INVITED PAPER

Haptic Communications

Haptic communications is the subject of this paper. As in audiovisual communications, users experience a strong central perceptual component. In recognition of this, there is a section on perceptual learning.

By Eckehard Steinbach, *Senior Member IEEE*, Sandra Hirche, *Senior Member IEEE*, Marc Ernst, Fernanda Brandi, *Student Member IEEE*, Rahul Chaudhari, *Member IEEE*, Julius Kammerl, *Student Member IEEE*, and Iason Vittorias, *Student Member IEEE*

ABSTRACT | Audiovisual communications is at the core of multimedia systems that allow users to interact across distances. It is common understanding that both audio and video are required for high-quality interaction. While audiovisual information provides a user with a satisfactory impression of being present in a remote environment, physical interaction and manipulation is not supported. True immersion into a distant environment and efficient distributed collaboration require the ability to physically interact with remote objects and to literally *get in touch* with other people. Touching and manipulating objects remotely becomes possible if we augment traditional audiovisual communications by the haptic modality. Haptic communications is a relatively young field of research that has the potential to substantially improve human–human and human–machine interaction. In this paper, we discuss the state-of-the-art in haptic communications both from psychophysical and technical points of view. From a human perception point of view, we mainly focus on the multimodal integration of video and haptics and the improved performance that can be achieved when combining them. We also discuss how the human adapts to discrepancies and synchronization errors between different modalities, a research area which is typically referred to as perceptual learning. From a technical perspective, we address perceptual coding of haptic information and the transmission of haptic data streams over resource-constrained and potentially lossy networks in the presence of unpredictable and time-varying communication delays. In this context, we also discuss the need for objective quality metrics for haptic communication. Throughout the paper, we stress the fact that haptic communications is not meant as a replacement of traditional audiovisual communications but rather as an additional dimension for telepresence that will allow us to advance in our quest for truly immersive communication.

KEYWORDS | Haptic communications; haptic compression; multimodal integration; perceptual coding; perceptual learning; psychophysics; telemanipulation; telepresence

I. INTRODUCTION

The field of audiovisual communications has witnessed tremendous growth and progress during the last decades. This progress has led to improved productivity and quality of experience in remote interaction scenarios such as video conferencing. With increasing quality, users feel more present, experience an improved feeling of togetherness, and are able to perform more subtle interactions. The resulting level of immersiveness can, for instance, be experienced in the commercial high-end teleconferencing products from CISCO (TelePresence) [1] and Hewlett Packard (Halo) [2], which have managed to partially fulfill the promise of connecting people remotely and giving them a feeling of presence and closeness that we usually can only experience when people are in the same room. Driving factors that made this improved telepresence possible are high-quality audio and video capturing and display devices, highly efficient audio and video coding standards, as well as the ever increasing transmission

Manuscript received May 16, 2011; revised August 16, 2011; accepted October 20, 2011. Date of publication February 17, 2012; date of current version March 21, 2012. This work was supported by the European/Brazilian Network for Academic Exchange (EUBRANEX), the German Research Foundation (collaborative research center SFB 453 and project STE 1093/4-1), and the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC under Grant 258941.
E. Steinbach, **F. Brandi**, **R. Chaudhari**, and **J. Kammerl** are with the Institute for Media Technology, Technische Universität München, Munich D-80333, Germany (e-mail: eckehard.steinbach@tum.de; fernanda.brandi@tum.de; rahul.chaudhari@tum.de; kammerl@tum.de).
S. Hirche and **I. Vittorias** are with the Institute for Automatic Control Engineering, Technische Universität München, Munich D-80333, Germany (e-mail: hirche@tum.de; vittorias@tum.de).
M. Ernst is with the Faculty of Biology/Kognitive Neurowissenschaften, Universität Bielefeld, Bielefeld D-33615, Germany (e-mail: marc.ernst@uni-bielefeld.de).

Digital Object Identifier: 10.1109/JPROC.2011.2182100

capacity of our communication networks. It is expected that 3-D video capture and display will further increase the level of immersiveness experienced by users in telepresence systems.

Despite all these advances, however, presence in a remote real or virtual environment and truly immersive communication cannot be complete without the ability to physically interact with distant objects and humans [3]–[17]. To achieve immersion into the task environment, telepresence systems have continually attempted to supply the user with comprehensive sensory feedback—auditory, visual, and haptic. The intricacies of involving haptics originate from the human-in-the-loop nature of haptic interaction. The human can not only perceive the haptic feedback offered—similar to other modalities—but also, in response, physically act upon an environment to alter it. Therefore, a human-centric design and analysis of haptic interaction systems is called for [18]–[20].

In particular, in shared cooperative (virtual) environments, the communication of multimodal sensory information plays a fundamental role as it enables the participants to communicate and interact through their actions [21]. In this context, the importance of the haptic modality and its positive influence on such shared user experience is discussed in [21] and [22]. The results clearly show that shared haptic interaction toward a common objective significantly improves task performance and the sense of togetherness. In addition to interpersonal communication, the integration of haptics significantly improves the perception of important information about the environment such as surface characteristics and shape of remote (virtual) objects [23]–[25]. The integration of haptic communication enables novel applications in the field of teaching/training, telerobotics, entertainment, gaming, etc.

In this invited paper, we complement and extend our discussion of haptic communications presented recently in [26]. Our selection of topics is such that overlap with [26] is avoided as much as possible without sacrificing the completeness and readability of this article. Although in the long run a joint treatment of all modalities (audio, video, and haptics) is required, we constrain our discussion in this paper mainly to the haptic modality and its role in telepresence and telemanipulation systems including virtual reality systems.

The remainder of this paper is organized as follows. We start in Section II with a discussion of the multimodal integration of video and haptics followed by a short description on how the human adapts to discrepancies and synchronization errors between different modalities. We finish Section II with a summary of the main properties of human haptic perception. In Section III, we begin our technical discussion of haptic communication with a focus on the characteristics of haptic telepresence, haptic control architectures, and performance evaluation metrics. Section IV is devoted to perceptual coding of haptic data streams for haptic telepresence. In Section V, we briefly touch upon communication protocols for haptic communications. Section VI introduces the field of error-resilient haptic communications. The discussion starts with an overview of typical artifacts that are introduced by packet loss in haptic telepresence followed by a proposal on how to perform error-robust perceptual haptic encoding. We stress in Section VII the fact that accelerated progress in haptic communications requires the availablity of objective quality metrics. In this context, we describe our recent proposal on how to perform objective quality assessment for haptic telepresence sessions in virtual environments. At the end of the paper, we provide a summary of selected challenges for future work in the area of haptic communications in Section VIII and conclude the paper in Section IX.

II. HAPTIC PERCEPTION AND MULTIMODAL INTEGRATION

Human haptic perception is highly multidimensional. For example, the shape of an object can be perceived through the force patterns generated by interacting with the object. Additionally, shape can also be perceived through the position information provided by the kinesthetic signals derived from the joints, tendons, and muscles. This multidimensional nature of haptics demands that the different sources of sensory information be combined in the human brain in order to arrive at a coherent and unified percept of the objects in the world [27]. It is, of course, not just the haptic modality that provides information about the environment. Foremost, vision and audition also display rich sources of information for the brain to learn about the objects in the environment. Therefore, the following question arises: How does the human brain combine sensory information across and within a sensory modality to construct a reliable and robust percept of the world? In general, sensory information derived from different sensory modalities can be complementary, such as color information from vision and force information from touch, giving rise to a rich representation of the objects in the human brain. Alternatively, information may be redundant, such as size information, which can be derived both, from vision and touch. In the following, we will elaborate on how the human brain integrates such redundant sources of information.

A. Mechanisms of Multisensory Integration

Just like information processing in any technical system, biological information processing is corrupted by noise. The optimal way to integrate unbiased but noisy sensory estimates S is to form a weighted average of the different redundant sources of information. Assuming the noises are independent and Gaussian distributed, the weights $\boldsymbol{w}$ should be set proportional to the inverse of the variances of these distributions. If we define the precision to be $r = 1/\sigma^2$, this can be written as

$$\hat{S} = \sum_i w_i \hat{S}_i \quad \text{with} \quad \sum_i w_i = 1 \tag{1}$$

and

$$w_j = \frac{r_j}{\sum_{i=1}^{N} r_i} \tag{2}$$

where i runs over all the different sensory estimates that should be combined. We consider this optimal, because this leads to an unbiased combined estimate, for which the noise is maximally reduced

$$r = \sum_i r_i. \tag{3}$$

Recent behavioral experiments with human participants have demonstrated that humans actually combine sensory information in such a statistically optimal manner. This was first demonstrated by Ernst and Banks [28] for the estimation of visual and haptic size estimates. For this, they used psychophysical methods, experimentally introduced small conflicts between the information provided by vision and touch, and manipulated the precision of the sensory channels by adding noise to the stimulus. These experiments confirmed that the combined percept was always a compromise between the information provided by vision and touch, as predicted by the weighted average. When there was little noise on the visual channel, it was the visual estimates that dominated the percept. However, when there was noise added to the visual signals, the combined size percept gradually shifted toward the size specified by the haptic modality, until finally, when there was a lot of noise added to the visual display, it was the haptic modality that dominated the combined percept. Importantly, this dynamic reweighting of information occurred on a trial-by-trial basis, which implies that the weights are immediately set correctly for the given perceptual situation. In other words, it seems as if the human perceptual system not only has an estimate of the magnitude of the sensory signals, but it jointly also derives an estimate of the amount of noise that is contained in the signal. Furthermore, by providing visual and haptic information together, Ernst and Banks [28] showed that perceptual performance increased compared to a situation when there was either only visual or only haptic information available. The increase in performance comes from the reduction of noise when combining sensory information and conforms to the value predicted by (3). Taken together these results demonstrate that the different sensory modalities cooperate by integrating multisensory information leading to an increase in perceptual performance.

Visual-haptic integration of size information as investigated in [28] was only the first example that demonstrated optimal cooperation between the senses. By now there are many other examples from various perceptual situations, which back this result up and which thereby demonstrate that multisensory integration is a general principle used by the human brain to reduce uncertainty and to increase perceptual performance [29]–[39].

Given that multisensory integration is such a general principle employed by the human brain, we may be able to exploit our understanding of it in technical applications, as discussed here in this paper in the context of signal communications for teleoperation. To continue on similar lines as the above discussion, let us consider the provision of video and the haptic feedback in a teleoperation system. Communication of sensory data in the teleoperation system subsume their acquisition, processing, and transmission. Each of these processes is prone to the introduction of noise into the signal in various ways, e.g., sensor noise, quantization noise, channel noise, etc. Let us consider the case where one of the two sensory channels—video or haptics—is known to be more noise-prone than the other. According to above findings the combined percept will emphasize the more reliable (less noisy) channel. Thus, by allocating a larger share of the available bit-rate budget to the more reliable modality, the overall performance can be improved.

B. Perceptual Consequences of Conflicting Multisensory Information

In order to integrate multisensory signals the brain first has to know which signals belong together. This is a nontrivial problem, also known as the correspondence problem. When there is no correspondence between the multisensory signals, for example, when the signals are unrelated or when there are large spatial or temporal discrepancies between the streams of multisensory information, integration will not occur (e.g., [40]). The correspondence problem becomes particularly severe in noisy and cluttered environments. Imagine walking down a busy shopping street, with lots of people talking, dogs barking, cars honking, music coming out of different shops, and church bells ringing. How does the brain associate one particular sound with its corresponding visual counterpart in order for those signals to be integrated? Our current understanding of this problem is still relatively poor. However, first progress has been made into using causal inference and cross correlation for solving this problem (e.g., [41]). This problem is complicated by the fact that the mapping between multisensory signals is not fixed, but can vary with exposure to altered sensory environments—a process also called perceptual recalibration [42]. Such altered sensory environments are very common even in everyday life, for example, when putting on or taking off a pair of spectacles, when using tools, or when wearing gloves while interacting with the world. The classical example for such recalibration processes goes back to Hermann von Helmholtz (1867) who first demonstrated that we quickly adapt our pointing behavior after wearing prism glasses. By now there is a multitude of demonstrations of such recalibration processes including both the spatial and temporal domains. When a discrepancy occurs between different streams of multisensory information—in space or in time—the perceptual system corrects for this mismatch, so that the discrepancy becomes less noticeable with time. One example that has recently been studied to some extent concerns the perceptual consequences when exposed to temporal delays between multisensory signals (e.g., [43]–[46]). Such adaptations have been found to occur quickly (within minutes) in a range of up to a couple of tenths of milliseconds. Exploiting these adaptation processes, the human perceptual system stays optimally adapted to the ever-changing statistical regularities of the environment it is currently exposed to. The mechanisms behind these recalibration processes, however, are still largely unknown. First attempts have recently been made using models based on Kalman filtering to describe these learning processes [47], [48].

Here, we discuss the technical implication of these observations for signal communications. As mentioned before, when conflicts between sources of sensory information occur, integration quickly breaks down. At the same time, it is also known that the human perceptual system quickly adapts to persistent conflicts. However, a technical system that requires the user to frequently adapt to novel conflicting situations will have an unsatisfactory performance in terms of quality-of-experience (QoE). Hence, in order to facilitate the coherent perception of an event across different sensory feedbacks, visual-haptic asynchrony should be systematically minimized in the teleoperation system, for instance, via intelligent statistical multiplexing of the audiovisual-haptic signals on the feedback communication channel.

C. Human Haptic Perception

Human haptic perception is concerned with the sense of touch. Haptic perceptions require direct contact with the environment. For exploring the environment, the sense of touch inherently involves action, arguably to a greater extend than any of the other senses do. The first systematic studies into human haptic perception go back to Ernst Heinrich Weber (1795–1878), one of the founding fathers of modern psychophysics [49]. Weber examined the precision of the sense of touch and established the well-known relationship named after him—Weber's law. This law states that the just noticeable difference (JND) between two stimuli, which is the minimum change in the magnitude of a stimulus that can be detected, is proportional to its magnitude. Thus, the sense of touch was the first sense that has been studied with a rigorous scientific method. Despite the fact that several prominent researchers have worked on the sense of touch since, over the years this picture has dramatically changed. In recent years, most perceptual research was devoted toward the study of vision and audition. Therefore, compared to the sense of touch today vision and audition are far better understood. This shift in focus from touch to vision and audition was boosted by the development and the availability of novel sensing and display technology. For vision, the development of cameras to record visual information and the progress in display technology from the first simple cathode ray tubes to high-fidelity 3-D virtual-reality theaters enabled us to generate and manipulate visual stimuli used in perceptual experiments in a very fine-controlled manner. Similarly, microphones and loudspeakers enabled us to record and display sounds with high-fidelity 3-D surround capability. In contrast to this, touch sensors and display devices are still in their infancy. This is most likely due to the highly multidimensional nature of the sense of touch and the perceptions originating from this sense. Touch can be considered multidimensional as it comprises kinesthetic as well as tactile inputs. Furthermore, touch is multidimensional as it involves not only sensations based on force and pressure distributions on the skin, but it includes sensations based on temperature as well as pain. Most importantly, however, touch is multidimensional because haptic perceptions are not only based on passively receiving information, but they are formed through interaction thereby actively gathering information; information that is used for recognizing and manipulating objects. When interacting with objects the central nervous system has to efficiently control all the many degrees of freedom (DoF) inherent in the biomechanical structure of our bodies, particularly the arms and the hands. Without haptic feedback this dexterous control quickly fails, as nicely demonstrated by Westling and Johansson [50]. This failure of precise control with the lack of haptic feedback can strikingly be demonstrated when trying to open a door using a key after an extensive snowball fight without gloves. Taken together, because of this multidimensional nature, touch is arguably the most complex sense to study.

Haptic technology today is still based on very crude force measurements, often measuring a single force vector for the interaction with objects, and limited to displaying position-dependent reaction forces to the finger or the entire hand, instead of differential force patterns, which would be necessary for inducing specific tactile sensation. Thus, typical haptic systems today mostly ignore temperature and pain sensations, tactile inputs, or the many DoF that are offered by the joints in our body. This is in gross contrast to our outstanding ability to haptically recognize objects, which is mostly based on tactile inputs, the interplay between all our fingers during exploration, and also the thermal properties of objects. For example, Lederman and Klatzky [51] showed that we use stereotypical exploratory procedures to efficiently recognize objects. These exploratory procedures critically involve all fingers of the hand; they are tuned to maximize tactile inputs, and to access the thermal properties of the objects. Equating the amount of information, Newell *et al.* [52] showed that haptic object recognition performance is not worse than that of vision and that humans use similar exploration strategies both for vision and touch [53]. Thus, if we want to make progress it is the great challenge of the years to come to increase the bandwidth of haptic display and sensing technology for it to be used as an efficient tool for simulations in telepresence and virtual reality scenarios. In turn, having such technology available would enable us to significantly further our understanding of human haptic perception similar to the fast progress made uncovering the perceptual principles underlying vision and audition by having such technology readily available.

III. HAPTIC TELEPRESENCE

The haptic sense is limited by the closeness between the subject and the object/person of interest since touch is necessary to allow perception. Therefore, to think about haptic interaction between remote individuals and objects might sound fantastic. However, with recent advances in haptic research, technology, and devices, it is possible for a person to physically participate in remote actions and also receive the according haptic feedback, as explicitly addressed in the next section.

A. Haptic Telepresence Technology

Haptic telepresence systems enable the human user to manipulate objects in remote environments and execute tasks without physically being there [3]. Aiming at the full immersion into the remote environment, the telepresence system is augmented with various displays providing multimodal sensor information. By multimodal we refer to the perceptual modalities of human beings, such as the visual, the auditory, and the haptic modality. The application field is broad ranging from underwater to space teleoperation and other hazardous, hardly accessible environments to tele/minimal invasive surgery and teleoperation in scaled environments (nano/micro/macromanipulation) [18], [54].

A typical haptic telepresence system consists of a human system interface (HSI), i.e., a robot able to display haptic interaction to the human, and a teleoperator (TO), i.e., the remote executing robot (see Fig. 1). Both are interconnected via a two-way communication link. While the human operator manipulates the HSI, it commands the motion of the teleoperator, which in turn interacts with the remote environment. The multimodal sensor data are fed back through the communication network and displayed to the human operator indicating that haptic communication is inherently bidirectional.

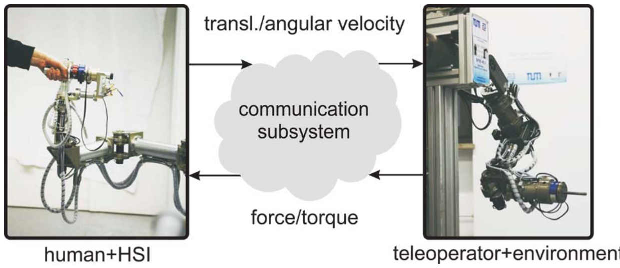

Fig. 1. ***A haptic telepresence system with multiple DoF consisting of a haptic interface device [55] and a teleoperator [56]. The haptic signals are transmitted over a communication network. In case of VR applications the block "teleoperator+environment" is replaced by a haptic rendering algorithm.***

Another important domain of haptic telepresence is virtual reality (VR) systems, which support physical interaction with a virtual environment. Such systems gain more and more relevance in education and training, e.g., as an experimental tool for neuroscience and cognitive science, for rapid prototyping in production, and in entertainment [10]. In haptic VR systems, the teleoperator and the remote environment in Fig. 1 are replaced by a haptic rendering engine. Many of the challenges regarding communication, control, and mechatronics are similar for haptic telepresence in terms of telerobotics and virtual reality and the respective results transfer from one domain to the other. Accordingly, in the following, an explicit distinction is made only where necessary.

A major challenge in the context of haptic telepresence systems is the design and control of haptic interfaces. While the measurement of haptic signals has become rather standard with today's advanced sensing technologies (position encoders, force and pressure sensors, etc.), the display of this information still remains a challenge both for the tactile as well as the kinesthetic feedback. Haptic interface devices behave like small robots that exchange mechanical energy with a user. While such an interface can be in contact with any part of the operator's body, hand interfaces have been the most widely used and developed systems to date [57].

Haptic devices differ in their kinematics including provided DoF, their output capability (in terms of displayed force/torque, velocity, and acceleration), their sensorial capability, their precision, backdrivability, and stiffness. The early commercialized haptic devices such as the Sensable Technologies PHANTOM [58] and the Force Dimension's Omega haptic device [59] are lightweight devices with a rather small workspace, high backdrivability, and good precision. Only few devices, e.g., the PHANTOM Premium and the Delta haptic device [60], show a moderate output capability. With increasing device size, as, for example, the HapticMaster [61], the output capability and workspace but also friction and inertia increase, which requires force sensing for compensation.

Innovative research prototypes are hyperredundant haptic interfaces such as the ViSHaRD10 [62] and mobile haptic telepresence systems for wide area interaction [63] including also bimanual interaction [64]. Other haptic devices such as the CyberGrasp/CyberGlove [65] are designed for multifingered interaction. More specialized devices are the DLR MIRO system [66], [67] developed to assist minimally invasive surgery. Recent advances in haptic actuator technologies are based on electroactive polymers, which enable area-based haptic sensing and skin surface actuation [68], [69]. For a comprehensive overview of haptic devices, the interested reader is referred to [70]; for guidelines on the development, control, and evaluation of kinesthetic haptic interfaces, see [71]; for aspects and challenges of tactile multipin displays, see [72].

B. Impact of Communication Delay

Haptic signals are exchanged between the operator side and the teleoperator side over a communication line, thereby closing a global control loop through the human operator, the HSI, the teleoperator, and the remote environment as interconnected subsystems. It is well known that communication-induced artifacts such as time delay and packet loss have detrimental effects on the performance and potentially the stability of the overall system as not only information but also energy is exchanged between the subsystems. In fact, even minor time delay may destabilize the system.

In order to account for communication artifacts, different haptic telepresence control architectures have been investigated. These architectures can be categorized based on the number and type of signals transmitted between the operator and the teleoperator side. In most approaches, motion (position or velocity) is sent to the teleoperator, and the environment force is transmitted back to the HSI; see also Fig. 2. There also exist architectures where velocity and force information is transmitted in both directions, i.e., four- and three-channel architectures. See [73] for the generic definition and a stability analysis. For illustration on how communication affects the overall system behavior, in the following, we will exemplarily study one of the most simple teleoperation control architectures: the so-called force-velocity control architecture.

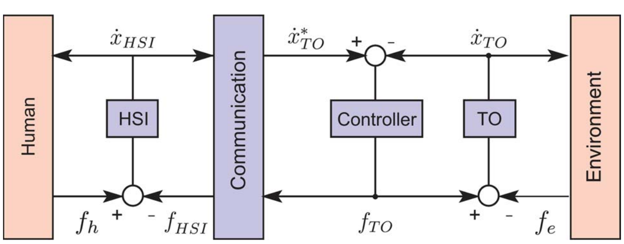

Fig. 2. *Standard control architecture for a haptic telepresence system.*

Example: We consider the 1-DoF haptic telepresence system architecture shown in Fig. 2. The human applies a force f_h to the HSI which reacts with the velocity $\dot{x}_{\mathrm{HSI}}$. The HSI velocity is transmitted over a communication network to the teleoperator side where it acts as the desired velocity $\dot{x}^*_{\mathrm{TO}}$ for the TO. The teleoperator moves with velocity $\dot{x}_{\mathrm{TO}}$, and the environment reacts in case of contact with a force f_e. The control effort f_{TO}, being a representative of the environment force, is transmitted to the HSI side where it is applied as motor torque of the HSI. The HSI is thus velocity controlled, while the TO is force controlled accounting for the name of force–velocity architecture. Assuming that the HSI/TO subsystems can be approximated around their operating points by local linear time-invariant mass-damper models the following dynamic equations describe the evolution of the HSI and TO position [15]:

$$\begin{aligned} m_{\mathrm{HSI}}\ddot{x}_{\mathrm{HSI}}(t) + b_{\mathrm{HSI}}\dot{x}_{\mathrm{HSI}}(t) &= -f_{\mathrm{HSI}}(t) + f_h(t) \\ m_{\mathrm{TO}}\ddot{x}_{\mathrm{TO}}(t) + b_{\mathrm{TO}}\dot{x}_{\mathrm{TO}}(t) &= f_{\mathrm{TO}}(t) - f_e(t) \end{aligned}$$

where m_i defines the mass, and b_i is the damping coefficient of the HSI and TO, $i \in \{\mathrm{HSI},\mathrm{TO}\}$. If the communication introduces latency, the transmitted signals arrive delayed at the receiver side

$$\dot{x}^*_{\mathrm{TO}}(t) = \dot{x}_{\mathrm{HSI}}(t - T_1) \quad f_{\mathrm{HSI}}(t) = f_{\mathrm{TO}}(t - T_2)$$

where $T_1, T_2 > 0$ are the time delays in the forward and backward channels, respectively.

In the following simulation study, it is demonstrated that even small time delay may jeopardize the stability of the overall system. We set the HSI and TO parameters to $m_{\mathrm{HSI}} = m_{\mathrm{TO}} = 1$ kg and $b_{\mathrm{HSI}} = b_{\mathrm{TO}} = 5$ Ns/m. The environment is represented by a linear spring with 10-N/m stiffness, and the human is represented by a linear time-invariant spring-damper system with 1-Ns/m damping and 30-N/m stiffness. The TO is velocity controlled with a proportional–integral controller with the 10-Ns/m P-gain and the 3000-N/m I-gain. The human applies a sinusoidal force on the HSI. Without time delay a satisfying position tracking between the HSI (solid line) and the TO (dashed line) is observed as displayed on the left side of Fig. 3. The same system with a time delay of just $T_1 = T_2 =$ 10 ms in each communication channel is already unstable.

A large variety of control architectures and designs have been proposed to guarantee stability in the presence of communication uncertainties such as time delay and packet loss; for a survey thereof, see [15] and [74]. A further challenge in the context of stability guarantees is the largely unknown dynamics of the human operator and

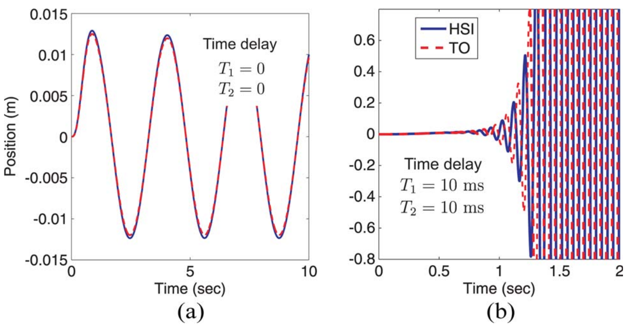

Fig. 3. ***Simulated HSI and TO position for the standard control architecture from Fig. 2 (a) without time delay and (b) with time delay $T_1 = T_2 = 10$ ms. Even a small time delay can lead to instability.***

the environment requiring particularly robust stability approaches.

One of the most successful robust stability-guaranteeing approaches is based on *passivity*, which is an energy-based concept characterizing the system by only analyzing its subsystems' input/output behavior and their interconnections. The dynamics of the HSI and the TO are usually passive or can be made passive by appropriate control; a trained human being can be considered to interact passively with passive environments [75]. If the communication system is passive as well, then stability of the overall interconnected system can be deduced. Depending on the communication characteristics particular control measures have to be employed to guarantee passivity such as the scattering transformation for constant but arbitrarily large constant time delay [76] and extensions thereof for time-varying time delay [77], [78] and packet loss [79], [80].

In summary, there is a strong dynamical coupling between the operator and teleoperator side subsystems via the communication channel. Any communication artifact will affect the performance and potentially the stability of the haptic telepresence system. Particularly challenging for stability is the time delay; even minor time delay can destabilize the overall system and appropriate control measures are required to compensate for that. Packet loss also has a detrimental effect on performance and potentially stability, in particular, when bursty loss patterns occur. As will be shown later, the effect of packet loss is dramatically amplified if the haptic data streams are compressed before transmission (see Sections IV and VI).

C. Performance Metrics and Evaluation

In order to evaluate the performance of haptic telepresence systems different criteria have been introduced representing measures of how close the ultimate goal of telepresence is achieved, namely, that the human operator feels like directly interacting with the remote real or virtual environment [81], [82]. The goal is achieved in a haptic telepresence system if the technical system between the human operator and the remote environment is *transparent*, which requires the positions and forces at the HSI and the teleoperator/environment to be equal [83] $x_{\mathrm{HSI}} = x_{\mathrm{TO}}, f_{\mathrm{HSI}} = f_e$, or alternatively, the mechanical impedance displayed to the human being equal to the environment impedance [73] $Z_{\mathrm{HSI}} = Z_e$. Ideal transparency in this sense is not achievable in practice, in particular, when communication effects such as time delay exist, as there is a fundamental tradeoff between robust stability and transparency [73], [84], [85]. Typical performance metrics derived from these transparency criteria are in terms of integrals (frequency or time) over position and/or force and/or impedance errors between the operator and the remote side.

A criterion which incoporates human haptic perception limits was first introduced in [74], [86], and [87], where the haptic telepresence system is called *perceived transparent* if the difference between the displayed impedance and the environment impedance is within the JND $Z_{\mathrm{HSI}} \in [Z_e - \mathrm{JND}, Z_e + \mathrm{JND}]$ (analogous for position and force errors). Both transparency and perceived transparency are objective quality metrics.

Other performance metrics evaluate the *task performance* of the operator, for example, in terms of task completion time and the operator's effectiveness in completing the assigned task like the sum of squared forces, e.g., in [88]. These metrics are in contrast to the concept of transparency task and operator dependent.

The feeling of *presence* is often measured through post-test questionnaires and subjective rating scales. In the *behavioral realism* approach, the difference of the user's reaction to the real and the mediated remote or virtual environment is evaluated based on reflexive or socially conditioned responses. Further investigated measures include physiological measures, postural responses, dual task measures, and social responses; see [89] for an overview on presence measurement.

In the context of telepresence, the experience of other humans (real or virtual) in terms of a *social presence* or *copresence* also plays an important role and has been investigated over the past years [10], [90].

IV. PERCEPTUAL CODING OF HAPTIC DATA

The transmission requirements of haptic information differ from those of audio and video in several important aspects. Haptic systems typically use a local 1-kHz control loop to overcome device dynamics and display high-frequency haptic effects to the user. Ideally this update rate should be maintained also across the communication channel. As latency can put the system stability at risk (see Section III-B) or at least deteriorate the system performance, the end-to-end delay should be kept as small as

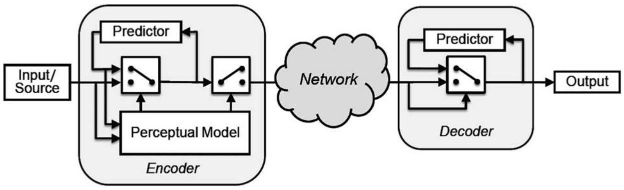

Fig. 4. *Overview of the perceptual haptic data reduction scheme presented in [91].*

possible. In this context, blockwise processing and retransmission of haptic samples is disfavored and haptic data packets are sent immediately once new sensor readings are available. The resulting high packet load on the network leads to substantial data overhead due to the transmission of packet header information. At a packet rate of 1 kHz, this can become a critical factor and overhead can even dominate compared to the actual payload data.

To address these issues, Otanez *et al.* [92] were first to propose a deadband-based data reduction scheme that specifically targets at a reduced packet rate for networked control. In their work, they compare new sensor readings to the most recently transmitted sample, and if the change is smaller than a predefined, fixed threshold, the sample is not transmitted. The proposed signal reconstruction at the receiver applies a hold-last-sample (HLS) algorithm. While the approach in [92] successfully reduces the packet rate, it does not explore the limitations of human perception. To this end, Hinterseer *et al.* [91] propose a perceptual coding scheme for haptic data; see Fig. 4. They deploy a prediction algorithm which estimates incoming haptic samples based on previously transmitted haptic information. In order to keep the reconstruction error imperceptible, a mathematical model of human haptic perception is employed. It allows for adaptively evaluating the quality of the predictor. As long as the difference between incoming and predicted haptic samples stays within the perception limits, no network transmissions are triggered. If the difference between incoming and predicted haptic samples exceeds the applied perception thresholds, additional signal information is sent over the network which updates the predictors at the encoder and decoder sides. In order to keep the introduced latency at a minimum, Hinterseer *et al.* [91] propose to use a zero-order predictor (hold-last-sample algorithm). In [93]–[96], predictors of higher order are used to improve the data reduction performance.

The compression of haptic data for both offline and online encoding has received significant attention during the last decade. Various lossy haptic compression schemes, which differ in their sampling, quantization, and entropy coding strategies, have been presented, for instance, in [97]–[102]. For a more detailed discussion of related work in the area of haptic data compression, we refer the interested reader to the *Short History of Haptic Data Reduction* in [26]. In the following, different perceptual models and compression techniques in the context of perceptual haptic data reduction for real-time haptic telepresence are discussed.

A. Perceptual Deadband-Based Data Reduction

The approach in [19], [91], [103], and [104] is based on Weber's law (see Section II-C). Applied in perceptual coding of haptic data it allows for keeping the introduced prediction errors in force and velocity signals below human haptic thresholds [103]. This scheme is described for haptic signals with 1-DoF by

$$\begin{array}{lll} \text{If: } |x(t') - x(t)| \leq k \cdot |x(t')| & \text{Do not transmit} & \\ \text{Else:} & \text{Transmit new value} & (4) \end{array}$$

where $x(t)$ is the current haptic sample, t' is the time when the most recent sample transmission happened, and k denotes the perceptual threshold parameter. We refer to the range of imperceivable changes as *deadband* in the following. Inspired by Weber's law, the deadband is adjusted for each transmitted sample, i.e., proportionally grows with the amplitude of the signal. The deadband parameter k has been identified in psychophysical studies [105] and has approximately the same size as the JND of the respective modality. A haptic telepresence system architecture including the deadband-based data reduction scheme is illustrated in Fig. 5. An impressive data reduction of up to 90% [105] without sacrificing the operator perceived transparency is achieved by this approach, and it has been extended also to time-delayed communication channels [106]. In the following example, the effect of the deadband-based perceptual coding approach is demonstrated in a simplified simulation.

Example: Deadband-Based Data Reduction: A sinusoidal signal, sampled at 1 kHz, is transmitted and the deadband approach is applied. The deadband parameter k in (4) is set to 20% resulting in the transmission of only 128 packets

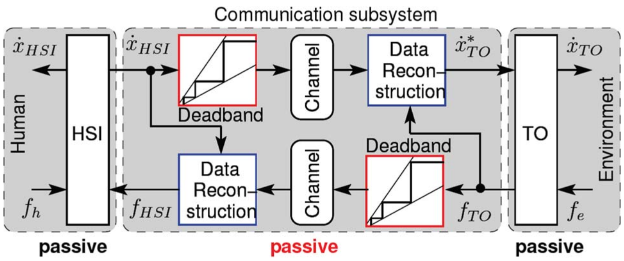

Fig. 5. *Example of a haptic telepresence system architecture with deadband-based data reduction. Figure adapted from [19].*

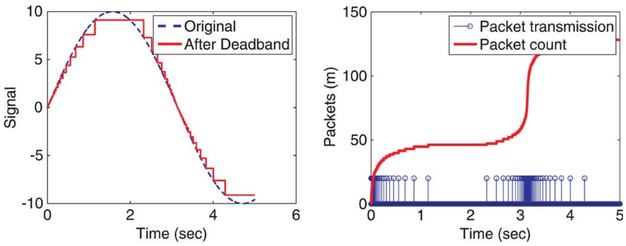

Fig. 6. ***Signal distortion with the deadband approach. A 20% deadband is applied resulting in a significant data reduction. Only 128 packets are transmitted compared to the case without perceptual data reduction where 5000 packets would be triggered.***

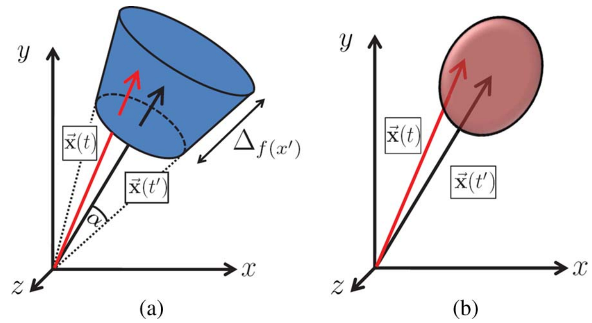

Fig. 7. ***(a) A nonisotropic discrimination zone (blue) and (b) a multidimensional deadband ellipse (red) formed by the last transmitted sample $\vec{\mathbf{x}}(t')$ (black arrow). The current vector $\vec{\mathbf{x}}(t)$ (red arrow) is not transmitted as it lies within the deadzone. Figure adapted from [111].***

compared to the case without perceptual data reduction where 5000 packets would be triggered. Consequently, for this signal, a data reduction of 97% is achieved. The signal before and after the deadband with an HLS-based reconstruction is shown in Fig. 6. We observe that the signal transmitted after the application of the deadband has minor artifacts, which can be even further reduced by using more advanced reconstruction strategies [91].

B. Velocity-Adaptive Perception Thresholds

Weber's law of JND is extended in [107] by exploiting the motion dependency of the human force perception. It is known that when a human operator interacts with a certain arm velocity, this reduces his/her force-feedback perception abilities [108], [109] allowing for further data reduction. This is captured in a simple relationship for the deadband factor

$$k_v = k + \alpha|\dot{x}(t)| \tag{5}$$

where k_v is the velocity-adaptive deadband parameter, which is now time varying, and k is the constant deadband parameter used in (4). The factor $\alpha > 0$ is identified in psychophysical experiments. The velocity-adaptive deadband approach proves to be successful at further reducing the packet rate (up to 30%) in experiments conducted in [107].

C. Multiple-DoF Haptic Data Reduction

In real-world teleoperation systems with multiple DoF, a haptic signal vector $\mathbf{x} \in \Re^n$ is transmitted instead of a scalar signal. Applying a psychophysical model to every single component of the representation is a straightforward extension, which, however, turns out to be very inefficient. The component with the smallest magnitude will trigger unnecessary transmissions, which might be in fact imperceivable. To address this issue, Hinterseer and Steinbach [110] propose an alternative approach: the so-called *multi-DoF "deadzone"* which is centered at the tip of the haptic sample vectors. In the 2-D case, this leads to a circular deadzone. Likewise, the 3-D case leads to a spherically shaped deadzone. In line with the 1-D deadband approach, the dimensions of the deadzone for the multidimensional case are defined by a fraction of the amplitude of the most recently transmitted haptic sample vector.

Interestingly, the assumption that the perception space of the human is isotropic turns out to be conservative as shown, for example, in [111]. It is found that the deadzone in a multidimensional space is nonisotropic; see, for example, the force discrimination deadzone in Fig. 7(a). It is defined by individual thresholds α and $\Delta_{f(x')}$, which describe independent perception thresholds for signal changes in direction and amplitude. Furthermore, more complex perceptual hypervolumes were found to apply for other modalities than only force; see [112] and the references therein, especially when two different stimuli are applied, e.g., a torque and a force; psychophysical studies still need to evaluate and exploit this field.

Another very flexible representation of the nonisotropic multi-DoF deadzone model is introduced in [113] using a deadzone hyperellipse [see Fig. 7(b)]. Instead of the deadband parameter k, the deadzone is represented by a positive–definite matrix $\mathbf{\Omega}$. Its diagonal elements indicate how much each component of the overall multidimensional signal contributes to the overall perception; its nondiagonal elements allow for possible masking effects to be captured, e.g., when a high amplitude force renders a small amplitude torque imperceivable. The multi-DoF perceptual data reduction algorithm extends the single-DoF algorithm (4) as follows:

If: $\|\mathbf{\Omega}(\vec{\mathbf{x}}(t') - \vec{\mathbf{x}}(t))\| \leq \|\vec{\mathbf{x}}(t')\|$	Do not transmit
Else:	Transmit new value.

If no masking effects are modeled, the deadzone matrix for n-dimensional haptic signals is diagonal

$$\text{with} \quad \mathbf{\Omega} = \begin{pmatrix} \frac{1}{k_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{k_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{k_n} \end{pmatrix} > 0. \tag{6}$$

Accordingly, the above multidimensional equation reduces to n single-DoF problems, and (4) is applied with k_i for each dimension $i = 1, \ldots, n$. With this technique, the problem of unnecessary transmissions triggered by signal components with small amplitudes is solved.

Data reduction and reconstruction introduce artifacts to the haptic signal within the global haptic closed loop and may potentially destabilize the overall system. One approach to guarantee stability is to perform the signal reconstruction at the receiver side such that it renders the communication subsystem in Fig. 5 passive [19], [86], [105]. Based on passivity arguments for the other subsystems (see also Section III-B), stability is deduced. In order to avoid conservatism, an optimization-based passive-rendering reconstruction technique is introduced in [113]. The proposed multi-DoF deadband approach with optimization-based reconstruction has been successfully validated in the multi-DoF telerobotic system shown in Fig. 1, resulting in a 30% more efficient data reduction (compared to the traditional Weber-based approach for each DoF) without impairing the transparency [113].

D. Event-Based Coding of Haptic Signals

Real-world haptic interactions with objects in our surroundings elicit a variety of haptic responses spanning the entire haptic perception bandwidth (of the order of a few hundred hertz) [114]. Tapping upon the surface of rigid objects generates sudden force transients followed by a nearly steady force balancing the force exerted by the human [115]. In comparison, the richness of feedback that we can invoke from virtual haptic models is quite limited. Haptic rendering with desktop haptic devices most commonly employs position-based resistive feedback, which is limited by the human manipulation bandwidth (of the order of a few hertz), due to the closed-loop nature of haptic interaction. Most conventional haptic rendering algorithms for virtual environments thus neglect high-frequency transients, essential to conveying physical information of the object realistically. To rectify this situation, the paradigm of event-based haptics has been proposed in [25] and [115]. Within this framework, discrete events of contact with an object described by the contact time and velocity can be used to trigger the display of high-frequency (HF) precomputed force histories in an open-loop manner. These HF contact transients

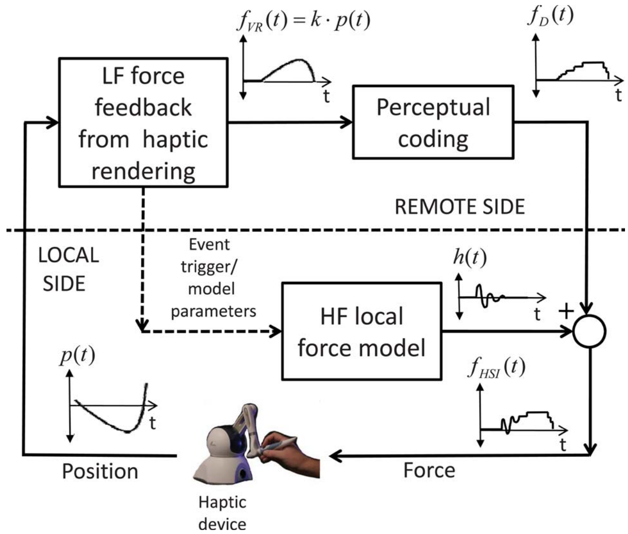

Fig. 8. ***Architectural overview of a virtual haptic interaction. Conventional low-frequency (LF) resistive feedback forces (computed based on a spring model with stiffness k) are augmented with high-frequency (HF) contact transients. The HF contact transients quickly remove the haptic device's momentum, slowing down penetration into the virtual object. The contact model for the transients is shifted to the client side for local rendering, thus preventing increased packet rates on the network on account of HF signal transmission (figure based on [115]).***

can also be described by parametric models easy for computation.

However, when transmitting such signals from the simulation server to a client over the network, the inherent high-frequency nature of contact transients results in increased sampling and packet rate requirements. Transmission delays occurring during communication make control-loop stability issues further critical when emitting high-frequency contact transients. To avoid additional packets being triggered by remotely superimposed contact transients, Kammerl *et al.* [116] propose to shift the contact model from the remote server to the local client side to enable local model-based haptic contact transient rendering (see Fig. 8). This is well supported by the open-loop nature of the proposed contact model. Thus, the model, triggered by event-of-contact messages received from the remote side, is deployed locally at the client side, resulting in high-fidelity haptic interactions along with efficient communication.

V. HAPTIC COMMUNICATION PROTOCOLS

A real-time haptic transport system should work with minimal protocol overhead and should be optimized with respect to the required high signal update rates. Furthermore, support for dejittering and stream synchronization would enable the haptic application to balance between disturbing additional latency and improved signal quality. In an IP-based packet-switched network scenario, the requirements for the haptic stream transport match to the widely applied RTP/UDP protocol. In [117] and [118], UDP-based data transport sessions for haptic telepresence have been successfully established and run across continents.

In addition to the transmission of the actual haptic transport streams, the exchange of system and session information is essential. Unlike in audio/video media connections where only a small set of system parameters like resolution, color depth, frame rate, codec, etc., is required to configure the system, systems for haptic telemanipulation vary in many aspects like the number of DoF, the number and type of devices, applied control architectures, workspace, sensors, and data representation. To enable a flexible connection between the haptic interfaces and devices, detailed knowledge of important system parameters and functional capabilities is essential. In contrast to the transmission of the haptic transport stream, communicating the session and system information is not restricted by hard delay constraints. Here, reliable signaling of the system description information and the system state updates are most important. In an Internet-based network scenario, the requirements for session control match to reliable TCP/IP-based remote procedure calls (RPCs). Fig. 9 illustrates the haptic session and transport streams typical to networked haptic telepresence.

Previous work by Tachi *et al.* [119] focused on autodescription of telerobotic configurations and dynamic teleoperation data types. Chat *et al.* [120] present a framework for haptic communication based on the MPEG-4 BIFS standard. In [121], the widely adopted Session Initiation Protocol (SIP) [122] is applied to haptic telemanipulation scenarios. During an initial call handshake, haptic system description information is exchanged and haptic codecs are negotiated according to the system capabilities and network parameters. After codec negotiation, RTP sessions are created for the audio, video, and an additional haptic transport stream. Interestingly, SIP provides a comprehensive architecture of standardized entities such as registry servers, redirect servers, and proxy servers, which provide important functionality for the haptic telepresence, such as name, address mapping, client localization, session forwarding and redirection, user management, capability negotiation, and security. However, these features are not unique to SIP, but are also common to other Internet session protocols like IAX2 and H.323.

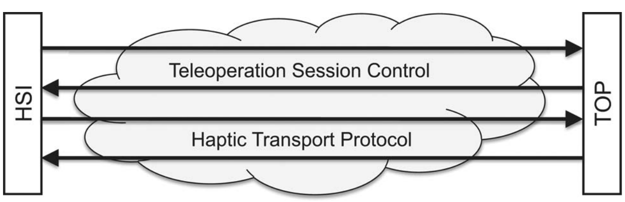

Fig. 9. ***Overview of session control and haptic data transport for networked haptic telepresence. TOP stands for teleoperator.***

VI. ERROR-RESILIENT HAPTIC COMMUNICATIONS

Haptic communication in the presence of packet loss suffers from erroneous input signals for the local control loops. If the packet loss happens on the forward path (from the operator to the teleoperator), this may lead to wrong position or velocity target values and hence a mismatch between the operator's commands and the resulting movement of the end effector. In some cases, this can cause unexpected contact with the environment or loss of contact and hence inconsistent force feedback to the operator. If the loss happens on the backward path, wrong force feedback values are displayed which in turn then affect the human's position commands. Due to the bidirectional nature of haptic communication and the global control loop that encompasses the human, the HSI, the communication, the end effector, and the environment, losses on one path also influence the values transmitted in the opposite direction. The stronger the signals are compressed, the more vulnerable the transmitted bitstream becomes against losses. For predictive coding approaches (e.g., motion compensated prediction in video coding) lost information leads to the infamous error propagation problem. This is also true for the haptic data reduction approach described in Section IV.

While for compressed audio and video the typical artifacts caused by packet loss are well understood, the impact of lost packets on compressed haptic data streams has been addressed only recently. The same holds for error-resilient haptic encoding. In audiovisual communication, error-resilient encoding, which deliberately introduces redundancy during the encoding process in order to improve the robustness against transmission errors, is a well-studied topic (see, e.g., [123] for an overview of error-resilient video encoding approaches). Again, for haptic communication, this has only recently been addressed. Brandi *et al.* study [124] both the impact of packet losses on the quality of experience during remote physical interaction with objects in a virtual environment and error-resilient encoding for haptic data streams.

A. Packet-Loss-Induced Artifacts

When using the perceptual data reduction scheme explained in Section IV in combination with predictive coding, haptic samples are selected to be either dropped or transmitted. The chosen samples to be transmitted are the ones that represent a perceivable change compared to their prediction. Moreover, these samples also assume an important role in predicting future samples. Hence, whenever sent samples are lost on the network, both

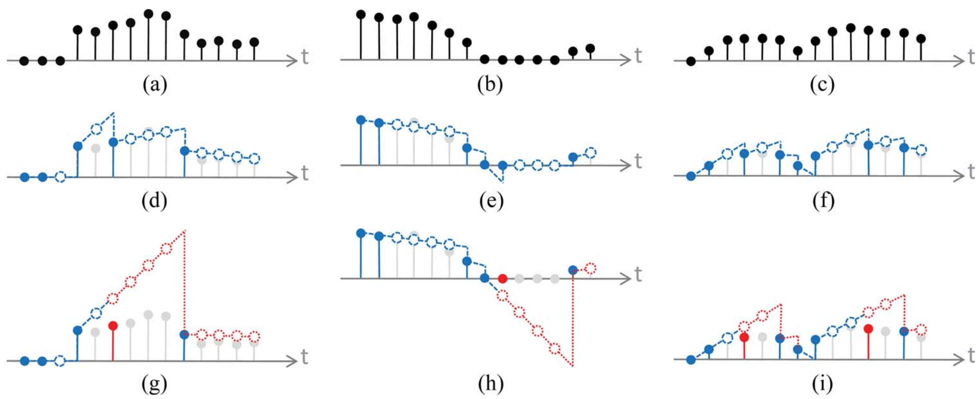

Fig. 10. ***Haptic artifacts due to packet losses for a perceptual deadband-based haptic encoder with linear prediction. Images (a)–(c) show the acquired haptic samples (black) at the encoder. Images (d)–(f) illustrate the predicted samples (dashed blue) and the update samples (blue) at the encoder. Images (g)–(i) depict the correctly predicted samples (blue), the lost samples (red), and the incorrectly predicted samples (dotted red). In the second and third rows, the original samples (light gray) are shown only for comparison. The bouncing artifact can be seen in (g) where a strong force is displayed pushing the end effector away from the contacted object. The glue effect can be observed in (h) in which a strong opposite force attracts the end effector toward the object. The roughness artifact is illustrated in (i) where mild forces are displayed in sequence provoking a granular texture sensation [124].***

encoder and decoder critically run out of synchronization and unexpected signal predictions occur displaying undesirable signals to both operator and teleoperator.

Three main artifacts due to packet losses were observed and explained in [124]. Although all of these artifacts are the result of wrong sample predictions at the decoder, the moment and the frequency with which the losses occur strongly influence the predictions and thus the displayed signal.

The *bouncing* artifact is characterized by a strong force displayed to the user in the occasion of a contact event between the end effector and an object. This artifact can be observed in Fig. 10(g).

The *glue effect* imposes an undesirable strong force in a reversed manner resulting in an attraction force toward the object. This artifact is illustrated in Fig. 10(h).

The *roughness* artifact provides an erroneous sensation of being in contact with a significantly rough surface. This artifact is illustrated in Fig. 10(i).

B. Error-Resilient Encoding

The error-resilient perceptual coding for networked haptic interactions proposed in [124] not only takes into account the psychophysical limitations of the human haptic perception to continue reducing the amount of samples to be transmitted such as seen in Section IV, but it also considers the packet loss probability on the communication channel to estimate the influence of lost packets on the reconstructed signal. In this way, a binary tree (inspired by [125]) enumerates at the encoder the cases of successful and unsuccessful transmissions and the respective resulting predictions. Whenever the predicted signals combined deviation is likely to disturb the system, redundancies are added to the transmission.

During the haptic session, update samples are also kept in the encoder buffer, which is consulted at every moment to calculate the possible samples combinations resulting in different predictions. On this occasion, three types of thresholds are proposed in [124] to be compared to the current estimates, namely, *expected deviation*, *sum of probabilities*, and *maximum deviation*. Each of these thresholds accounts for different impacts on the displayed signal due to the losses. As an example, if one wants to avoid that the displayed signal deviates more than a certain predefined amount—no matter how likely it is to occur—the *maximum deviation* trigger can be employed. In the opposite case, if one wants to minimize the overall occurrence probability of incorrect predictions—no matter how much they exceed the detectable perceptual thresholds—they should use the *sum of probabilities* trigger. If one wants to combine both packet loss occurrence probability and relative deviation, the *expected deviation* trigger can be applied. A simplified block diagram depicting the error-resilient haptic data reduction approach can be seen in Fig. 11.

The update samples need to be kept at the encoder buffer until packet acknowledgments (ACK) arrive certifying which were the last packets received by the decoder. As a result, no estimations concerning the samples previous to the acknowledged sample need to be considered anymore and the tree can be rebuilt from that instant onwards.

Brandi *et al.* recently further improved the approach in [124] by proposing a low-complexity error-resilient haptic compression scheme [126] where the number of states to be estimated grows linearly instead of exponentially as in [124].

The error-resilient haptic data reduction schemes proposed in [124] and [126] showed to be very efficient radically minimizing the perceivable artifacts adding

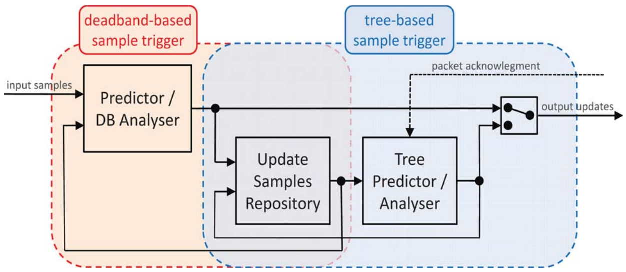

Fig. 11. ***Schematic diagram of the error-resilient haptic data reduction scheme at the encoder. The red box on the left represents the data reduction approach described in Section IV. The blue right box involves the proposed approach in [124] wherein update samples are saved and used to determine a set of predictions, which can be compared on-the-fly to predefined thresholds and trigger additional packets whenever the signal distortion must be minimized.***

redundancies in the haptic communication while still presenting compression ratios comparable to the standalone deadband-based data reduction approach in Section IV.

VII. OBJECTIVE QUALITY METRICS FOR HAPTIC COMMUNICATION

In multimedia communication applications, the sink of information is typically the human who consumes and evaluates the media content. Assessing the user experience, often also referred to as QoE, requires by definition to explicitly involve the human observer in the evaluation process, or alternatively, to replace the user by a mathematical model of human perception. The former leads to subjective tests which are time consuming, expensive, hard to reproduce, and require carefully controlled experimental setups, but lead to reliable judgments of user satisfaction. The latter leads to objective quality metrics, which are often questioned in terms of how well they correlate with the actual user satisfaction over a wide range of users, media content, and test conditions.

For haptic communication, as of today, quality evaluation is almost exclusively performed via subjective tests. Objective quality evaluation strategies for haptic communication hence have a huge potential to propel advances in haptic communications.

The starting point for haptic objective quality evaluation is fundamentally different when compared to objective quality evaluation for audio and video as a result of the bidirectional nature of physical interaction. First, in a telemanipulation session, there is no original, undistorted signal we could compare the recorded haptic signals (position and force samples) to. Physical interaction in a remote environment is only possible with the human in the loop and hence even if we define a specific task (e.g., asking the operator to follow a predefined trajectory with the end effector), unpredictable manipulative actions performed by the user lead to position changes, which are reflected back through the remote environment as force feedback, which in turn influence the actions of the human and hence again his position commands, and so on and so forth. The same human operator will not generate identical haptic signals when performing the same task twice. In other words, it is the unpredictable behavior of the human in this globally closed control and interaction loop that leads to irreproducible haptic signal sequences, which makes a sample-by-sample comparison of haptic signals from different runs impossible. This situation becomes even worse when the operator performs the same task twice, once with and once without haptic compression switched on. The reaction of the human to perceivable compression artifacts influences the position commands and hence in turn through the interaction with the environment the force feedback, and so on.

For this reason, QoE evaluation for haptic interaction has been mainly restricted to subjective tests. Basdogan *et al.* evaluate the role of haptic feedback for human–human and human–machine interactions in collaborative virtual environments in [10]. They study the impact of force feedback on task performance and show that support for haptic communication leads to an improved feeling of togetherness. The results in [10] are obtained via subjective tests.

First steps toward an objective QoE evaluation of haptic interaction with virtual environments have been introduced in [127]. The authors combine relevant, measurable QoS parameters with user-experience-related parameters (perception measures, rendering quality, physiological measures, and psychological measures) and map these to a QoE value using a fuzzy inference system. Users, however, still have to provide their quality estimation for selected user-experience-related parameters and hence the approach still heavily relies on user involvement.

In [128], Ruffaldi *et al.* address the need for objective validation and comparison of haptic rendering algorithms. They provide force data sets for real-world objects together with corresponding computer graphics models and use these to compare the force response generated by a haptic rendering system with the previously measured physical forces.

Sargardia *et al.* propose in [129] an objective quality measure for the evaluation of haptic rendering algorithms. To this end, the forces and torques generated by the haptic rendering algorithms while moving an object along specified trajectories in a virtual environment are compared to analytically computed reference values. While this is a signal-based quality evaluation approach where a distorted signal is compared with an undistorted reference signal, it focuses exclusively on haptic rendering algorithms and does not include in the evaluation the impact of human action, the human–system interface, and limits of human haptic perception.

Based on the notion of perceived transparency (see also Section III-C for the defnition) Hirche *et al.* [74], [86],

[87] derived an objective quality metric for haptic telepresence systems. The performance metric measures the transparency under consideration of human haptic perceptual limits and serves for the analysis of the effect of control and communication parameters on the perceived transparency. For simple settings, even closed-form solutions are obtained, for example, for the relationship between the displayed stiffness, the environment stiffness, and the time delay in the communication channel. In addition, it allows for the optimization of control and communication parameters with respect to stability robustness under the constraint that the system remains perceived transparent.

Chaudhari *et al.* [130] present a fully automatic objective quality evaluation framework for haptic communication. The main idea in [130] is to replace the real human with its unpredictable behavior by a virtual human with predictable behavior (i.e., a computer model of the operator). The approach uses the human operator model for compensatory manual control in tracking tasks described in [131] and is composed of partial models for the central nervous system and the neuromuscular arm of the operator. This model has been used by Penin *et al.* [132] for the simulation of a teleoperation system with kinesthetic feedback. Additionally, Chaudhari *et al.* [130] also introduce a model of the human arm coupled to the haptic interaction device (HSI) such that the entire telemanipulation experiment can be executed as a software simulation. Fig. 12 illustrates this concept. Some fundamental differences between the objective quality evaluation (OQE) approaches in [86], [87], and [130] are highlighted in Table 1.

In [130], first the interaction with the environment is simulated without compression of the haptic data. The recorded haptic sample sequence represents the undistorted reference signal for the quality evaluation. In a second run, the haptic compression module is switched on and this time the recorded haptic samples represent the distorted signal. These two sequences can now be compared sample by sample in the sense of a full-reference quality metric. By this, the undesired intersubject and intertestrun variabilities are removed. The approach in [130] goes one step further and proposes to compare the recorded haptic samples in the perceptual domain to compensate for known limitations of human haptic perception. The reference signal and the distorted signal are transformed into the perceptual domain using Weber–Fechner's law [133]

$$S = c \cdot \ln\left(\frac{I}{I_0}\right) \tag{7}$$

with S being the sensation the human experiences as a function of the applied haptic stimulus, c being a scaling constant that needs to be determined experimentally, I being the magnitude of the applied stimulus, and I_0 being the absolute detection threshold. Chaudhari *et al.* [130] then define for the reference sensation S and the distorted sensation $\hat{S}$ over a sequence of N values the *perceptual mean squared error* (PMSE) as

$$\text{PMSE} = \frac{1}{N}\sum_{i=0}^{N-1}\left(S(i) - \hat{S}(i)\right)^2 \tag{8}$$

which can be then written in terms of the actual haptic values as

$$\text{PMSE} = \frac{c^2}{N}\sum_{i=0}^{N-1}\ln\left(\frac{I(i)}{\hat{I}(i)}\right). \tag{9}$$

To the best of our knowledge, the approach in [130] is the first approach for an objective quality assessment for

Table 1 Comparison of OQE Approaches: Approach Based on Perceived Transparency [86], [87] and Human Haptic Action/Perception Model [130]

	perceived transparency [86], [87]	**human model** [130]
human action model	no, out of scope for this work	yes, for compensatory tracking tasks
haptic perception model	yes, exploits the concept of JNDs for stiffness perception to derive wave impedance values for haptic transparency	yes, a psychophysical model for measuring quality degradation based on the Weber-Fechner law
fully automatic	no	yes

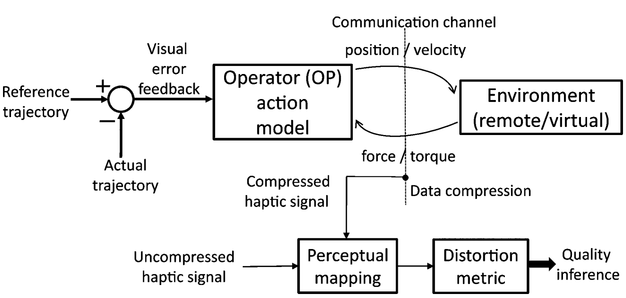

Fig. 12. ***Overview of the objective quality assessment framework introduced in [130]. The human operator model is identified via subjective tests. It is then simulated with a given input reference trajectory producing haptic signals. The uncompressed and compressed versions of the haptic signals are compared in the perceptual domain using the PMSE metric [see (8)].***

haptic interactions in virtual environments which is fully automatic and which explicitly considers limits of human haptic perception.

VIII. SELECTED CHALLENGES FOR THE FUTURE

Haptic communications is a relatively young field and many challenges need to be addressed in the coming years. The following list summarizes some of these challenges without claiming completeness.

A. Sensing and Actuation

Most haptic interaction devices in use today work with single-point end effectors. The human, however, perceives touch sensations across the entire body. Although the spatial resolution varies greatly (e.g., between the finger tips and the back) a more distributed or area-based sensing and actuation is required to reach the next level of immersion. Although first prototypes for an artificial skin with built-in area-based sensing and actuation exist, progress in this area is eagerly awaited.

B. Perceptual Coding

The perceptual deadband-based haptic data reduction scheme presented in Section IV is based on Weber's law and a haptic signal predictor. Further gains can be expected by integrating additional findings from psychophysics. Particularly, the integration of cross-modal dependencies of the visual, audio, and haptic modality promises significant improvements in detecting and removing irrelevant signal content. Also very relevant in this context is the extension of the models of human perception from today's standard stationary psychophysical quantities (e.g., Weber's law) toward models that capture the characteristics of human dynamic haptic perception. Additionally, the performance of perceptual haptic data reduction can be improved by extending the signal predictor. The prediction of haptic signals is not necessarily required to be based on statistical knowledge of the signal trajectories. By detecting points of contact in the haptic signals, we are able to estimate geometric shape and physical properties of remote objects in contact [134]. Integrated into the perceptual coding scheme, the haptic force-feedback signals can then be predicted using haptic rendering. In this context, the integration of remote depth sensor information and the analysis of visual feedback can further support such geometric model estimation.

C. Objective Quality Evaluation

While the approach by Chaudhari *et al.* [130] discussed in Section VII represents an important step toward objective quality evaluation for haptic interaction, it needs to be extended in several directions. More sophisticated and comprehensive models for human manual control action and haptic perception need to be incorporated. The approach is currently designed for haptic interactions in virtual environments and needs to be extended for real-world telemanipulation systems. Finally, the QoE for the auditory and visual feedback needs to be considered and fused with the QoE measure for the haptic modality into a joint QoE metric.

D. Error Resiliency and Stability

Stability and performance-guaranteeing approaches for haptic telepresence systems have been studied for lossy and time-delayed communication but only for the case without haptic compression. A stability analysis for a combination of these adverse channel conditions with haptic compression has not been addressed so far but would be of particular importance for wireless haptic communications. Also, the joint design of control architectures and communication protocols needs to be addressed in future work.

E. Multiplexing of Multimodal Sensory Information

The performance of a telepresence system with manipulation capability critically depends on the exchange of multimodal sensory information between the operator and the teleoperator. As each of the involved modalities is characterized by individual sampling and data rate requirements as well as latency constraints, the development of a multiplexer dedicated for multimodal data streams becomes highly relevant. By investigating and integrating upper and lower latency bounds for each modality as well as incorporating human cross-modal temporal integration [135], a high degree of immersiveness and transparency can be achieved even in scenarios with capacity-limited communication channels. First work toward multimodal multiplexing can be found in [136] and [137].

F. Shared Cooperative Virtual Haptic Environments

Collaborative haptic virtual environments enable multiple users to collaborate toward a common objective. In this context, a low-latency communication system for distributing the multimodal sensory information is of fundamental importance. As soon as users are haptically coupled during events of joint manipulation, the network load and communication latency becomes a critical factor. This can, for instance, be addressed by detecting joint haptic interaction and decentralizing the system by local grouping of active users.

IX. CONCLUSION

This paper presents recent advances in the area of haptic communications with a special focus on communication challenges for haptic telepresence and telemanipulation. We show that the communication requirements differ substantially from those in audiovisual communication as a result of the human-in-the-loop nature of haptic interaction. Time-delayed communication and packet loss put the system stability at risk and need to be handled appropriately using specific control architectures and error-resiliency mechanisms. We show that Internet-based haptic communication suffers from high packet rates, and present perceptual haptic data reduction approaches which address this issue and lead to substantial packet rate reductions without noticeable impact on the system transparency. We further argue that accelerated progress in the field of haptic communications is hindered by the lack of objective quality metrics, and present our own recent proposal on how to design such a metric. We complement our discussion of recent technological advances with an overview of the properties and limits of human haptic perception. We also address multimodal integration of audio, video, and haptics as haptic communications cannot be treated independently from audiovisual communication. Truly immersive telepresence requires the availability of all three modalities with a proper understanding of their cross-modal interaction. ■

Acknowledgment

Many people have contributed to the work presented in this paper over the past years. The authors kindly acknowledge the contributions made by S. Chaudhuri, P. Hinterseer, H. King, A. Peer, M. Rank, and B. Stanczyk.

REFERENCES

[1] Cisco Systems, Inc., *TelePresence*, 2011. [Online]. Available: http://www.cisco.com/en/US/products/ps7060/index.html

[2] Hewlett-Packard Development Company, L.P., *Introducing Halo*, 2011. [Online]. Available: http://www.hp.com/halo/introducing.html

[3] W. R. Ferrell and T. B. Sheridan, "Supervisory control of remote manipulation," *IEEE Spectrum*, vol. 4, no. 10, pp. 81–88, Oct. 1967.

[4] K. M. C. Loeb, "Human factors and behavioral science: Membrane keyboards and human performance," *Bell Syst. Tech. J.*, vol. 62, no. 6, pp. 1733–1749, Jul. 1983.

[5] T. B. Sheridan, *Telerobotics, Automation, and Human Supervisory Control*. Cambridge, MA: MIT Press, 1992.

[6] C. A. Lawn and B. Hannaford, "Performance testing of passive communication and control in teleoperation with time delay," in *Proc. IEEE Int. Conf. Robot. Autom.*, Atlanta, GA, May 1993, vol. 3, pp. 776–783.

[7] G. Hirzinger, B. Brunner, J. Dietrich, and J. Heindl, "ROTEX—The first remotely controlled robot in space," in *Proc. IEEE Int. Conf. Robot. Autom.*, San Diego, CA, May 1994, vol. 3, pp. 2604–2611.

[8] M. Srinivasan and C. Basdogan, "Haptics in virtual environments: Taxonomy, research status, and challenges," *Comput. Graph.*, vol. 21, no. 4, pp. 393–404, Jul.–Aug. 1997.

[9] J. T. Dennerlein, D. B. Martin, and C. Hasser, "Force-feedback improves performance for steering and combined steering-targeting tasks," in *Proc. ACM SIGCHI Conf. Human Factors Comput. Syst.*, 2000, pp. 423–429.

[10] C. Basdogan, C.-H. Ho, M. A. Srinivasan, and M. Slater, "An experimental study on the role of touch in shared virtual environments," *ACM Trans. Computer-Human Interaction*, vol. 7, no. 4, pp. 443–460, Dec. 2000.

[11] M. L. McLaughlin, G. Sukhatme, and J. Hespanha, *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*. Upper Saddle River, NJ: Prentice-Hall, 2001.

[12] T. Debus, T. Becker, P. Dupont, T. Jang, and R. Howe, "Multichannel vibrotactile display for sensory substitution during teleoperation," in *Proc. SPIE Int. Symp. Intell. Syst. Adv. Manuf.*, Newton, MA, Oct. 2001, pp. 42–49.

[13] A. Cockburn and S. Brewster, "Multimodal feedback for the acquisition of small targets," *Ergonomics*, vol. 48, no. 9, pp. 1129–1150, 2005.

[14] G. Tholey, J. Desai, and A. Castellanos, "Force feedback plays a significant role in minimally invasive surgery," *Ann. Surgery*, pp. 102–109, Jan. 2005.

[15] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, Dec. 2006.

[16] G. Robles-De-La-Torre, "The importance of the sense of touch in virtual and real environments," *IEEE Multimedia*, vol. 13, no. 3, pp. 24–30, Jul.–Sep. 2006.

[17] A. El Saddik, "The potential of haptics technologies," *IEEE Instrum. Meas. Mag.*, vol. 10, no. 1, pp. 10–17, Feb. 2007.

[18] M. Ferre, M. Buss, R. Aracil, C. Melchiorri, and C. Balaguer, *Advances in Telerobotics: The Human Role in Telerobotics*, 1st ed. New York: Springer-Verlag, 2007.

[19] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss, "Towards deadband control in networked teleoperation systems," in *Proc. IFAC World Congr.*, Prague, Czech Republic, Jul. 2005, DOI: 10.3182/20050703-6-CZ-1902.01414.

[20] S. Hirche and M. Buss, "Human perception oriented control aspects of networked telepresence and teleaction systems," in *Proc. Int. Conf. Instrum. Control Inf. Technol.*, Okayama, Japan, 2005, pp. 3430–3435.

[21] C. Basdogan, C.-H. Ho, M. A. Srinivasan, and M. Slater, "An experimental study on the role of touch in shared virtual environments," *ACM Trans. Computer-Human Interaction*, vol. 7, pp. 443–460, Dec. 2000.

[22] C. Basdogan, C. Ho, M. Slater, and M. Srinivasan, "The role of haptic communication in shared virtual environments," in *Proc. 3rd PHANToM Users Group Workshop*, Dedham, MA, pp. 85–89, Oct. 1998.

[23] J. Norman, H. Norman, A. Clayton, J. Lianekhammy, and G. Zielke, "The visual and haptic perception of natural object shape," *Attention Perception Psychophys.*, vol. 66, pp. 342–351, 2004.

[24] M. Heller, "Visual and tactual texture perception: Intersensory cooperation," *Attention Perception Psychophys.*, vol. 31, pp. 339–344, 1982.

[25] K. J. Kuchenbecker, J. Fiene, and G. Niemeyer, "Improving contact realism through event-based haptic feedback," *IEEE Trans. Vis. Comput. Graph.*, vol. 12, no. 2, pp. 219–230, Mar./Apr. 2006.

[26] E. Steinbach, S. Hirche, J. Kammerl, I. Vittorias, and R. Chaudhari, "Haptic data compression and communication," *IEEE Signal Process. Mag.*, vol. 28, no. 1, pp. 87–96, Jan. 2011.

[27] M. O. Ernst and H. H. Bülthoff, "Merging the senses into a robust percept," *Trends Cogn. Sci.*, vol. 8, pp. 162–169, 2004.

[28] M. O. Ernst and M. S. Banks, "Humans integrate visual and haptic information in a statistically optimal fashion," *Nature*, vol. 415, no. 6870, pp. 429–433, Jan. 2002.

[29] D. Alais and D. Burr, "The ventriloquist effect results from near optimal crossmodal integration," *Current Biol.*, vol. 14, pp. 257–262, 2004.

[30] K. Drewing and M. Ernst, "Integration of force and position cues for shape perception through active touch," *Brain Res.*, vol. 1078, no. 1, pp. 92–100, Feb. 2006.

[31] K. Drewing, T. Wiecki, and M. Ernst, "Material properties determine how force and position signals combine in haptic shape perception," *Acta Psychologica*, vol. 128, no. 2, pp. 264–273, Jun. 2008.

[32] G. Robles-de-la Torre and V. Hayward, "Force can overcome object geometry in the perception of shape through active touch," *Nature*, vol. 412, pp. 445–448, 2001.

[33] D. C. Knill and J. A. Saunders, "Do humans optimally integrate stereo and texture information for judgments of surface slant?" *Vis. Res.*, vol. 43, pp. 2539–2558, 2003.

[34] J. M. Hillis, S. Watt, M. S. Landy, and M. S. Banks, "Slant from texture and disparity cues: Optimal cue combination," *J. Vis.*, vol. 4, pp. 1–24, 2004.

[35] H. B. Helbig and M. O. Ernst, "Optimal integration of shape information from vision and touch," *Exp. Brain Res.*, vol. 166, pp. 595–606, 2007.

[36] N. Roach, J. Heron, and P. Mcgraw, "Resolving multisensory conflict: A strategy for balancing the costs and benefits of audio-visual integration," *Proc. Biol. Sci.*, vol. 273, pp. 2159–2168, 2011.

[37] L. Shams, Y. Kamitani, and S. Shimojo, "Visual illusion induced by sound," *Cogn. Brain Res.*, vol. 14, pp. 147–152, 2002.

[38] J.-P. Bresciani, F. Dammeier, and M. O. Ernst, "Vision and touch are automatically integrated for the perception of sequences of events," *J. Vis.*, vol. 6, pp. 554–564, 2006.

[39] H. McGurk and J. MacDonald, "Hearing lips and seeing voices," *Nature*, vol. 264, pp. 746–748, 1976.

[40] S. Gepshtein, J. Burge, M. S. Banks, and M. O. Ernst, "Optimal combination of

vision and touch has a limited spatial range," *J. Vis.*, vol. 5, pp. 1013–1023, 2005.

[41] K. P. Körding, U. Beierholm, W. Ma, S. Quartz, J. Tenenbaum, L. Shams, and O. Sporns, "Causal inference in multisensory perception," *PLoS ONE*, vol. 2, no. 9, Sep. 2007e943.

[42] R. B. Welch, *Perceptual Modification: Adapting to Altered Sensory Environments*. New York: Academic, 1978.

[43] W. Fujisaki, S. Shimojo, M. Kashino, and S. Nishida, "Recalibration of audiovisual simultaneity," *Nature Neurosci.*, vol. 7, pp. 773–778, 2004.

[44] V. Harrar and L. R. Harris, "Simultaneity constancy: Detecting events with touch and vision," *Exp. Brain Res.*, vol. 166, no. 3, pp. 465–473, 2005.

[45] J. Vroomen, M. Keetels, B. de Gelder, and P. Bertelson, "Recalibration of temporal order perception by exposure to audio-visual asynchrony," *Cogn. Brain Res.*, vol. 22, pp. 32–35, 2004.

[46] M. Di Luca, T.-K. Machulla, and M. O. Ernst, "Recalibration of multisensory simultaneity: Crossmodal transfer coincides with a change in perceptual latency," *J. Vis.*, vol. 9, no. 12, pp. 1–16, 2009.

[47] R. J. Baddeley, H. A. Ingram, and R. C. Miall, "System identification applied to a visuomotor task: Near-optimal human performance in a noisy changing task," *J. Neurosci.*, vol. 23, pp. 3066–3075, 2003.

[48] J. Burge, M. O. Ernst, and M. S. Banks, "The statistical determinants of adaptation rate in human reaching," *J. Vis.*, vol. 8, no. 4, pp. 1–19, 2008.

[49] E. Weber, *Die Lehre vom Tastsinn und Gemeingefuehl, auf Versuche gegruendet*. London, U.K.: Verlag Friedrich Vieweg und Sohn, 1978.

[50] G. Westling and R. S. Johansson, "Factors influencing the force control during precision grip," *Exp. Brain Res.*, vol. 53, no. 2, pp. 277–284, 1984.

[51] S. J. Lederman and R. L. Klatzky, "Hand movements: A window into haptic object recognition," *Cogn. Psychol.*, vol. 19, no. 3, pp. 342–368, Jul. 1987.

[52] F. N. Newell, M. O. Ernst, B. S. Tjan, and H. H. Bülthoff, "Viewpoint dependence in visual and haptic object recognition," *Psychol. Sci.*, vol. 12, pp. 37–42, Jan. 2001.

[53] M. O. Ernst, C. Lange, and F. N. Newell, "Multisensory recognition of actively explored objects," *Can. J. Exp. Psychol.*, vol. 61, no. 3, pp. 242–253, Sep. 2007.

[54] T. Sato, J. Ichikawa, M. Mitsuishi, and Y. Hatamura, "Micro teleoperation system concentrating visual and force information at operators' hand," in *Experimental Robotics III*, vol. 200, T. Yoshikawa and F. Miyazaki, Eds. Berlin, Germany: Springer-Verlag, 2007, pp. 353–365.

[55] A. Peer and M. Buss, "A new admittance-type haptic interface for bimanual manipulations," *IEEE/ASME Trans. Mechatron.*, vol. 13, no. 4, pp. 416–428, Aug. 2008.

[56] M. Buss, A. Peer, T. Schauss, N. Stefanov, U. Unterhinninghofen, S. Behrendt, J. Leupold, M. Durkovic, and M. Sarkis, "Development of a multi-modal multi-user telepresence and teleaction system," *Int. J. Robot. Res.*, vol. 29, no. 10, pp. 1298–1316, Sep. 2010.

[57] K. Salisbury, F. Conti, and F. Barbagli, "Haptic rendering: Introductory concepts," *IEEE Comput. Graph. Appl.*, vol. 24, no. 2, pp. 24–32, Mar.–Apr. 2004.

[58] SensAble, *PHANTOM Haptic Devices*, 2011. [Online]. Available: http://www.sensable.com/products-haptic-devices.htm

[59] Force Dimension, *Omega.6 Haptic Device*, 2011. [Online]. Available: http://www.forcedimension.com/omega6-overview

[60] S. Grange, F. Conti, P. Rouiller, P. Helmer, and C. Baur, "Overview of the Delta haptic device," in *Proc. Eurohaptics*, Birmingham, U.K., Jul. 2001.

[61] R. Q. Van Der Linde, P. Lammertse, E. Frederiksen, and B. Ruiter, "The HapticMaster, a new high-performance haptic interface," in *Proc. Eurohaptics*, Edinburgh, U.K., Jul. 2002, pp. 1–5.

[62] M. Ueberle, N. Mock, and M. Buss, "ViSHaRD10, a novel hyperredundant haptic interface," in *Proc. Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst./IEEE Virtual Reality Conf.*, Chicago, IL, Mar. 2004, pp. 58–65.

[63] N. Nitzsche, U. Hanebeck, and G. Schmidt, "Design issues of mobile haptic interfaces," *J. Robot. Syst.*, vol. 20, pp. 549–556, Aug. 2003.

[64] M. Buss, K. Lee, N. Nitzsche, A. Peer, B. Stanczyk, M. Ueberle, and U. Unterhinninghofen, "Advanced telerobotics: Dual-handed and mobile remote manipulation," in *Advances in Telerobotics*. Berlin, Germany: Springer-Verlag, 2007, pp. 471–498.

[65] CyberGlove Systems, *CyberGlove II Wireless Data Glove*, 2011. [Online]. Available: http://www.cyberglovesystems.com/products/cyberglove-ii/overview

[66] U. Hagn, M. Nickl, S. Jörg, G. Passig, T. Bahls, A. Nothhelfer, F. Hacker, L. Le-Tien, A. Albu-Schäffer, R. Konietschke, M. Grebenstein, R. Warpup, R. Haslinger, M. Frommberger, and G. Hirzinger, "The DLR MIRO: A versatile lightweight robot for surgical applications," *Ind. Robot, Int. J.*, vol. 35, no. 4, pp. 324–336, Oct. 2008.

[67] U. Hagn, T. Ortmaier, R. Konietschke, B. Kübler, U. Seibold, A. Tobergte, M. Nickl, S. Jörg, and G. Hirzinger, "Telemanipulator for remote minimally invasive surgery," *IEEE Robot. Autom. Mag.*, vol. 15, no. 4, pp. 28–38, Dec. 2008.

[68] F. Carpi, D. De Rossi, R. Kornbluh, R. Pelrine, and P. Sommer-Larsen, *Dielectric Elastomers as Electromechanical Transducers*. Amsterdam, The Netherlands: Elsevier, 2008.

[69] M. Matysek, P. Lotz, T. Winterstein, and H. Schlaak, "Dielectric elastomer actuators for tactile displays," in *Proc. 3rd Joint Eurohaptics Conf./Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Mar. 2009, pp. 290–295.

[70] V. Hayward, O. R. Astley, M. Cruz-Hernandez, D. Grant, and G. Robles-De-La-Torre, "Haptic interfaces and devices," *Sensor Rev.*, vol. 24, no. 1, pp. 16–29, 2004.

[71] M. W. Ueberle, "Design, control, and evaluation of a family of kinesthetic haptic interfaces," Dept. Electr. Eng. Inf. Technol., Ph.D. dissertation, Technische Universität München, Munich, Germany, 2006.

[72] K. Drewing, M. Fritschi, R. Zopf, M. Ernst, and M. Buss, "First evaluation of a novel tactile display exerting shear force via lateral displacement," *ACM Trans. Appl. Percept.*, vol. 2, no. 2, pp. 118–131, Apr. 2000.

[73] D. A. Lawrence, "Stability and transparency in bilateral teleoperation," *IEEE Trans. Robot. Autom.*, vol. 9, no. 5, pp. 624–637, Oct. 1993.

[74] S. Hirche, M. Ferre, J. Barrio, C. Melchiorri, and M. Buss, "Bilateral control architectures for telerobotics," in *Advances in Telerobotics*, Berlin, Germany: Springer-Verlag, 2007, pp. 163–176.

[75] N. Hogan, "Controlling impedance at the man/machine interface," in *Proc. IEEE Int. Conf. Robot. Autom.*, Scottsdale, AZ, May 1989, vol. 3, pp. 1626–1631.

[76] R. J. Anderson and M. W. Spong, "Bilateral control of teleoperators with time delay," *IEEE Trans. Autom. Control*, vol. 34, no. 5, pp. 494–501, May 1989.

[77] N. Chopra, M. Spong, S. Hirche, and M. Buss, "Bilateral teleoperation over the internet: The time varying delay problem," in *Proc. Amer. Control Conf.*, Denver, CO, Jun. 2003, vol. 1, pp. 155–160.

[78] Y. Yokokohji, T. Tsujioka, and T. Yoshikawa, "Bilateral control with time-varying delay including communication blackout," in *Proc. Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Los Alamitos, CA, Apr. 2002, pp. 285–292.

[79] P. Berestesky, N. Chopra, and M. W. Spong, "Discrete time passivity in bilateral teleoperation over the internet," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2004, vol. 5, pp. 4557–4564.

[80] S. Hirche and M. Buss, "Packet loss effects in passive telepresence systems," in *Proc. IEEE Conf. Decision Control*, Atlantis, Paradise Island, Bahamas, Dec. 2004, vol. 4, pp. 4010–4015.

[81] M. Minsky, "Telepresence," *Omni Mag.*, Jun. 1980.

[82] T. B. Sheridan, "Musings on telepresence and virtual presence," *Presence, Teleoper. Virtual Environ.*, vol. 1, no. 1, pp. 120–126, 1992.

[83] Y. Yokokohji and T. Yoshikawa, "Bilateral control of master-slave manipulators for ideal kinesthetic coupling-formulation and experiment," *IEEE Trans. Robot. Autom.*, vol. 10, no. 5, pp. 605–620, Oct. 1994.

[84] B. Hannaford, "Stability and performance tradeoffs in bi-lateral telemanipulation," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 1989, vol. 3, pp. 1764–1767.

[85] K. Hashtrudi-Zaad and S. E. Salcudean, "Analysis of control architectures for teleoperation systems with impedance/admittance master and slave manipulators," *Int. J. Robot. Res.*, vol. 20, no. 6, pp. 419–445, 2001.

[86] S. Hirche, "Haptic telepresence in packet switched communication networks," Dept. Electr. Eng. Inf. Technol., Technische Universität Munchen, Ph.D. dissertation, VDI-Verlag, Düsseldorf, Germany, 2005.

[87] S. Hirche and M. Buss, "Transparenz haptischer Telepräsenzsysteme mit konstanter Zeitverzögerung (Transparency of haptic telepresence systems with constant time delay)," *Automatisierungstechnik*, vol. 54, no. 2, pp. 51–59, Feb. 2006.

[88] B. Hannaford, L. Wood, D. McAffee, and H. Zak, "Performance evaluation of a six-axis generalized force-reflecting teleoperator," *IEEE Trans. Syst. Man Cybern.*, vol. 21, no. 3, pp. 620–633, May–Jun. 1991.

[89] W. A. Ijsselsteijn, H. de Ridder, J. Freeman, and S. E. Avons, "Presence: Concept, determinants and measurement," *Proc.*

SPIE—Int. Soc. Opt. Eng., vol. 3959, pp. 520–529, 2000.

[90] N. Durlach and M. Slater, "Presence in shared virtual environments and virtual togetherness," *Presence, Teleoper. Virtual Environ.*, vol. 9, no. 2, pp. 214–217, 2000.

[91] P. Hinterseer, S. Hirche, S. Chaudhuri, E. Steinbach, and M. Buss, "Perception-based data reduction and transmission of haptic data in telepresence and teleaction systems," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 588–597, Feb. 2008.

[92] P. G. Otanez, J. R. Moyne, and D. M. Tilbury, "Using deadbands to reduce communication in networked control systems," in *Proc. Amer. Control Conf.*, Anchorage, AK, Nov. 2002, vol. 4, pp. 3015–3020.

[93] P. Hinterseer and E. Steinbach, "Model-based data compression for 3D virtual haptic teleinteraction," in *Proc. IEEE Int. Conf. Consumer Electron.*, Las Vegas, NV, Jan. 2006, pp. 23–24.

[94] N. Sakr, J. Zhou, N. Georganas, J. Zhao, and X. Shen, "Prediction based haptic data reduction and compression in tele-mentoring systems," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 5, pp. 1727–1736, May 2008.

[95] S. Clarke, G. Schillhuber, M. Zaeh, and H. Ulbrich, "Prediction based methods for teleoperation across delayed networks," *Multimedia Syst.*, vol. 13, no. 4, pp. 253–261, May 2008.

[96] Y. Yonghee and M. Sung, "Haptic data transmission based on the prediction and compression," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, May 2008, pp. 1824–1828.

[97] K. Hikichi, H. Morino, I. Fukuda, S. Matsumoto, Y. Yasuda, I. Arimoto, M. Iijima, and K. Sezaki, "Architecture of haptics communication system for adaptation to network environments," in *Proc. IEEE Int. Conf. Multimedia Expo*, Los Alamitos, CA, Aug. 2001, pp. 563–566.

[98] C. Shahabi, M. Kolahdouzan, G. Barish, R. Zimmermann, D. Yao, K. Fu, and L. Zhang, "Alternative techniques for the efficient acquisition of haptic data," *ACM SIGMETRICS Performance Eval. Rev.*, vol. 29, no. 1, pp. 334–335, Jun. 2001.

[99] C. Shahabi, A. Ortega, and M. R. Kolahdouzan, "A comparison of different haptic compression techniques," in *Proc. IEEE Int. Conf. Multimedia Expo*, Lausanne, Switzerland, Aug. 2002, vol. 1, pp. 657–660.

[100] C. W. Borst, "Predictive coding for efficient host-device communication in a pneumatic force-feedback display," in *Proc. 1st Joint Eurohaptics Conf./Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Pisa, Italy, Mar. 2005, pp. 596–599.

[101] J. Kammerl and E. Steinbach, "Deadband-based offline-coding of haptic media," in *Proc. ACM Multimedia*, Vancouver, ON, Canada, Oct. 2008, pp. 549–558.

[102] F. Brandi and E. Steinbach, "Minimizing the number of support points in perceptual offline haptic compression," in *Proc. IEEE Int. Workshop Haptic Audio Visual Environ. Games*, Phoenix, AZ, Oct. 2010, pp. 7–12.

[103] P. Hinterseer, E. Steinbach, S. Hirche, and M. Buss, "A novel, psychophysically motivated transmission approach for haptic data streams in telepresence and teleaction systems," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Philadelphia, PA, Mar. 2005, vol. 2, pp. 1097–1100.

[104] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss, "Network traffic reduction in haptic telepresence systems by deadband control," in *Proc. IFAC World Congr.*, Jul. 2005, DOI: 10.1.1.61.7372.

[105] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss, "Transparent data reduction in networked telepresence and teleaction systems—Part I: Communication without time delay," *Presence, Teleoper. Virtual Environ.*, vol. 16, no. 5, pp. 523–531, Oct. 2007.

[106] S. Hirche and M. Buss, "Transparent data reduction in networked telepresence and teleaction systems—Part II: Time-delayed communication," *Presence, Teleoper. Virtual Environ.*, vol. 16, no. 5, pp. 532–542, Oct. 2007.

[107] J. Kammerl, I. Vittorias, V. Nitsch, B. Faerber, E. Steinbach, and S. Hirche, "Perception-based data reduction for haptic force-feedback signals using adaptive deadbands," *Presence, Teleoper. Virtual Environ.*, vol. 19, no. 5, pp. 450–462, Oct. 2010.

[108] X.-D. Yang, W. F. Bischof, and P. Boulanger, "Perception of haptic force magnitude during hand movements," in *Proc. IEEE Int. Conf. Robot. Autom.*, Pasadena, CA, May 2008, pp. 2061–2066.

[109] M. H. Zadeh, D. Wang, and E. Kubica, "Perception-based lossy haptic compression considerations for velocity-based interactions," *Multimedia Syst.*, vol. 13, no. 4, pp. 275–282, 2008.

[110] P. Hinterseer and E. Steinbach, "A psychophysically motivated compression approach for 3D haptic data," in *Proc. Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Arlington, VA, Mar. 2006, pp. 35–41.

[111] J. Kammerl, R. Chaudhari, and E. Steinbach, "Exploting directional dependencies of force perception for lossy haptic data reduction," in *Proc. IEEE Int. Symp. Haptic Audio-Visual Environ. Games*, Phoenix, AZ, Oct. 2010, DOI: 10.1109/HAVE.2010.5623975.

[112] J. Drössler, "An n-dimensional Weber law and the corresponding Fechner law," *J. Math. Psychol.*, vol. 44, pp. 330–335, 2000.

[113] I. Vittorias, H. Ben Rached, and S. Hirche, "Haptic data reduction in multi-DoF teleoperation systems," in *Proc. IEEE Int. Symp. Haptic Audio-Visual Environ. Games*, Phoenix, AZ, Oct. 2010, DOI: 10.1109/HAVE.2010.5623991.

[114] K. Boff and J. Lincoln, *Engineering Data Compendium: Human Perception and Performance*. Wright-Patterson AFB, OH: Armstrong Aerospace Research Laboratory, 1988.

[115] J. D. Hwang, M. D. Williams, and G. Niemeyer, "Toward eventbased haptics: Rendering contact using open-loop force pulses," in *Proc. Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Los Alamitos, CA, Apr. 2004, pp. 24–31.

[116] J. Kammerl, R. Chaudhari, and E. Steinbach, "Combining contact models with perceptual data reduction for efficient haptic data communication in networked VEs," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 1, pp. 57–68, Jan. 2011.

[117] J. Arata, H. Takahashi, P. Pitakwatchara, S. Warisawa, K. Tanoue, K. Konishi, S. Ieiri, S. Shimizu, N. Nakashima, K. Okamura, Y. Fujino, Y. Ueda, P. Chotiwan, M. Mitsuishi, and M. Hashizume, "A remote surgery experiment between Japan and Thailand over Internet using a low latency CODEC system," in *Proc. IEEE Int. Conf. Robot. Autom.*, Rome, Italy, Apr. 2007, pp. 953–959.

[118] A. Peer, S. Hirche, C. Weber, I. Krause, M. Buss, S. Miossec, P. Evrard, O. Stasse, E. S. Neo, A. Kheddar, and K. Yokoi, "Intercontinental cooperative telemanipulation between Germany and Japan," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Nice, France, Sep. 2008, pp. 2715–2716.

[119] S. Tachi, "Real-time remote robotics-toward networked telexistence," *IEEE Comput. Graph. Appl.*, vol. 18, no. 6, pp. 6–9, Nov./Dec. 1998.

[120] J. Cha, Y. S. Ho, Y. Kim, J. Ryu, and I. Oakley, "A framework for haptic broadcasting," *IEEE Multimedia*, vol. 16, no. 3, pp. 16–27, Jul.–Sep. 2009.

[121] H. King, J. Kammerl, B. Hannaford, and E. Steinbach, "Establishing multimodal telepresence sessions using the session initiation protocol (SIP) and advanced haptic codecs," in *Proc. IEEE Haptics Symp.*, Waltham, MA, Mar. 2010, pp. 321–325.

[122] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, *SIP: Session Initiation Protocol, RFC 3261 (Proposed Standard)*, Internet Engineering Task Force, Jun. 2002, updated by RFCs 3265, 3853, 4320, 4916, 5393, 5621. [Online]. Available: http://www.ietf.org/rfc/rfc3261.txt

[123] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, "Error resilient video coding techniques," *IEEE Signal Process. Mag.*, vol. 17, no. 4, pp. 61–82, Jul. 2000.

[124] F. Brandi, J. Kammerl, and E. Steinbach, "Error-resilient perceptual coding for networked haptic interaction," in *Proc. ACM Multimedia*, Firenze, Italy, Oct. 2010, pp. 351–360.

[125] P. A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 390–404, Apr. 2006.

[126] F. Brandi and E. Steinbach, "Low-complexity error-resilient data reduction approach for networked haptic sessions," in *Proc. IEEE Int. Workshop Haptic Audio Visual Environ. Games*, NanchangJiangxiChina, Oct. 2011, pp. 135–140.

[127] A. Hamam, M. Eid, A. Saddik, and N. D. Georganas, "A fuzzy logic system for evaluating quality of experience of haptic-based applications," in *Proc. Int. Conf. Haptics, Perception Devices Scenarios*, Madrid, Spain, Jun. 2008, pp. 129–138.

[128] E. Ruffaldi, D. Morris, T. Edmunds, F. Barbagli, and D. K. Kai, "Standarized evaluation of haptic rendering systems," in *Proc. Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Alexandria, VA, Mar. 2006, pp. 33–40.

[129] M. Sargardia, T. Hulin, C. Preusche, and G. Hirzinger, "A benchmark of force quality in haptic rendering," in *Proc. Int. Conf. Human-Computer Interaction*, San Diego, CA, Jul. 2009, pp. 19–24.

[130] R. Chaudhari, E. Steinbach, and S. Hirche, "Towards an objective quality evaluation framework for haptic data reduction," in *Proc. IEEE World Haptics Conf.*, Istanbul, Turkey, Jun. 2011, pp. 539–544.

[131] R. A. Hess, "Structural model of the adaptive human pilot," *J. Guid. Control Dyn.*, vol. 3, no. 5, pp. 416–423, 1980.

[132] L. F. Penin, A. Caballero, R. Aracil, and A. Barrientos, "Human behavior modeling in master-slave teleoperation with kinesthetic feedback," in *Proc. IEEE Int. Conf. Robot. Autom.*, Leuven, Belgium, May 1998, pp. 2244–2249.

[133] H. Yilmaz, "On the laws of psychophysics," *Bull. Math. Biol.*, vol. 26, pp. 235–237, Sep. 1964.

[134] P. Mitra and G. Niemeyer, "Model-mediated telemanipulation," *Int. J. Robot. Res.*, vol. 27, pp. 253–262, Feb. 2008.

[135] Z. Shi, H. Zou, M. Rank, L. Chen, S. Hirche, and H. J. Mueller, "Effects of packet loss and latency on temporal discrimination of visual-haptic events," *IEEE Trans. Haptics*, vol. 3, no. 1, pp. 28–36, 2010.

[136] M. Eid, J. Cha, and A. El Saddik, "An adaptive multiplexer for multimodal data communication," in *Proc. Int. Workshop Haptic Audio-Visual Environ. Games*, Lecco, Italy, Nov. 2009, pp. 111–116.

[137] J. Cha, Y. Seo, Y. Kim, and J. Ryu, "An authoring/editing framework for haptic broadcasting: Passive haptic interactions using MPEG-4 BIFS," in *Proc. 2nd Joint Eurohaptics Conf./Int. Symp. Haptic Interfaces Virtual Environ. Teleoperator Syst.*, Washington, DC, 2007, pp. 274–279.

ABOUT THE AUTHORS

Eckehard Steinbach (Senior Member, IEEE) studied electrical engineering at the University of Karlsruhe, Karlsruhe, Germany, the University of Essex, Essex, U.K., and ESIEE, Paris, France. He received the Engineering Doctorate from the University of Erlangen-Nuremberg, Germany, in 1999.

From 1994 to 2000, he was a member of the research staff of the Image Communication Group, University of Erlangen-Nuremberg. From February 2000 to December 2001, he was a Postdoctoral Fellow with the Information Systems Laboratory, Stanford University, Stanford, CA. In February 2002, he joined the Department of Electrical Engineering and Information Technology, Technische Universität München, Munich, Germany, where he is currently a Full Professor for Media Technology. His current research interests are in the area of audiovisual-haptic information processing and communication as well as networked and interactive multimedia systems.

Dr. Steinbach has been serving as an Associate Editor for the IEEE Transactions on Circuits and Systems for Video Technology since 2006 and for the IEEE Transactions on Multimedia since 2011.

Sandra Hirche (Senior Member, IEEE) received the Diploma Engineer degree in mechanical engineering and transport systems from the Technical University Berlin, Berlin, Germany, in 2002 and the Doctor of Engineering degree in electrical engineering and computer science from the Technische Universität München, Munich, Germany, in 2005.

From 2005 to 2007, she has been a Japanese Society for the Promotion of Science (JSPS) Postdoctoral Researcher at the Tokyo Institute of Technology, Tokyo, Japan. Since 2008, she has been an Associate Professor heading the Associate Institute for Information-oriented Control in the Department of Electrical Engineering and Information Technology, Technische Universität München. Her research interests include control over communication networks, networked control systems, control of large-scale systems, cooperative control, human-in-the-loop control, human–machine interaction, robotics, and haptics.

Dr. Hirche has been serving as Chair for Student Activities in the IEEE Control System Society (CSS) since 2009, as Chair of the CSS Awards Subcommittee on "CDC Best Student-Paper Award" since 2010, and became an elected member of the Board of Governors of IEEE CSS in 2010.

Marc Ernst studied physics at the Ruprecht-Karls-Universität in Heidelberg and the Johann Wolfgang Goethe-Universität Frankfurt am Main. He received the Ph.D. degree for investigations on human visuomotor behavior from the Max Planck Institute for Biological Cybernetics, Tübingen, Germany, in 2000.

Then, he spent almost two years as a Research Associate at the University of California at Berkeley, Berkeley, working on psychophysical experiments and computational models investigating the integration of visual-haptic information. In 2001, he returned to the Max Planck Institute and became Principal Investigator of the Sensorimotor Lab in the Department of Human Perception, Cognition, and Action. In 2007, he became leader of the Max Planck Research Group on Human Multisenory Perception and Action at the Max Planck Institute for Biological Cybernetics. Since 2011, he has been Chair of the Department of Cognitive Neuroscience, University of Bielefeld, Bielefeld, Germany. He is interested in human multisensory perception, sensorimotor integration, and man–machine interaction. He has published well over 50 papers and conference proceedings in high-profile journals including *Nature*, *Science*, and *Nature Neuroscience*. He was involved in several international collaborative grants, including different European Union Projects focussing on the human–machine interaction.

Dr. Ernst is President and founding member of the EuroHaptics Society and Vice Chair of the IEEE Technical Committee on Haptics. He is an Editor for the *Springer Series on Touch and Haptic Systems* and an Associate Editor for the *ACM Transactions on Applied Perception*, *PLoS ONE*, and the journal *Seeing & Perceiving*.

Fernanda Brandi (Student Member, IEEE) was born in Brazil in 1981. She received the B.S. degree in electrical engineering and the M.S. degree from the Universidade de Brasilia, Brasilia, Brazil, in 2006 and 2009, respectively. She is currently working towards the Ph.D. degree at the Institute of Media Technology, Technische Universität München, Munich, Germany.

Her research interests include video and haptic communication with focus on optimizing the signal compression in the presence of communication unreliabilities.

Rahul Chaudhari (Member, IEEE) received the B.Eng. degree (graduating as the top engineering student in class) in electronics and telecommunications from the University of Pune, Pune, India, June 2006 and the M.S. degree in communication systems, with a focus on signal processing, from the Technische Universität München, Munich, Germany, in October 2009. During his M.S. studies, he was supported by a scholarship from Siemens AG, Germany, under the "Siemens Master's Program."

Since January 2010, he has been a member of the research and teaching staff at the Institute for Media Technology, Technische Universität München. His current research interests include perceptual coding of haptic data and objective evaluation of quality-of-experience for haptic communication.

Julius Kammerl (Student Member, IEEE) studied computer science at the Technische Universität München, Munich, Germany, where he received the degree "Dipl.-Inf. (Univ)" in January 2005.

After working at the Audio and Multimedia Group at Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany, he joined the Institute for Media Technology, Technische Universiät Müchen in May 2006, where he is currently working as a member of the research and teaching staff. His current research interests are in the field of haptic communication with focus on perceptual compression of haptic data streams. He is a member of the interdisciplinary research cluster on "high-fidelity telepresence and teleaction" which is funded by the German Research Foundation DFG.

Iason Vittorias (Student Member, IEEE) was born in Rhodes, Greece, in 1985. He received the Diploma Engineer degree in electrical and computers engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2007. He is currently working towards the Ph.D. degree at the Technische Universität München, Munich, Germany.

Since 2008, he has been a Research Assistant at the Institute of Automatic Control Engineering, Technische Universität München. His research interests include haptic teleoperation systems over networks, passivity-based control, and haptic data reduction.