. Dialogue and Human–Robot Interaction

1) *How Dialogue Affects Perception of Robot:* A significant side-effect of dialogue is that it directly affects how users perceive and treat the robot. In particular, it may cause users to personify the robot and to attribute qualities to it, which do not have a basis in reality. For example, consider this comment:

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. Under collaborative control, the robot is supposed to have sufficient awareness to recognize when human assistance would be beneficial. However, sometimes the robot has difficulty assessing the situation 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.

As in the prior case, this assessment is only partially accurate. While the robot is not explicitly designed to be obstinate, its safeguards are intended to keep the robot away from harm. Thus, in situations where its safety is at risk, such as driving fast in a cluttered area, the robot may appear to be inflexible.

2) *Asking the Robot a Question:* A large part of our research has focused on robot-initiated dialogue: questions asked by the robot and answers given by the human. This is because the concept of “human as resource for the robot” is central to collabo-

orative control. However, since good dialogue is two-way, let us briefly examine the human-initiated side.

Of course, a significant fraction of human-initiated dialogue relates to commands, tasks, and directives the human wants the robot to carry out. This type of human–robot dialogue has been studied by others, particularly in terms of interpreting and sequencing commands based on the robot’s capabilities and the dialogue grammar [6]. 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. None of the subjects, however, dedicated much time to doing this. When questioned, many of the users said that the questions did not seem very pertinent, or helpful, to the tasks at hand. Several users also pointed out that there was no need to ask the robot “What is the progress with the current command?”, since it was obvious from the displays when the robot had completed (or failed) executing a command. The most interesting comment was:

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

In terms of collaborative control system design, therefore, it may not be necessary to have full dialogue, but only to provide support for the human to issue commands and the robot to ask questions.

VI. DISCUSSION

A. Human–Robot Collaboration

We have found that there are two key factors for achieving effective human–robot collaboration. First, roles and responsibilities must be assigned according to the capabilities of both the human and the robot. Although this might seem easy to do, in practice it is not. In particular, many vehicle teleoperation tasks, such as identifying obstacles in an unfamiliar environment, can be highly situation dependent. Thus, even if the robot has previously accomplished a task by itself, it may not be able to the next time without some amount of human assistance.

In the case of multi-robot remote driving by a single operator, we need to reduce, as much as possible, the level of attention and control the operator must dedicate to each robot. This is true whether the human controls the robots individually or as a group (e.g., in formation). Moreover, even if one or more robots work together (i.e., robot–robot collaboration), we must still find ways to direct the human’s attention to where it is needed, so that he can help solve problems.

One way to achieve this is for the human to focus on global strategy (e.g., where to go) and to allow the robots to handle the low-level details (i.e., how to get there safely). Then, whenever a robot completes a task or encounters a problem, it notifies the operator. If multiple robots, working individually or as a team, encounter problems at the same time, we arbitrate among the requests to identify the most urgent one for the human to address.

Given this approach, the second factor is clear: we must make it easy for the human to effect control and to rapidly assess the

situation. Human–robot dialogue is particularly important because it allows the operator to review what has happened, to understand problems each robot has encountered, and to be notified when his assistance is needed. Dialogue also improves context switching: enabling the human to quickly change his attention from robot to robot, directing and answering questions as needed.

B. Benefits of Collaborative Control

Collaborative control is an effective system model for teleoperation because it allows autonomy and human–robot interaction to vary as needed. This is particularly useful for environments that are unknown, dynamic, or difficult to plan for in advance. In these environments, system configuration (e.g., module parameters) may need continuous or repeated modification as the situation varies. Although such changes could be performed autonomously, in many cases human decision or approval is needed. Thus, providing dialogue between the human and robot is advantageous.

Being able to dynamically modify autonomy and interaction is also useful when an operator (or operators) must work with multiple robots. In such situations, especially when simultaneous control or supervision is required, the level of autonomy for each robot may need to be adjusted. For example, an event may occur which forces the operator to limit interaction to a single robot for an extended period of time. As a consequence, to prevent deadlock or blocking (i.e., robots waiting for human input), the level of autonomy for the other robots will have to be increased.

Finally, we have seen that dialogue allows the human to be highly effective. By focusing attention on where it is most needed, dialogue helps to coordinate and direct problem solving. In particular, we have found that in situations where the robot does not know what to do, or in which it is working poorly, a simple human answer (a single bit of information) is often all that is required to get the robot out of trouble.

ACKNOWLEDGMENT

The authors would like to thank S. Grange for his assistance in developing the PdaDriver.

REFERENCES

- [1] D. McGovern, "Experiences and results in teleoperation of land vehicles," Sandia Nat. Lab., Albuquerque, NM, Tech. Rep. SAND90-0299, 1990.
- [2] T. Fong, C. Thorpe, and C. Baur, "Advanced interfaces for vehicle teleoperation: collaborative control, sensor fusion displays, and remote driving tools," *Auton. Robots*, vol. 11, no. 1, pp. 77–85, July 2001.
- [3] T. Fong, "Collaborative control: A robot-centric model for vehicle teleoperation," Ph.D. dissertation, Robotics Inst., Carnegie Mellon Univ., Pittsburgh, PA, 2001.
- [4] T. Sheridan, "Eight ultimate challenges of human-robot communication," in *Proc. IEEE Int. Workshop Robot-Human Interactive Communication (RO-MAN)*, 1997, pp. 9–14.
- [5] G. Baltus *et al.*, "Towards personal service robots for the elderly," in *Proc. AAAI Workshop Interactive Robotics and Entertainment*, Apr. 2000.

- [6] A. Green and K. Severinson-Eklundh, "Task-oriented dialogue for CERO: a user-centered approach," in *Proc. IEEE Int. Workshop Robot-Human Interactive Communication*, Bordeaux and Paris, France, 2001, pp. 146–151.
- [7] I. Nourbakhsh *et al.*, "An affective mobile robot educator with a full-time job," *Artif. Intell.*, vol. 114, no. 1–2, pp. 95–124, 1999.
- [8] T. Laengle, T. Hoeniger, and L. Zhu, "Cooperation in human-robot-teams," in *Proc. Conf. Informatics and Control*, St. Petersburg, Russia, 1997, pp. 1306–1314.
- [9] P. Bonasso, "Issues in providing adjustable autonomy in the 3T architecture," in *Proc. AAAI Spring Symp.: Agents With Adjustable Autonomy*, 1999, pp. 11–16.
- [10] J. Rosenblatt, "DAMN: a distributed architecture for mobile navigation," in *Proc. AAAI Spring Symp. Lessons Learned From Implemented Software Architectures for Physical Agents*, Stanford, CA, 1995, pp. 167–178.
- [11] J. Albus, "NASA/NBS standard reference model for telerobot control system architecture (NASREM)," NIST, Gaithersburg, MD, Tech. Note 1235, 1987.
- [12] E. Krotkov *et al.*, "Safeguarded teleoperation for lunar rovers," in *Proc. Int. Conf. Environmental Systems*, Monterey, CA, 1996, SAE Tech. Paper 961585.
- [13] K. Holtzblatt and S. Jones, "Contextual inquiry: A participatory technique for system design," in *Participatory Design: Principles and Practice*, D. Schuler and A. Namioka, Eds. Hillsdale, NJ: Lawrence Erlbaum, 1993.



Terrence Fong received the B.S. and M.S. degrees in aeronautics and astronautics from the Massachusetts Institute of Technology, Cambridge, in 1988 and 1990, respectively, and the Ph.D. degree in robotics from Carnegie Mellon University, Pittsburgh, PA, in 2001.

He is currently a Postdoctoral Fellow at the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. His research interests include human–robot interaction, Web-based interfaces, virtual environments, and field mobile robots.



Charles Thorpe received the Ph.D. degree in computer science from Carnegie Mellon University, Pittsburgh, PA, in 1984.

He is Director of the Robotics Institute at Carnegie Mellon University. He has published over 120 peer-reviewed papers on 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 received the Ph.D. degree in micro-engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1992.

He is Adjoint Scientifique at EPFL and Director of the Virtual Reality and Active Interfaces Group, which he created in 1993. In addition to his work at EPFL, he was founder and CEO of 2C3D, a startup company specializing in real-time 3-D visualization for medical imaging and endoscopic applications.