On the Role of Multimodal Communication in Telesurgery Systems

Robert Bauernschmitt\textsuperscript{5}  Eva U. Braun\textsuperscript{2}  Martin Buss\textsuperscript{2}  Florian Fröhlich\textsuperscript{6}  Sandra Hirche\textsuperscript{2}
Gerhard Hirzinger\textsuperscript{6}  Julius Kammerl\textsuperscript{1}  Alois Knoll\textsuperscript{1}  Rainer Konietzschke\textsuperscript{6}  Bernhard Kübler\textsuperscript{6}
Rüdiger Lange\textsuperscript{6}  Hermann Mayer\textsuperscript{4}  Markus Rank\textsuperscript{2}  Gerhard Schillhuber\textsuperscript{3}  Christoph Staub\textsuperscript{4}
Eckehard Steinbach\textsuperscript{1}  Andreas Tobergte\textsuperscript{6}  Heinz Ulbrich\textsuperscript{3}  Iason Vittorias\textsuperscript{2}  Chen Zhao\textsuperscript{3}

\textsuperscript{1} Institute for Media Technology, Technische Universität München, Munich, Germany
kammerl@tum.de eckehard.steinbach@tum.de
\textsuperscript{2} Institute of Automatic Control Engineering, Technische Universität München, Munich, Germany
mb@tum.de hirche@tum.de m.rank@tum.de vittorias@tum.de
\textsuperscript{3} Institute of Applied Mechanics, Technische Universität München, Munich, Germany
schillhuber@amm.mw.tu-muenchen.de ulbrich@amm.mw.tu-muenchen.de zhao@amm.mw.tu-muenchen.de
\textsuperscript{4} Institute of Robotics and Embedded Systems, Technische Universität München, Munich, Germany
knoll@in.tum.de mayerh@in.tum.de staub@in.tum.de
\textsuperscript{5} German Heart Center Munich, Clinic for Cardiovascular Surgery, Munich, Germany
bauernschmitt@dhm.mhn.de brauneva@dhm.mhn.de lange@dhm.mhn.de
\textsuperscript{6} German Aerospace Center, Institute of Robotics and Mechatronics, Oberpfaffenhofen, Germany
Florian.Froehlich@dlr.de Gerhard.Hirzinger@dlr.de Rainer.Konietzschke@dlr.de
Bernhard.Kuebler@dlr.de Andreas.Tobergte@dlr.de

Abstract—Telesurgery systems integrate multimodal communication and robotic technologies to enable surgical procedures to be performed from remote locations. They allow human surgeons to intuitively control laparoscopic instruments and to navigate within the human body. In this paper, we present selected topics on multimodal interaction in the context of telesurgery applications. These are results from the collaborative research project SFB 453 on “High-Fidelity Telepresence and Teleaction” which is funded by the German Research Foundation in the larger Munich area. The focus in this paper is on multimodal information processing and communication including simulation of surgical targets in the human body. Furthermore, we present an overview of our advanced multimodal telesurgery demonstrators that provide a comprehensive platform for our collaborative telepresence research.

I. INTRODUCTION

Minimally invasive surgery was a revolutionary technique when it appeared in the early 1980s as an effective alternative to open surgery. Large incisions required by conventional, open surgery are avoided by using laparoscopic tools and cameras inserted into the human body through small puncture wounds. As a result, the pain, recovery time, blood loss, risk of infection and other intra-operative complications are reduced. However, the loss of haptic feedback, natural hand-eye coordination and the limited workspace and degrees of freedom when using laparoscopic equipment poses a serious challenge. To overcome these limitations and in order to extend capabilities of the surgeons, surgical telepresence and teleaction systems (TPTA) have been deployed to virtually immerse the surgeon into the human body through multimodal feedback.

In a surgical TPTA system, a surgical teleoperator (TOP) - a surgery robot equipped with multimodal sensors and actuators - is controlled by a human operator (OP) connected to a human system interface (HSI). The HSI reflects the sensory information acquired by the robot back to the OP using displays for visual and haptic information. The multimodal feedback allows the surgeon to intuitively control the laparoscopic equipment and navigate within the human body. In this context, the haptic modality can play a central role. Position and velocity of the OP’s movements, typically the hands, are tracked by the HSI and used to command movements of the surgery robot. When contact within the surgery environment is encountered, corresponding contact forces are transmitted over the communication system to the surgeon, helping him to reduce damage to tissue, breaks of suture material, etc. due to excessively high contact force. As human life is at stake, the TPTA system must be robust, fail-safe, and resilient to communication problems.

The collaborative research project SFB 453 on “High-Fidelity Telepresence and Teleaction” funded by the German Research Foundation (DFG) has a focus on overcoming the “barrier of human skin” to allow surgeons greater access to the body with less trauma induced to the patient. In this paper, we discuss the importance of multimodal communication for telesurgery, including a description of two surgical teleoperators developed within the SFB 453 that provide a comprehensive platform for our collaborative research project.
The remainder of this paper is organized as follows. Section II discusses the integration of multimodal sensory information by the OP when time delay is different for each modality. Upper bounds for temporal discrepancy are considered. In Section III, we present methods for efficiently communicating haptic signals over a band limited network. Advanced techniques for improving real and simulated telesurgery scenarios using multimodal feedback are discussed in Section IV. Section V deals with haptic rendering of tool contacts with tissues. In Section VI we give an overview of two surgery telemanipulation robots developed in the context of the SFB 453. These teleoperators are highly integrated with rich sensor information to provide high-fidelity multimodal feedback to the OP. Section VII concludes the paper.

II. TEMPORAL MULTIMODAL INFORMATION PROCESSING
A realistic impression of the properties of the environment inside the body is important in a telesurgery context as it affects the task performance of the surgeon. Note that especially in telesurgery the haptic modality plays an important role, e.g., in diagnostic techniques as palpation. The quality of the multimodal in-body-environment is challenged by diverse factors such as signal noise, unequal scaling between visual and haptic feedback, and unimodal as well as intermodal time delay. The latter arises from multiple sources like data compression algorithms and communication latency between the HSI and the surgery robot. Delayed visual feedback is primarily a perceptual issue for the human surgeon. Delay in the haptic signals, however, decreases the stability margins of the control loop that spans over the operator, the system interface, the communication channel, and the teleoperator. Sophisticated control methods have been developed in the past to overcome this issue, see e.g. [1], [2] for an overview. In addition to these stability issues, the question arises, to which extend multimodal feedback delays are tolerable without impairing the performance of the surgeons. It was suggested in [3] that the dominant factor for decreasing performance in a surgical task through a time-delayed teleoperator is not the delay in the visual modality but in the force-feedback loop. While they had comparably large time delays of 0.5-1.2 s and focused on the influence of different delays on task performance, we performed experiments to determine the actual discrimination thresholds for the detection of force-feedback latency. Our investigation revealed that a delayed linear spring environment could reliably be distinguished from a non-delayed spring of the same stiffness when the delay was larger than 24 ms [4]. Consistent with this, Jay et al. [5] found that the operator performance decreases linearly with haptic time delay greater than 25 ms. This allowable latency has to be considered as an upper bound because medical personal are known to have haptic capabilities superior to untrained people [6].

When the visual feedback from the teleoperator side is delayed, the discrimination threshold is around 33 ms [7], i.e., a latency below 33ms is hardly noticeable. The situation changes when head-tracking systems and head-mounted displays are used to allow for more intuitive control, e.g., of a laparoscope. This technique, though increasing the performance of the surgeons, comes at the cost of a tolerable time delay of only 15ms [8]; the tolerance to delay increases with decreasing head motion speed. This is in line with earlier findings on the influence of visuomotor action on the simultaneity perception [6].

Due to different encoding or signal processing delay, haptic and visual impressions may not always be temporally consistent. Experimental studies in [3] suggest, however, that task performance is much more affected by haptic feedback latency and showed no statistically significant deterioration of performance even with 0.6 s of asynchrony between haptic and visual feedback. A general rule of thumb for transmission of multimodal information is therefore to send haptic information as soon as it is available and as quickly as possible and not to buffer the data to ensure a consistent impression of the environment.

III. EFFICIENT COMMUNICATION OF HAPTIC DATA
In telesurgery, the surgeon performs robotic surgery on a physically separated patient using a communication link. Therefore, surgeon and patient do not have to be at the same location and patients can access surgeons around the world without traveling beyond their local hospital. In 2001, the first complete remote surgery on a human patient was performed over a distance of 7000 km between Strasbourg and New York with a latency of 135 ms [9].

In TPTA systems with a large number of degrees of freedom the amount of generated data is high and the transmission can be challenging. Haptic information is transmitted bidirectionally between the surgeon and the surgery robot, closing a control loop over the communication channel. As a consequence, any delay in the communication channel impairs the stability of the involved control loops and may reduce the quality of the user feedback and performance of the TPTA system. To enable remote surgery telemanipulation in the presence of significant communication delays, several control architectures from the field of control engineering are known to solve these stability issues, see e.g. [1], [2] for an overview.

Nevertheless, the existence of significant communication and/or computation time delay in the haptic control loops always has a noticeable impact on the system performance. Therefore, buffering, block based signal processing and transmission approaches which are widely applied in the field of audio and video communication, are not applicable for haptic signals. To minimize introduced latency, haptic events need to be sent immediately resulting in high packet rates of up to the device sampling rates (usually 1 kHz for haptic signals). In packet-based networks, this strategy consumes expensive network resources due to network protocol overhead.

To address these issues, we proposed in our earlier work a perceptual coding approach for haptic data, the so called "deadband approach" [10], [11]. It exploits psychophysical findings, namely the Weber’s Law of Just Noticeable Differences, which defines the smallest amount of change within a haptic signal that can be detected just as often as it cannot be
detected. It can be described by

$$\frac{\Delta I}{I} = k = \text{const.}$$

(1)

where $I$ is the stimulus intensity, $\Delta I$ is the so-called Difference Threshold or the Just Noticeable Difference (JND) and $k$ is a constant, which describes the linear relationship between the JND $\Delta I$ and the initial stimulus $I$.

In conjunction with haptic data reduction, it allows us to define a region in which signal changes are too small to be perceivable. As only noticeable changes need to be transmitted over the communication channel, the perceptual deadband approach allows for a significant, signal adaptive packet rate reduction. At the decoder, the transmitted irregular signal updates are input to a reconstruction algorithm to upsample the haptic signal to a constant rate which is needed for the local control loops. Applying the Weber-based deadband scheme on haptic signals achieves a data reduction of up to 96% without impairing the performance of the TPTA system [10]. Moreover, specifically designed passive reconstruction algorithms ensure the stability and performance of the overall system, without and with time delay [12], [13]. Furthermore, the deadband approach can be combined with predictive coding, which leads to an additional reduction of haptic data on the communication channel [14].

IV. VIRTUAL REALITY IN TELESURGERY

In telesurgery systems, virtual reality models enable us to simulate the conditions at the surgery environment. Visual and haptic rendering of the patient’s organs and tissues including its interaction with the surgery instruments provide an important assistance to doctors in their operation of surgical tools. Furthermore, it allows for developing educational and training applications which enable surgeons to simulate surgery from or in remote places [15]. In addition, in telepresence systems with latency on the communication channel, a local simulation can be used to compensate for time delays [16]. By substituting the remotely received feedback with the locally rendered multimodal information, the actions of the surgeon can be kept synchronous to the displayed feedback. In addition, the deployment of virtual walls allows us to prevent accidental contacts with sensitive objects within the workspace, which results in increased safety for minimal invasive procedures.

According to the surface properties of the virtual body model, diverse techniques for visual and haptic rendering such as Finite Element Method (FEM) [17] and mass-spring models are applicable. The advanced FEM copes with complex object’s geometries and tissue layers and hence, it is more suitable for detailed surgery applications. To take small and large deformations into account, different FEM modeling approaches can be applied. The nonlinear, linear or corotated FEM approaches describe deformations of an object differently and result also in different computational costs for evaluating the system equations. Thus, the choice of FEM modeling scales between the accuracy of the virtual object’s deformation and the computing time.

Compared to visual feedback, haptic rendering has demanding computational requirements. In order to keep local control loops stable and to provide realistic haptic feedback, a typical update rate of approx. 1000 Hz is applied which is often unfeasible due to limited computational resources and high complexity of haptic rendering algorithms. To address this issue, haptic rendering techniques are often running at reduced rate, followed by applying a low-complexity inter- and extrapolation algorithm for upsampling the haptic feedback to a constant and high update rate.

V. HAPTIC RENDERING OF TOOL CONTACTS FOR DELAY COMPENSATION IN TELESURGERY

As discussed earlier in this paper, latency on the communication channel quickly endangers the stability of local control loops within telesurgery systems. This issue can be addressed by designing mathematical models which describe haptic feedback during tissue contact. By substituting delayed force measurements with locally generated haptic feedback we can compensate delay on the communication channel. Interaction during contacts of the laparoscopic equipment can be modeled with deformable bodies, which describe the mechanical behavior of the operated tissues. Different modeling approaches have been proposed in this context, such as the mass-spring models, the boundary element method (BEM) and the FEM. According to [18], FEMs of deformable bodies show great performance in this area and are deployed in our work using tetrahedral elements. The geometry of the object can be imported from the computed tomography scanner or be measured by laser scanning and image processing. Then using the surface reconstruction and mesh generation, the geometric model can be created. The elasticity of most materials can be considered to be linear up to a certain degree. However, soft tissues require to model the viscoelasticity, known from continuum mechanics, which can be also numerically modeled [19]. Here, the contact force evaluation is based on the following equation:

$$K \cdot u = f^e - f^h,$$

(2)

where $K$ is the model stiffness matrix, inclusive of the material parameters, $u$ and $f$ are the nodal displacement and nodal force vectors. The effects of the loading history are involved in the term $f^h$. These viscoelastic models are generally computationally expensive. Thus, to apply these models in the surgery TPTA system, an extended Kalman filter (EKF) and an analytical contact force model have been implemented to estimate the contact force during the force evaluation of the FE models. By this means the computing time of the FE model can be reduced in order to achieve a high update rate required by the local haptic control loops. To adapt the locally deployed contact model to the currently touched tissues, its interaction with the surgery instruments provide an impor-

---

\[ \Delta I \]

\[ \text{JND} \]
The Technische Universität München showed in the new haptic collaborative work with the Institute of Embedded Systems of feedback for the telemanipulative working heart surgeon. The system is implemented showing the must-have of haptic feedback systems for minimally invasive surgery is presented.

In the following, an overview of our telesurgery demonstrator systems for minimally invasive surgery is presented.

A. Surgical telemanipulation system for heart surgery

In the German Heart Center Munich a telemanipulation system is implemented showing the must-have of haptic feedback for the telemanipulative working heart surgeon. The collaborative work with the Institute of Embedded Systems of the Technische Universität München showed in the new haptic system significantly reduced working forces in haptic experiments. Therefore, haptic feedback for the robotic working heart surgeon and for the safety of the patients is required.

The implementation of telemanipulator systems for cardiac surgery enabled heart surgeons to perform delicate minimally invasive procedures with high precision under stereoscopic view. Heart surgeons require telemanipulation for numerous minimally invasive heart surgeries. Until now the commercially available telemanipulators do not offer haptic feedback to the surgeon.

The introduction of master/slave systems for minimally invasive and endoscopic operations showed the limitations and difficulties of the system shortly after the first series of operations. For further improvement and easier handling of the telemanipulator, we evaluated the feedback of force on surgical skills dependent on different surgical experience [21]. In this work we present the value of surgical experience for telemanipulation including force feedback. The amount of force in the dominant and non-dominant hand, its dependency on the surgical experience and necessity for consequences in the design of surgical telemanipulators was not yet under examination in previous studies.

The experiments have been conducted with a system for robot-assisted minimally invasive surgery, which was developed by the German Heart Center Munich (see Figure 1). The slave manipulators of the system consist of four robots (Mitsubishi MELFA 6SL), which are mounted on a gantry on the ceiling. The robots are equipped with minimally invasive instruments originally deployed by the da Vinci surgical system. The robot is concatenated with the surgical instrument by a magnetic coupling, which prevents the instrument from damages in case of a severe collision. The instruments are powered by small servo motors, which are integrated into the coupling mechanism. The user’s place with 3D display is located in front of the main in-/output devices, two PHANToM haptic displays [22]. The PHANToMs themselves are assembled upside down. This arranges for less constricted flexibility of the stylus pen. Forces occurring at the instruments are measured by strain gauge sensors and fed back by means of the haptic devices.

The maximum forces not only in surgical cutting, but also in suturing are significantly reduced with haptic feedback. Significant results were shown in surgical skills. This platform represents a simulation environment for modeling and testing haptic interaction with a tissue model. The tissue trauma is also significantly reduced. Students and surgeons appreciate the opportunity to work with haptic feedback in suturing, knotting and handling the needle over. The application and the amount of forces can be restricted to tolerable amplitude to prevent harmful collisions not only between the instruments but also between the instruments and the thorax of the patient.

B. MiroSurge System for minimally invasive robotic surgery

The DLR MiroSurge system for research in minimally invasive robotic surgery (MIRS) is composed of the versatile, lightweight robot MIRO and is fully torque controlled in seven DoF [23]. The MICA instruments dedicated to MIRS serve as the surgical front-end of the MIRO in MIS extending the robots dexterity beyond the human skin barrier. The MICA features a universal joint, restoring the two DoF lost at the fulcrum point, a six DoF force/torque sensor on the tool-tip and an additional gripper for manipulation of tissue [24]. In MIRS three MIROs can be mounted to an operation table, as shown in Figure 2. Two MIROs are carrying MICAs and one is guiding a stereo endoscope. The robots and entry point positions are optimized by planning software. The surgeon specifies the operating volume in a 3D model, e.g. generated from CT-scan, and the coarse positions of the entry points. He

\[ f(x) = \frac{1}{2}||f_r - f_s(x)||, \]

where \( f_r \) and \( f_s \) are the reference and the simulation force. The vector \( x \) is the material parameter vector of the model, which includes the Young’s modulus, the Poisson’s ratio and the relaxation parameters, such as the relaxation amplitudes and the relaxation time coefficients. Furthermore, the tissues should be considered as anisotropic and inhomogeneous materials. The anisotropy can be taken into account by defining the material parameters separately in each coordinate direction, and a modeling method is also developed for inhomogeneous materials [20].
MiroSurge provides force feedback with two haptic hand controllers (omega.7, by Force Dimension, Inc., Lausanne, Switzerland). The motions of the user’s finger tips are projected into the instruments’ tip inside the patient. The instruments follow the user’s command with an arbitrary scaling factor. Measured forces and the gripping force are fed back to the haptic device and can be felt by the user. An autostereoscopic display provides the endoscopic view from inside the patient without special glasses, as shown in Figure 2. Tracking of the user’s eyes enables stereo rendering according to the current position of the user, so that the user can move in front of the display. An alternative interaction modality is implemented with a contact free interface. Conventional surgical pincers are equipped with marker balls of an optical tracking system, giving the surgeon his used instruments in hand. Force information is augmented in the stereo scene, as shown in Figure 4. An arrow is displayed at each instrument, showing the force exerted by the environment. The augmented stereo stream is distributed via Ethernet by a vision server. Clients can connect from any PC in the network with different mono and stereo output modalities.

Presently, an expansion of the MiroSurge system towards tactile feedback is under development to minimize the risk of accidental ruptures of (arterial) vessels. In open surgery slight pulsations within tissue usually indicate vessels to the surgeon, perceivable by palpation. Blunt dissection in these regions might cause hardly controllable bleedings. In MIRS, where direct manual access is impossible, vessels can alternatively be traced by ultrasound. The processed data indicating a vessel are intended to be modality substituted and displayed kinesthetically on the haptic interfaces.

**VII. Conclusion**

This paper discussed selected topics on multimodality for telesurgery in the context of the Collaborative Research Center SFB 453 on “High-Fidelity Telepresence and Teleaction”. Surgery TPTA systems enable surgeons to remotely control robotics systems to perform minimally invasive surgical procedures on patients at a remote location. Processing and communication of multimedia information between the surgeon and the robot are therefore essential to provide quality surgical service.

Understanding the capability of human surgeons to process multimodal feedback allows us to define tolerable bounds for latencies in each deployed modality. Our studies show that task performance is most affected by haptic feedback latency. Therefore, haptic data has to be transmitted immediately and displayed with minimal delay. Since this strategy leads to high packet rates and stress on the underlying network resources, we propose a perceptual coding approach for haptic data that reduces the number of haptic samples for transmission without affecting the performance of the telesurgery system. Furthermore, aspects of designing and developing virtual telesurgery scenarios are discussed. The goal of locally simulating the surgery environment is to increase the safety of minimally
invasive procedures through task planning which can make the real surgery more accurate. Also, local modeling and rendering of the remote multimodal feedback helps to compensate for latency in time-delayed telesurgery scenarios. By designing contact models based on the finite element method, mechanical behavior of tissues during tool contact can be locally simulated to support stability of the involved control loops. By graphical modeling of the remote environment and teleoperator, visual feedback delay to the OP can be reduced. In addition, we present an overview of our multimodal surgery telemanipulation demonstrators used to evaluate the performance and quality of our developed methods. The MiroSurge system for minimally invasive robotic surgery developed by the German Aerospace Center provides a comprehensive research platform with bimanual manipulation ability and multimodal feedback. Furthermore, experiments conducted with the heart surgery demonstrator system at the German Heart Center revealed that haptic feedback results in significantly reduced applied forces when remotely interacting with tissues.

Current work focuses on multiple-operator-multiple-teleoperator surgery systems, which allow for distributed co-operation of multiple human surgeons (telecooperation). New challenges arise in this context in the distribution of multimodal information among the different operators and teleoperators. Additionally, the integration of autonomous functionality to assist and support the surgeon is subject of current research.

ACKNOWLEDGMENT

This work is supported by the German Research Foundation (DFG) within the Collaborative Research Center SFB 453 on “High-Fidelity Telepresence and Teleaction”.

REFERENCES


