Guest Editor's Introduction: Toward High-Definition Telepresence

The concepts of telepresence (and teleaction) refer (i) to technologies that allow human operators to feel (and act) as if they were present in a remote location and (ii) to the set of methods that allow information to travel in both directions between the remote location and the operator's location. In order to attain the maximum level of the operator's perception of immersion in the situation at the remote location, a number of goals must be achieved: the communication channels between the two sites must provide the necessary bandwidth and limits on latency times; sensors and actuators must have task-adapted and sufficient resolutions, precisions, and ranges; and the number of modalities (vision, touch, slip, positions, etc.) as well as the number of stimuli provided to the operators, must be adequate to maintain the illusion of "being there." Ideally, the equipment used by the operators is unobtrusive enough not to interfere with their movements and not to require them to adopt unnatural positions to pick up the sensory information provided from the remote site.

Due to recent developments in methodology and technology, but also owing to the ever-increasing computing power and available bandwidth, we are currently witnessing two trends: (i) telepresence is on the brink of becoming a very desirable technology for mainstream applications, for example, within the computer-games industry; and (ii) the most advanced equipment makes it possible to handle complex tasks with delicate objects, ranging from miniature parts, for example, in terms of handling hazardous materials, and so forth, to manipulation inside the human body in areas that were otherwise inaccessible. Obviously, a number of the following barriers could be overcome by highly sophisticated devices:

Firstly, there is *scalability:* in principle, nano-sized objects (but also very large and heavy parts) can be handled with similar or identical high-definition gear on the operator side. Clearly, neither of these situations could be mastered by humans directly without more or less adapted tools.

Secondly, *accessibility* is another barrier that can be overcome with new technologies. Here the simultaneous capabilities of reaching and maneuvering in confined spaces, as well as simultaneous navigation are important. Typically, they can only be achieved through highly adaptive actuators and high resolution sensors.

Thirdly, there is the barrier of *distance:* while there are numerous applications in which scalability and, in particular, accessibility play a major role, and distance is not an issue, long-distance uses will become more and more important. While space applications (e.g., robot control on other planets) are the most spectacular and compelling, the problems of communication delays and breakdowns, as well as lack of communication services quality, apply to internet-based terrestrial telepresence setups as well. Using the internet as a communication infrastructure for teleaction scenarios is an obvious and inexpensive choice and will become more and more attractive. However, such scenarios will only meet with acceptance when the related problems (e.g., variable time-delays, communication breakdowns, etc.) have been resolved. Research in this direction is therefore also very important.

Clearly, these three barrier types can be found in any combination in the diverse applications that will become possible based on the progress that has been achieved within the last few years. One of the most active centers in this field of research is the collaborative research center "High-Fidelity Telepresence and Teleaction (SFB453)" funded by the German National Science Foundation (DFG). The center is located at various sites in the greater Munich area (see www.sfb453.de for more information).

The activities in the last three years have focused on the development of new methodologies for teleaction, the evaluation of teleaction setups for their "proximity to reality" with respect to human psychology and the development of complete application scenarios, for example, for minimally invasive surgery. The main technical emphasis has been on the exploration of multi-modality issues—that is, in addition to the visual and acoustic sensory impressions, haptic impressions, in particular, are needed. This pertains to both tactile (pressure, temperature, roughness, vibrations, etc.) and kinesthetic (proprioception, inertia effects, gravity) channels. The main conceptual focus has been on the development of a new framework, called the "Haptic Visual Workspace (HVW)." The HVW should be understood as a commonly perceived space, in which several geographically distributed operators in a common environment have to solve a complex task through teleoperation and teleaction. All of them receive visual, auditory, and haptic impressions. Each of them also has available a teleoperator to implement the manipulations in the common work space. There may be one-to-one allocations between operators and teleoperators, but other mappings among tasks, operators, and teleoperators are possible as well, for example, several operators share one teleoperator, or one operator controls several teleoperators. In order to make these interactions possible, the actions of the individual teleoperators must be coordinated and synchronized. To achieve this, procedures have to be provided that are able to predict and/or prevent possible actions with several teleoperators. That is, they touch on the issue of partial autonomy and shared control of complex multi-operator, multi-teleoperator scenarios.

The articles in this special issue are all reports on the developments taking place within the HVW "umbrella." They are organized in such a way that we proceed from complete implementations to more specific and theoretical issues. They are grouped as follows:

In the first article, by Hermann Mayer, Istvan Nagy, Alois Knoll, Eva U. Braun, Robert Bauernschmitt, and Rüdiger Lange, the technology and use of haptic feedback in a teleaction system is explored. The authors have designed an experimental telemanipulator system for cardiac surgery that makes it possible for surgeons to perform delicate, minimally invasive procedures with high precision under stereoscopic view. It provides both force-feedback and Cartesian control (as a prerequisite for partial automation). Omitting technical details, the authors focus on the inclusion of force feedback and its evaluation by surgeons. Based on research experience with a two-arm version of the system, the design of the new robot setup presented in the article is based on four multi-purpose arms mounted on a gantry above the operating table. The force-feedback enables the surgeon to palpate arteriosclerosis, to tie surgical knots with real suture material, and to feel the rupture of the suture material, should it occur. Within the realm of this newly available technology, the hypothesis that haptic feedback in the form of sensory substitution facilitates performance of surgical tasks was evaluated. In addition, a further hypothesis was explored: the eminently high fatigue of surgeons during and after robotic operations, which may be caused by the surgeon's visual compensation due to the lack of force feedback.

Closely related to that work is the research described in the second article, by Barbara Deml. Here, human factors involved in the design of telepresence systems are investigated in order to provide guidelines from a human factors point of view. Up to now, there have been few standard input devices that could be used in telepresence systems, nor have there been any recognized design principles. There are two conflicting objectives in the design of telepresence systems: they should offer a high degree of performance and a high sensation of presence. However, the relationship between these goals is still very vague. Two experimental studies are presented. The first one is related to the control of multiple degrees of freedom and the second one refers to bimanual input control. Moreover, a study is presented that describes the relationship between presence and performance more precisely. Toward the end of the article, a computer-based design guide is suggested to provide a framework for these and further human factor aspects.

The theme of human factors is taken up in the third article as well: Helena Pongrac, Jan Leupold, Stephan Behrendt, Berthold Färber, and Georg Färber explore how to enhance live video streams with virtual reality, with respect to the variables of performance, situation awareness, and feeling. They have designed a system that presents virtual reality images combined with camera images captured at the remote teleoperator's site. The authors show that presenting a widened field of view to the human operator enhances the human performance and the feeling of telepresence. In a second experiment, they explore how the transition between video and virtual views should be implemented. Relevant criteria of this transition were defined. Their results show that the operator's ratings of quality, feeling of telepresence, and situation awareness are positively affected by variations of the transition parameters. Furthermore, a trade-off between the rating of quality and situation awareness was observed. As a result, a parameter selection scheme is presented which can serve as a design guideline for combining video and virtual views depending on the desired application.

The following article, by Fakheredine Keyrouz and Klaus Diepold, profoundly investigates a more specific aspect of telepresence systems: binaural sound source localization for telepresence. The authors focus on acoustic telepresence at both the operator and teleoperator sites. For installation at the remote site, they have built an innovative, binaural sound source localizer using generic head-related transfer functions that provides estimates for the direction of a single sound source with a very efficient algorithm. For adequate localization on the operator's side, the problem of spatially interpolating, head-related transfer functions for densely sampled high-fidelity 3D sound synthesis is addressed. In their application scenario, the synthesized 3D sound is presented to the operator over headphones to achieve high-fidelity acoustic immersion. A comparison with existing methods shows that the matrix-valued interpolation of the transfer functions in particular, results in superior performance for the reconstruction.

The article by Sandra Hirche and colleagues is in two parts (Part I by Sandra Hirche, Peter Hinterseer, Eckehard Steinbach, and Martin Buss; Part II by Sandra Hirche and Martin Buss). It investigates transparent data reduction from different aspects, both in the case of communication without time delay as well as with time delay. The focus is on the investigation of methods to reduce data transmission bandwidth for the haptic channel. The authors introduce a novel approach to reduce traffic by exploiting limits in human haptic perception. The idea behind the deadband control approach is that data are transmitted only if the signal change exceeds a signal amplitude-dependent perception threshold. Experimental user studies show that significant network traffic reduction can be achieved without impairing the perception of the remote environment. If time delay is also considered, however, stability issues can severely deteriorate the system's performance. To cope with this problem, the scattering transformation concept, a well-known time delay approach, is applied. Experimental user studies show that for this case, an average network traffic reduction of up to 96% can be achieved.

The final article, by Stella Clarke, Gerhard Schillhuber, Michael F. Zaeh, and Heinz Ulbrich, deals with the effects of simulated inertia and force prediction, including the time delay issues. Two methods are presented that aim to alleviate the negative effects of network delays on teleoperation. The first method investigates the hypothesis that simulated inertia in the haptic input device can be a supporting factor across delayed networks during teleoperation. Experiments were carried out under various network and inertia conditions. Simulated inertia was found to be neither a supporting factor nor a detrimental factor to operator performance and immersion in the presence of both delayed and non-delayed networks. The second method is a force prediction approach that extends the teleoperation system with a local force model. Instead of relying on the delayed force signals from the teleoperator side, haptic information can be extracted from this local force model. An experimental design is presented to demonstrate the benefits of this approach as compensation for the instabilities due to time delay.

A special thanks from the guest editor to the papers' authors for their excellent research on such a wide range of topics. We would also like to respectfully and gratefully acknowledge the comments and suggestions for improvement from the many reviewers. Their input is invaluable.

Professor Alois Knoll, Guest Editor
Department of Informatics
Chair of Robotics, Department of Robotics and
Embedded Systems, Technische Universität München

We would like to express our thanks to both Professor Knoll and Dr. Gerhard Schrott of the Department of Informatics at Technische Universität München, who shared the responsibilities of guest editing this special issue. Thanks also are due to Professor Martin Buss, Chair of the Collaborative Research Center (SFB) 453, who initiated this special issue.

The Editors