[Eckehard Steinbach, Sandra Hirche, Julius Kammerl, Iason Vittorias, and Rahul Chaudhari]

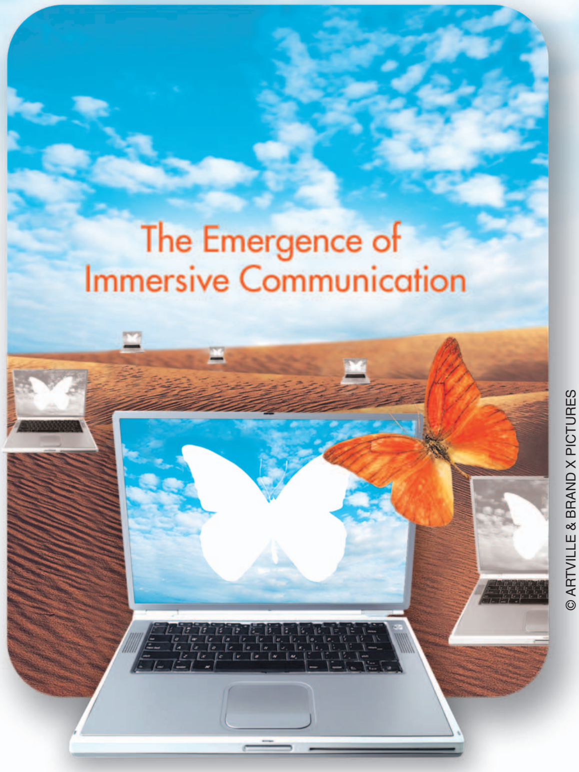
Haptic Data Compression and Communication

[A review of the challenges for telepresence and teleaction]

The past decade has witnessed how audio-visual communication has shaped the way humans interact with or through technical systems. In contemporary times, the potential of haptic communication has been recognized as being compelling to further augment human-to-human and human-to-machine interaction. In the context of immersive communication, video and audio compression are considered key enabling technologies for high-quality interaction. In contrast, the compression of haptic data is a field of research that is still relatively young and not fully explored. This disregards the fact that we as humans rely heavily on the haptic modality to interact with our environment. True immersion into a distant environment and efficient collaboration between multiple participants both require the ability to physically interact with objects in the remote environment. With recent advances in virtual reality, man-machine interaction, telerobotics, telepresence, and teleaction, haptic communication is proving instrumental in enabling many novel applications. The goal of this overview article is to summarize the state of the art and the challenges of haptic data compression and communication for telepresence and teleaction.

INTRODUCTION

In today's multimedia systems, visual and auditory information are predominant. Efficient data compression methods and standards for digital audio and video have paved the way for widespread adoption of modern multimedia technology. Especially by investigating the characteristics and limits of human perception, important steps in the development of efficient data reduction schemes have been taken.



© ARTVILLE & BRAND X PICTURES

On the other hand, the compression and communication of haptic data has not been a field of intense research so far. This is in contrast to the fact that we as humans rely heavily on the haptic modality in our daily life. Haptic perception, handled by the human somatosensory system, consists of the kinaesthetic and the tactile sense. Kinaesthetic information refers to the sensation of muscle movements and helps to determine joint positions. Tactile information refers to touch, pressure, temperature, and pain and is a prerequisite for nearly all neuromuscular activities like perception of objects and their positions in case of limited visibility, identification of materials, and other surface properties. In particular, in so-called telepresence and teleaction (TPTA) systems, the haptic modality plays a central role. Telepresence systems allow us to immerse into environments that are distant, inaccessible (including virtual environments), scaled to macro- or nano-dimensions, or hazardous for a human being. In case of an additional manipulative ability in the remote environment, we refer to teleaction or

---

Digital Object Identifier 10.1109/MSP.2010.938753
Date of publication: 17 December 2010

1053-5888/11/$26.00©2011IEEE

telemanipulation systems. When multiple humans are immersed into the same real or virtual remote environment, collaborative task execution becomes feasible.

THE COMPRESSION OF HAPTIC DATA IS A FIELD OF RESEARCH THAT IS STILL RELATIVELY YOUNG AND NOT FULLY EXPLORED.

Many exciting areas of application for TPTA technologies exist. Examples can be found in tele-education, teleconferencing, telemanipulation in dangerous environments, telerobotics, tele- and minimally invasive surgery, microassembly, and on-orbit teleservicing. In all scenarios, the realistic presentation of remote environment properties determines the performance and quality of experience for the user.

TELEPRESENCE AND TELEACTION SYSTEMS

A TPTA system, as visualized in Figure 1, consists of three main components: the human system interface (HSI), the teleoperator (TOP), and a communication link connecting them. The HSI represents both the input device, typically consisting of haptic devices for position and orientation input, and output devices for displaying multiple modalities, e.g., a head-mounted-display (HMD) for stereo-video, headphones for audio and haptic display devices for force and torque feedback. By means of the HSI, the human operator (OP) commands the motion of the TOP during the observation of and the interaction with the remote environment. The TOP itself is a robot equipped with multiple sensors like video cameras, microphones, as well as force and torque sensors. To be able to physically interact with the remote environment, it may be equipped with haptic actuators like grippers or more anthropomorphic limbs. In case the remote environment is virtual, physical interaction is simulated to generate multimodal feedback. Ideally, the TPTA system is transparent to the operator, i.e., the operator feels as if he/she were in place of the robot interacting with the environment. In other words, he/she feels completely immersed.

The architecture of a TPTA system can be decomposed into two separate subsystems, i.e., a HSI/master and a TOP/slave system, bidirectionally exchanging haptic signals over the network. Which haptic signals are transmitted depends on the deployed control architecture. The choice of the control architecture determines stability, robustness, and transparency, which are influenced by the master and slave device dynamics, the communication characteristics and the available sensors. A detailed discussion of these factors exceeds the scope of this article and the reader is referred to the survey articles [2], [3]. The velocity-force architecture, as shown in Figure 2, is one of the most commonly applied architectures for haptic telepresence systems. Here, the human controls the HSI that transmits its velocity $\dot{x}_m$. Naturally, the human end-effector velocity $\dot{x}_h$ is equal to the HSI velocity assuming contact between the two. A major advantage of exchanging velocity signals rather than absolute position coordinates is that the local coordinate systems of the OP and the TOP do not have to be identical. The TOP receives the desired velocity signal $\dot{x}_s^d$ with a communication delay of T_1. It is fed to a local control loop for controlling the velocity $\dot{x}_s$ of the slave robot. As soon as the TOP encounters contact with the environment, an environment force F_e occurs. This environment force is measured by the TOP using force sensors and the resulting signal F_s is transmitted back from the TOP to the OP. The HSI receives the force-feedback information F_m^d (delayed by T_2) that becomes the reference for the local force control loop at the HSI. This results in a force F_h displayed to the human. The impedance of the master/slave subsystems can be approximated around their operating point by local linear time-invariant models

$$\begin{aligned} M_m\ddot{x}_m(t) + B_m\dot{x}_m(t) &= \tau_m(t) + F_h(t) \\ M_s\ddot{x}_s(t) + B_s\dot{x}_s(t) &= \tau_s(t) - F_e(t), \end{aligned} \tag{1}$$

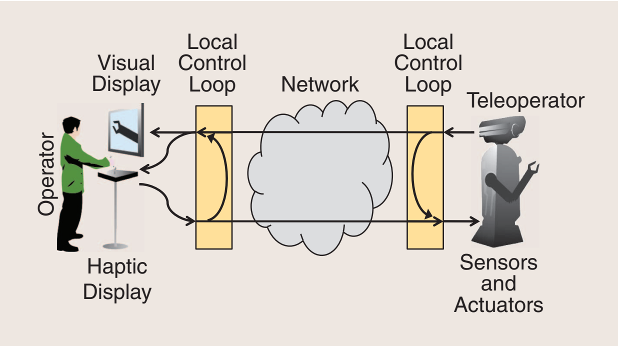
where M_* defines a positive inertia matrix and B_* is the corresponding damping matrix, and τ_m, τ_s are the forces resulting from the corresponding motor torques that are a function of the local control operations at the HSI and TOP. Ignoring the data reduction and data reconstruction blocks in Figure 2, the communication latency is accounted for by

$$\dot{x}_s^d(t) = \dot{x}_m(t - T_1) \quad \text{and} \quad F_m^d(t) = F_s(t - T_2). \tag{2}$$

The dynamics of the human H and the environment E are unknown functions, which close the global control loop:

$$\dot{x}_h(t) = f_H(H, F_h, t), \quad F_e(t) = f_E(E, \dot{x}_s, t). \tag{3}$$

For real-time processing and communication of haptic data in TPTA systems, very strict delay constraints are imposed by the involved control loops. Already milliseconds of time delay may destabilize the overall system. This prevents us from using signal processing and compression approaches that introduce algorithmic delay. Hence, when a TPTA session is run across a packet-switched network (e.g. the Internet), new haptic samples are immediately packetized and injected into the network to minimize end-to-end delay. This results in packet rates of up to the applied sampling rates (typically 1 kHz). The high packet rates as well as the overhead due to the transmission of packet header information become critical factors [4]. Furthermore,



[FIG1] Schematic overview of a TPTA system (adapted from [1]).

IN SO-CALLED TELEPRESENCE AND TELEACTION SYSTEMS, THE HAPTIC MODALITY PLAYS A CENTRAL ROLE.

recent advances in haptic multimedia systems clearly show the trend of integrating more and more degrees of freedom (DoF) to improve haptic immersion and manipulative flexibility. To enable extensive haptic interaction, position tracking and force feedback display are required for many joints. For instance, Immersion's CyberGrasp/CyberGlove HSI device integrates 22 DoF to enable the haptic modality for a human hand. A comprehensive introduction to haptic input and output devices can be found in [5]. Future HSIs will be richly equipped with sensors and actuators to provide a skin-type sensing with a large number of DoF. As each DoF needs to be continuously sampled and controlled, the amount of sensor/actuator data quickly increases and novel data reduction methods are required to ensure an efficient usage of the available communication resources.

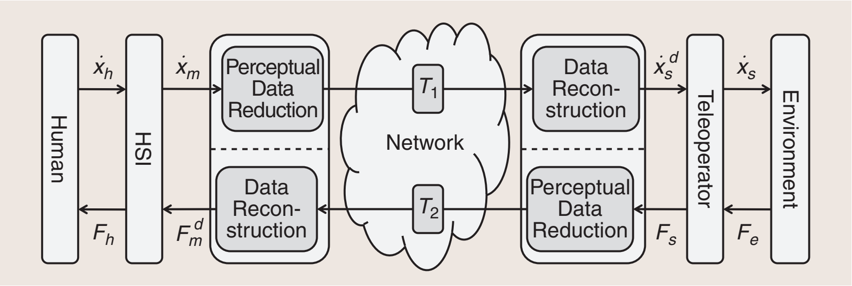
Besides real-time communication of haptic information, several application scenarios exist where efficient off-line coding and playback of multimodal TPTA sessions are of great relevance. For instance, this can be used to give the operator the possibility to record a haptic TPTA-session. From the recorded interaction session, the system and operator performance can be determined, which is especially interesting in applications like on-orbit servicing and minimally invasive surgery. Additionally, the replay of recorded haptic manipulation traces is of great interest in teaching scenarios. While a teacher demonstrates important manipulation steps on a TPTA system, his/her activity is recorded. Later, the students can be guided along the teacher's movements while being simultaneously served with recorded media streams of all modalities.

ORGANIZATION

In this article, we discuss haptic data communication with a special focus on perceptual data reduction in TPTA systems. As we go along, we highlight the challenges that should be addressed by the signal processing community to advance the state of the art in this emerging interdisciplinary research field.

A SHORT HISTORY OF HAPTIC DATA REDUCTION

Data compression for efficient storage and/or transmission has been investigated extensively by the information theory and source coding communities. For the compression of multimedia signals, lossy compression algorithms are most suitable, as they explicitly detect and remove irrelevant information, which either is not perceivable by the human or cannot be displayed due to hardware limitations. If compression is done without affecting the perceptual quality and without disturbing system performance, the applied signal processing methods are said to be transparent. Lossy compression achieves high compression rates and is therefore also of great interest for the compression of haptic signals. Data reduction for haptic signals is, however, fundamentally different from the compression of audio and video. As a result of the strict delay and stability requirements, approaches that compress haptic signals without algorithmic delay are required.



[FIG2] Schematic overview of haptic signal communication with perceptual data reduction in TPTA systems with velocity/force control.

Early approaches for haptic data compression can be found in [6]–[8], where different sampling and quantization techniques for haptic data are introduced. Applying DPCM and ADPCM with Huffman coding on haptic signals has been treated in [9] and [10]. Here, the concept of perceptual lossy compression for haptic data has been introduced and the authors propose to adjust the quantization coarseness such that the introduced quantization noise stays below absolute human haptic perception thresholds.

None of the aforementioned approaches addresses the reduction of the high packet rates, which is the main challenge for real-time haptic interaction across packet-switched networks. The first proposal that targets packet rate reduction for networked control systems can be found in [11]. If the difference between the most recently sent update and the current input value exceeds a fixed threshold, signal updates are triggered. The receiver reacts to a missing sample by holding the value of the most recently received sample.

First perceptual data reduction approaches for real-time haptic communication across the Internet have been presented in [12]–[18] and investigated in terms of stability criteria in [15]–[18]. The perceptual deadband (PD) data reduction approach in [12]–[18] successfully addresses the problem of high packet rates in networked TPTA systems and satisfies the strict delay constraints. A first theoretical analysis of the performance of the PD approach appeared in [19]. In [20]–[23], the PD approach is combined with predictive coding. In [24], the PD approach is extended by incorporating a model of human force-feedback discrimination during relative hand movements.

Haptic data reduction makes the TPTA system more vulnerable to packet losses. Reference [25] proposes a TCP-compatible transport protocol with a congestion and rate control mechanism that is optimized for haptic Internet communication. Further work on transport schemes for haptic communication can be found in [26], where haptic samples are prioritized and grouped according to their predictability, network condition, and the detected haptic system activity. An adaptive packetization scheme for haptic data transmission is proposed in [27].

BESIDES REAL-TIME COMMUNICATION OF HAPTIC INFORMATION, SEVERAL APPLICATION SCENARIOS EXIST WHERE EFFICIENT OFF-LINE CODING AND PLAYBACK OF MULTIMODAL TPTA SESSIONS ARE OF GREAT RELEVANCE.

Off-line compression of haptic data for training, documentation and performance analysis is fundamentally different from real-time haptic communication in TPTA systems. Here, at a given time, haptic data is either sent from the haptic sensors to the recording device or played back. In both cases, a unidirectional data transfer of haptic information is established without strict delay requirements and the use of specific stability measures. This allows for pre- and/or post-processing, including block processing of the corresponding haptic data streams, which would otherwise be unfeasible during delay critical real-time processing of haptic data in an online telemanipulation scenario. Therefore common coding approaches, such as DCT, DWT, and Vector Quantization, which are widely used in the field of audio and image/video coding are suitable for the off-line compression of haptic signals. Nevertheless, the PD data reduction approach is also applicable to off-line compression and has been shown to achieve excellent compression performance for recorded TPTA sessions [28].

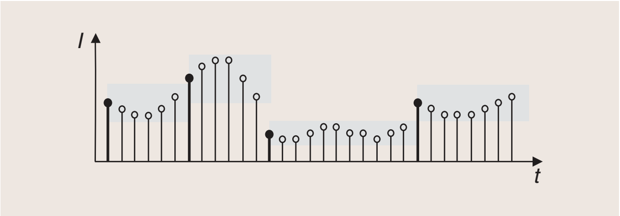
PSYCHOPHYSICS

To investigate whether a human can detect a stimulus, differentiate between two stimuli, and quantify the magnitude or nature of this difference, psychophysical experiments are performed. In this context, typically two different types of thresholds are of interest. Absolute (sometimes called detection) thresholds refer to the smallest stimulus amplitude that can be preceived by a subject. Difference thresholds refer to the smallest difference in stimulus magnitude that can be perceived.

In 1834, the experimental physiologist Ernst Weber was among the first to propose a mathematical relationship between the physical intensity of a stimulus and its phenomenologically perceived intensity [29], [30]. Specifically, he proposed the size of the difference threshold [or just noticeable difference (JND)] to be a linear function of stimulus intensity. This became known as the Weber's law of JND. It can be represented by the following equation:

$$\frac{\Delta I}{I} = \kappa = \text{constant}, \qquad (4)$$

where I is the stimulus intensity, ΔI is the so-called difference threshold or the JND and κ is a constant called the Weber fraction. With some variation, Weber's law has been found to apply to almost every human sense, including haptic perception, and over a wide stimulus range [30]–[34].



[FIG3] **Principle of PD data reduction. The perception thresholds (boundaries of gray zones) are a function of the haptic stimulus intensity.**

HAPTIC DATA REDUCTION USING PERCEPTUAL DEADBANDS

Weber's law provides a mathematical model of human haptic perception that allows for the detection of imperceptible changes in haptic signals. The principle of PDs that dynamically adapt to human perception thresholds is shown in Figure 3. Samples with black solid circles represent the output of the PD data reduction scheme. They define perception thresholds represented by a deadband, illustrated as gray zones. Samples with empty circles fall within the currently defined deadband and can be dropped as the associated signal change is too small to be perceptible. Please note that the size of the applied deadband Δ at discrete time i is a function of a deadband parameter k, the type of the haptic signal (velocity/force information) and the signal amplitude of the recently transmitted haptic sample value, where n samples back in time the last perception threshold violation took place

$$\begin{aligned} \Delta^{\dot{x}_m}(i) &= \cdots = \Delta^{\dot{x}_m}(i-n) = k^{\dot{x}} \cdot |\dot{x}_m(i-n)| \\ \Delta^{F_s}(i) &= \cdots = \Delta^{F_s}(i-n) = k^{F} \cdot |F_s(i-n)|. \end{aligned} \qquad (5)$$

The deadband parameter k is closely related to the Weber threshold parameter κ in (4). Once the deadband is violated by a new input sample, it is considered to be relevant for encoding/transmission and redefines the applied perception thresholds. The blocks denoted with "Perceptual Data Reduction" in Figure 2 show where the data reduction is applied in a networked TPTA system. As only the samples that violate the deadband are considered to constitute perceptible signal changes, the PD coding scheme allows for signal-adaptive downsampling and therefore dramatically reduces the number of samples to be transmitted. In case of noisy sensors, a minimum absolute deadband size exceeding the current noise level should be defined, to avoid unnecessary packet triggers at small signal amplitudes. Alternatively, low-delay low-pass filtering can be applied [35].

At the receiver side, the signal needs to be upsampled to a constant rate before it can be applied to the local haptic rendering and control loop. To this end, a simple hold-last-sample approach has been proposed for data reconstruction in [11] and [12].

In real-world TPTA experiments with one DoF, packet rate reductions of more than 85% have been reported for this approach [13]. In [13], the PD data reduction scheme operates such that most users cannot feel the difference between a TPTA session with and without data reduction being applied. A further increase of the deadband parameter k in (5) leads to noticeable distortion but does not necessarily affect the task performance, which refers to the speed and accuracy with which a user can complete a given task in a TPTA scenario.

WHETHER THE ISOTROPIC DEAD ZONE LEADS TO THE BEST POSSIBLE PERFORMANCE FROM A PERCEPTUAL OR TASK PERFORMANCE POINT OF VIEW HAS NOT YET BEEN CONCLUSIVELY ANSWERED.

To the best of our knowledge, as of today there is no comprehensive model available that directly maps the distortion introduced by data reduction to task performance. Such a model would be of great value to parameterize the data reduction scheme. First steps towards designing an objective metric for haptic signal distortion have been presented in [36]. Deriving such a mapping is one of the open challenges that should be addressed jointly by the control engineering and signal processing communities.

VELOCITY-ADAPTIVE PERCEPTUAL DEADBANDS

Very few of the empirical studies that investigate the limits of human perception consider the effects of dynamic movements that are important aspects of real-life task performance. Along these lines, the previously discussed data reduction principle has been extended for force-feedback signals by incorporating an additional perceptual dimension—specifically, the velocity of the operator's hand movement [24]. In comparison to (5), the deadband parameter for the force signal F_s is now modeled as a function of velocity

$$\phi = k^F + \alpha \cdot |\dot{x}_s|. \qquad (6)$$

Psychophysical experiments show that the velocity-adaptive deadband approach leads to better performance when compared to the PD approach in the section "Haptic Data Reduction Using Perceptual Deadbands." Reference [24] reports further packet rate reductions of up to 30% without perceptibly impairing the quality of the force-feedback signal or significantly affecting task performance accuracy.

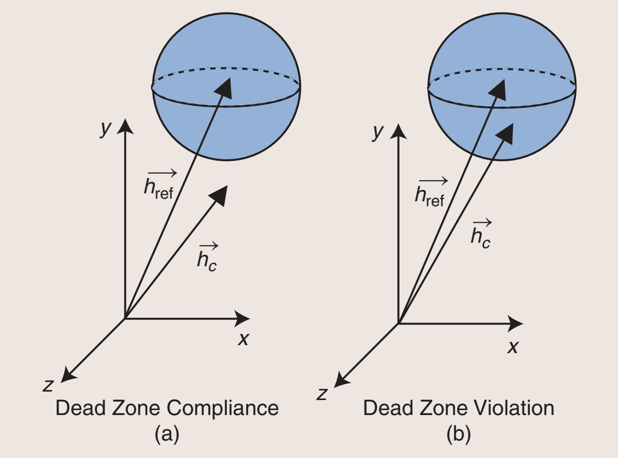
EXTENSION FOR SIGNALS WITH MULTIPLE DEGREES OF FREEDOM

In real-world TPTA scenarios, the haptic data streams typically consist of multiple DoF. The extension of the single-DoF data reduction approach to multi-DoF can be designed in many ways, one being to apply the 1-DoF PD data reduction scheme to each component of the multidimensional data individually. This approach is highly inefficient, as the resulting signal update rate would always be determined by the component with the lowest signal amplitude. An alternative approach, the construction of a so called multi-DoF dead zone is more promising in this context. To this end, [37] proposes an isotropic dead zone. In the 2-DoF case, this leads to a circular dead zone, centered at the tip of haptic sample vectors. Likewise, the 3-DoF case leads to a spherically shaped dead zone. In line with the 1-DoF PD data reduction approach, the volume of the dead zone for the multidimensional case is defined by the amplitude of the most recently transmitted haptic vector (see Figure 4). The experimental results in [13] show that with this extension to three DoF, a data reduction performance similar to that in the 1-DoF case can be achieved.

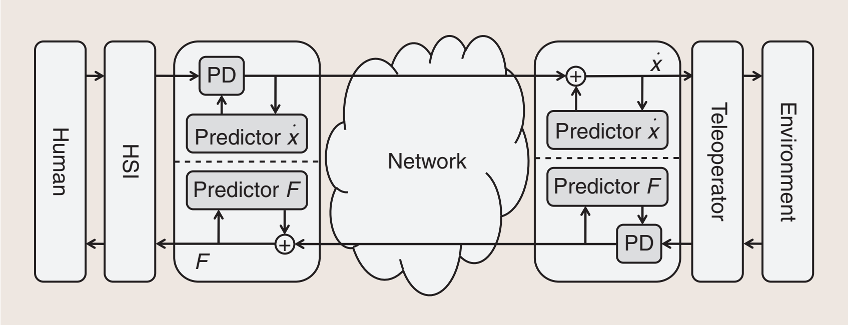
Whether the isotropic dead zone leads to the best possible performance from a perceptual or task performance point of view has not yet been conclusively answered. Preliminary investigations in [38] suggest that in fact the multi-DoF equivalent of the 1-DoF dead zone is not isotropic. Another unsolved issue is the joint compression of Cartesian and orientation variables such as position/angles and forces/torques. Further investigations, in particular, for dynamically changing haptic stimuli are necessary.

PREDICTIVE CODING

To further reduce the number of packets to be transmitted, a prediction model for haptic signals can be used to estimate future haptic sample values from previous data. The same predictor runs in parallel at both the OP and TOP sides; the OP side being fed with the incoming force feedback signals and the TOP side with position and/or velocity values. Both predictors being kept strictly coherent at any given instant of time, only a sample that differs from the current prediction value by more than the JND needs to be transmitted.



[FIG4] An isotropic dead zone for haptic signals with three DoF. (a) Dead zone compliance. (b) Dead zone violation. $\vec{h}_{\text{ref}}$ denotes the most recently transmitted haptic sample vector that defines the applied spherical dead zone. $\vec{h}_c$ is the current haptic sample vector.



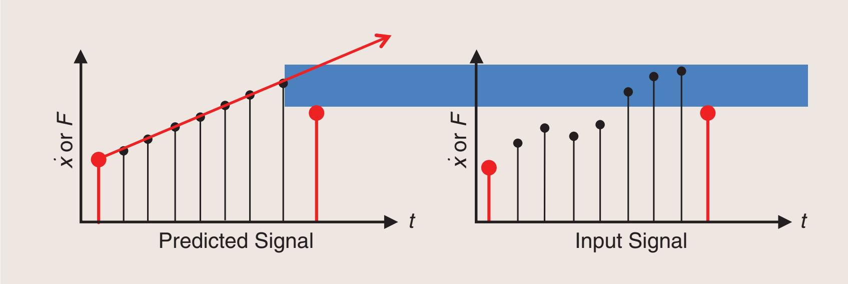
[FIG5] **System architecture for the combination of predictive coding with PD data reduction. Predictor $\dot{x}$ and Predictor F represent the prediction processes for the haptic velocity and force-feedback signals, respectively.**

Figure 5 shows the system architecture for such a combination of predictive coding and PD data reduction. In case no update is received at the TOP side, simply the locally predicted velocity samples are used. Similarly, if no update is received at the OP side, locally predicted force samples are used for display to the user.

SIGNAL-BASED PREDICTIVE CODING

Different kinds of predictors can be used to estimate the future haptic values. In [20], a linear predictor of the first order has been used in conjunction with PDs. Even this simple predictor leads to a significant decrease in packet rate without deteriorating the immersiveness of the system.

Figure 6 illustrates the combination of a linear predictor with the PD data reduction scheme in the section "Haptic Data Reduction Using Perceptual Deadbands." It can be seen that only in case the predicted value and the actual sample differ by more than the perception threshold (deadband violation) a signal update has to be sent to the receiver. In virtual environments, accumulated position errors can be simply corrected by regular absolute position updates. Even more efficient is to piggyback these position updates onto velocity packets triggered by the PD scheme. For real TOP systems, position updates are more critical as the TOP has to be moved to the new position. This quickly leads to noticeable artifacts and limits the usefulness of predictive coding on velocity signals for real world TPTA systems. For force signals this limitation does not apply.



[FIG6] **Illustration of combining PD data reduction with first-order predictive coding. Only samples illustrated in red are transmitted and used for updating the predictor.**

MODEL-BASED PREDICTIVE CODING

In a TPTA system, kinaesthetic motion commands are sent to and executed by the TOP. As soon as contact with the remote environment is encountered, corresponding force-feedback is generated to be reflected to the human operator. Haptic feedback is thus associated with certain end-effector positions within the remote environment. By combining absolute position and force-feedback information received from the remote side, the construction of a local model of the geometric structure and physical properties of the currently touched remote object surface becomes feasible [39]. As soon as enough position-force feedback pairs are collected to compute the model parameter values, a haptic rendering routine can locally simulate the physical interaction with the remote environment. This allows for the local substitution of the remotely generated or sensed haptic force-feedback values. To ensure compliance of the predicted/generated signal with the actual input signal, a coherent model has to be built on the TOP side. In case of a deviation of more than the perceptual threshold, the actual current haptic force-feedback signals are transmitted and are used to update the surface models.

To minimize the number of samples necessary for approximating remote object surfaces, models with a small set of parameters should be used. Therefore, simple geometric models, such as a plane and/or a sphere are well suited for this purpose. This allows for modeling planar as well as concave and convex surface structures in the remote environment. The geometric model has to be assigned with a set of physical properties also estimated from the position-force pairs. In particular, impedance characteristics such as stiffness and damping are relevant for locally predicting remote force-feedback.

Figure 7 shows a screenshot of our first experiments for geometry-based prediction in a virtual environment with static objects. Our preliminary results show improved performance of geometry-based prediction in terms of data reduction and immersiveness compared to signal-based prediction.

One of the main challenges to be addressed in future work is to extend the prediction models by including additional surface properties, such as friction. Furthermore, to model movements of remote objects, translation and/or rotation are to be determined based on the received feedback and used for the local simulation.

EVENT-BASED CODING OF HAPTICS

Ideally, interacting with rigid objects through a TPTA system would be as simple as directly interacting with real rigid objects. For example, upon each tap on the surface of a piece of wood, immediately after first contact, sudden force and acceleration transients are felt followed by a more or less steady force to balance the force exerted by the human. These transients are characterized by the material impedance properties of the object like its mass, density, damping factor, and so on. Most conventional haptic rendering algorithms for virtual environments completely ignore any such high-frequency transients and only concentrate on displaying closed-loop quasi-static restoring forces according to Hooke's law or potential fields.

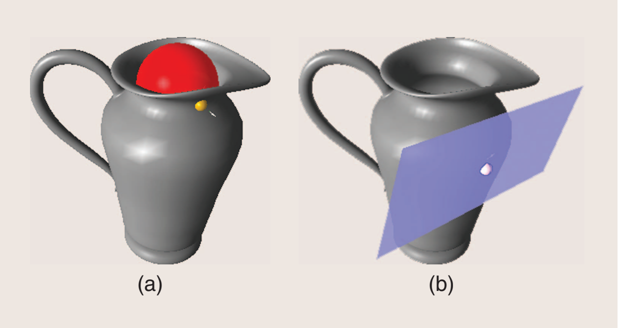
In this context, the paradigm of event-based haptics (EVBH) has been proposed [40], [41]. The authors propose and validate the idea that discrete events of contact with an object described by the contact time and velocity can be used to trigger the display of pre-computed force histories in an open-loop manner. Vibration waveforms arising out of a tapping task can be represented by an exponentially decaying sinusoid whose parameters depend on the material characteristics that we wish to display [42]

$$f_{\mathrm{acc}}(t) = A \cdot \dot{x} \cdot e^{-Bt} \cdot \sin(2\pi f t), \tag{7}$$

where $f_{\mathrm{acc}}(t)$ is the vibration waveform produced by first contact with the object and $A \cdot \dot{x}$ is the attack amplitude that is a function of the attack (or contact) velocity $\dot{x}$. For a given material, B is a decay constant matched to the apparent decay of the vibration waveform and f is the frequency of the attack portion of the wave (in Hz).

Transmitting these high-frequency contact transients over the network leads to increased sampling/packet rate requirements due to their high-frequency nature. Furthermore, in networked haptic environments, transmission delays occurring during communication particularly accentuate the criticality of control-loop stability issues when emitting high-frequency contact transients.

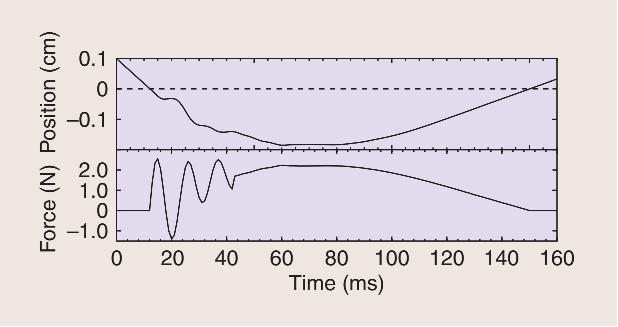
In [43], the Wellman and Howe model in [42] is combined with the idea of EVBH to develop a mathematical model for haptic contacts. To avoid additional packets being triggered by remotely superimposed contact transients, [43] proposes to shift the contact model from the remote server to the local client side to enable local model-based haptic contact transient rendering. This is supported by the open-loop nature of the proposed contact model. Locally deployed at the client side, the model is triggered by event-of-contact messages received from the remote side. The generated force transients are then locally superimposed onto the incoming low-frequency (LF) proportional feedback forces remotely generated, as shown in Figure 8. With these local operations, the transmission of high-frequency (HF) haptic signals over the network can be avoided while simultaneously improving the perceptual quality of haptic contacts. In other words, only the remotely generated LF-proportional feedback forces undergo PD data reduction, which leads to low update rates and hence high data reduction performance.



[FIG7] **Simulation of geometry-based prediction of haptic force-feedback based on simple geometric surface models. The little spheres represent the virtual endeffector. In (a) and (b), a spherical and planar surface approximation are illustrated, respectively.**

STABILITY ISSUES AND STABILITY-ENSURING CONTROL ARCHITECTURES

In real-world telepresence scenarios, transmission delays between the human operator and the TOP have to be taken into account. Even small time delays of several milliseconds may jeopardize the system stability resulting in dangerous destructive unbounded oscillations of the robotic devices. Several control architectures have been proposed to enable stable TPTA sessions in the presence of delay; see [2] and [3] for an overview. In particular, the so-called wave variable transformation (also called scattering transformation) [44], [45] is successfully deployed. Using this approach, stability for arbitrarily large, constant time delays can be ensured. Extensions for the stabilization in presence of varying time delays [46] and packet loss are available [47]. These approaches ensure the stability by preserving passivity within all elements of the communication system and the control architecture of the TPTA system. The haptic signals are transformed into the so-called wave domain where latency between the operator and the TOP no longer produces destabilizing effects. Wave variables consist of linear combinations of local haptic input



[FIG8] **Illustration of improving contact realism by superimposing acceleration forces during contact events. The upper part of the figure shows the trajectory of the haptic interface point as a function of time. The dashed horizontal line indicates the virtual object surface (object penetration occurs below this line). The bottom part of the figure shows the superimposition of event-triggered force contact transients on traditionally rendered penetration-based feedback forces.**

and output, including received remote wave variables, and can be written as follows:

$$u_m = \frac{1}{\sqrt{2b}}(F_m^d + b\dot{x}_m), \quad u_s = \frac{1}{\sqrt{2b}}(F_s + b\dot{x}_s^d),$$
$$v_m = \frac{1}{\sqrt{2b}}(F_m^d - b\dot{x}_m), \quad v_s = \frac{1}{\sqrt{2b}}(F_s - b\dot{x}_s^d), \qquad (8)$$

where $b > 0$ defines a characteristic impedance associated with the wave variables and represents a tuning parameter that directly affects the system behavior.

The wave variables u_m (forward path) and v_s (backward path) are transmitted over the communications network and arrive at the corresponding receiver with a time delay $T_1, T_2 > 0$, hence,

$$u_s(t) = u_m(t - T_1) \text{ and } v_{\mathrm{m}}(t) = v_{\mathrm{s}}(t - T_2).$$

Although the wave variables originate from local and remote haptic signals, they do not directly relate to human haptic perception anymore. Hence, the previously discussed perceptual haptic data reduction methods are no longer directly applicable. They need to be redesigned and adjusted when applied in the wave domain.

This is nontrivial because it is unknown so far whether perceptual resolution limits like Weber's law can be appropriately defined in the wave variable domain. In fact, first investigations [18] suggest that the Weber-based deadband approach is not the most efficient one, but others might be superior. Perceptual data reduction in the wave variable domain is thus challenging and should be addressed in future work.

RECORDING AND PERCEPTUAL OFF-LINE CODING OF HAPTIC DATA

Quite the contrary to real-time processing of haptic datastreams, recording and off-line compression of haptic data furnishes the operator with the possibility to store and replay a haptic TPTA session.

Although during replay, the displayed force feedback does not relate to the user's interaction anymore, the human perception system can easily associate the previously recorded movements and actions with the visual content and relate the displayed force feedback to it [28]. In this manner, the users can immerse into the previously recorded telemanipulation session.

In Figure 9, a schematic overview of a haptic recorder/haptic player demonstrator setup is illustrated (also see [28]). During recording, the visual and force feedback information received from the remote environment is synchronously stored to a local data storage system. To enable posterior playback of the recorded multimodal TPTA session, the previously recorded and compressed multimedia content is decoded and redisplayed through the HSI.

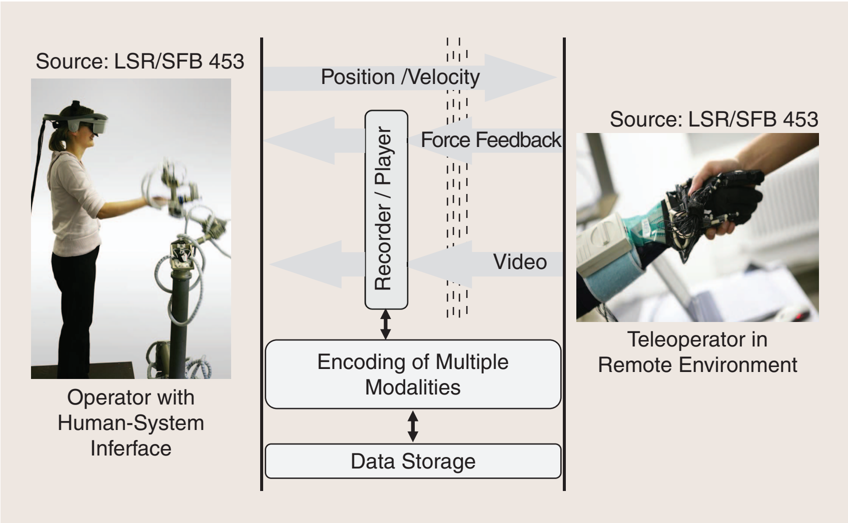
For off-line haptic compression, due to the relaxed delay constraints, the challenge of designing a haptic compression scheme can be addressed in a fundamentally different way. For instance, we are no longer restricted to process the signals sample by sample. Instead we can now perform a comprehensive investigation of the whole haptic recording at once or in parts in the time and/or frequency domain. Here, analysis-by-synthesis techniques applied in audio and video coding are of interest. By simulating the process of decoding during encoding, the distortion that would be introduced can be exactly determined. As a result, the coding parameters can be optimized such that the distortion stays below human perception thresholds.

Furthermore, the signal reconstruction process at the decoder strongly influences the irregular sampling process of the PD data reduction approach introduced in the section "Haptic Data Reduction Using Perceptual Deadbands." This means that a minimum set of samples required for lossy compression can be successively and iteratively generated. This approach is closely related to the theory of irregular sampling that determines optimum subsets of signal samples at minimal information loss [48].

Extending this approach to incorporate a mathematical model of human haptic perception is one of the major objectives of our future work. Furthermore, for off-line compression, the processing of large blocks of discrete samples is possible and in most cases favorable. It allows for adopting and adjusting widely used methods from audio, image, and video compression. In particular, the integration of temporal domain and frequency domain masking in the psychophysical model of human haptic perception is promising in this context and should be investigated by the signal processing community.

MULTIMODAL SIGNAL MULTIPLEXING

In a TPTA system, several modalities need to be simultaneously encoded and



[FIG9] **Schematic overview of a multimodal recording and playback architecture.**

transmitted. The properties of visual, auditory, and haptic data transmission do not only differ with respect to the required transmission rate, but also vary dramatically in terms of delay constraints. Delay introduced in the haptic modality is most critical and hence the transmission of haptic information needs to be prioritized. Furthermore, due to different processing delays for each modality, cross-modal asynchrony becomes an issue and can lead to temporally inconsistent representations. By investigating human cross-modal temporal integration, upper and lower latency bounds for each modality can be determined [49]. By applying prioritization rules to each modality, this allows us to design a multiplexer for improving the performance of multimodal communication in haptic telepresence scenarios. A recent example for multiplexing of audio, video, and haptics can be found in [50]. First steps towards the use of standard protocols for multimodal media streaming can be found in [51], where the use of MPEG-4 BIFS for haptic signal transmission is proposed.

CONCLUSIONS

In this article, we discuss the emerging field of haptic data compression and communication. We lay special emphasis on haptic data transmission in the context of telepresence and teleaction scenarios, where high packet rates and communication delay are the main challenges. Recent advances in this field employ a mathematical model of human haptic perception to determine which haptic samples to transmit. This allows us to significantly reduce the packet rate while at the same time keeping the introduced distortion below human perception thresholds. Perceptual coding of haptic signals has been successfully combined with predictive coding and EVBH. As a result, both data reduction performance and user experience could be further improved. The article also briefly discusses off-line coding and multiplexing of audio-video-haptic data streams.

Although significant progress has been made over the past few years, a multitude of challenges remains to be addressed. We have outlined some of the major issues in the corresponding sections of this article and hope that the signal processing community will address these during the years to come.

ACKNOWLEDGMENTS

This work has been supported, in part, by the German Research Foundation (DFG) within the Collaborative Research Centre SFB 453 on high-fidelity telepresence and teleaction and project STE 1093/4-1.

AUTHORS

Eckehard Steinbach (eckehard.steinbach@tum.de) studied electrical engineering at the University of Karlsruhe, Germany, University of Essex, Colchester, United Kingdom, and ESIEE, Paris, France. He received the engineering doctorate from the University of Erlangen-Nürnberg, Germany, in 1999. From 1994 to 2000, he was a member of the research staff of the Image Communication Group, University of Erlangen-Nürnberg. From February 2000 to December 2001, he was a postdoctoral fellow with the Information Systems Laboratory, Stanford University, California. In February 2002, he joined the Department of Electrical Engineering and Information Technology, Technische Universität München, Munich, Germany, as a professor of media technology. His current research interests are in the area of audio-visual-haptic information processing, image and video compression, error-resilient video communication, and networked multimedia systems. He is a Senior Member of the IEEE.

Sandra Hirche (hirche@tum.de) received the diploma engineer degree in mechanical engineering and transport systems in 2002 from the Technical University Berlin, Germany, and the doctor of engineering degree in electrical engineering and computer science in 2005 from the Technische Universität München, Munich, Germany. From 2005 to 2007, she was a postdoctoral at the Tokyo Institute of Technology, 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 networked control systems, cooperative control, human-in-the-loop control, and haptics. Since 2009, she has been the chair for student activities in the IEEE Control System Society. She is a Member of the IEEE.

Julius Kammerl (kammerl@tum.de) studied computer science at the Technische Universität München in Munich, Germany. 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 in Erlangen, Germany, he joined the Institute for Media Technology at the Technische Universität München in 2006, where he is currently working as a member of the research and teaching staff. His research interests are in the field of haptic communication with a focus on perceptual coding 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. He is a Member of the IEEE.

Iason Vittorias (vittorias@tum.de) received his diploma engineer degree in electrical and computers engineering in 2007 from the Aristotle University of Thessaloniki, Greece. Since 2008, he has been a research assistant at the Institute of Automatic Control Engineering, Technische Universität München, Munich, Germany, pursuing his Ph.D. degree. His research interests include teleoperation systems over networks, passivity-based control, and haptic data reduction.

Rahul Chaudhari (rahul.chaudhari@tum.de) received the M. Sc. degree in communication systems from the Technische Universität München, Munich, Germany in 2009, focusing on signal processing and compression/reduction of data for haptic communication. He received an undergraduate degree (bachelor of engineering) in electronics and telecommunications from the University of Pune, India, graduating in 2006 as the top engineering student in his class.

REFERENCES

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

[2] S. Hirche, M. Ferre, J. Barrio, C. Melchiorri, and M. Buss, "Bilateral control architectures for telerobotics," in *Advances in Telerobotics: Human System Interfaces, Control, and Applications (STAR series)*, M. Ferre, M. Buss, R. Aracil, C. Melchiorri, and C. Balaguer, Eds. New York: Springer-Verlag, 2007, pp. 163–176.

[3] P. Hokayem and M. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 49, no. 12, pp. 2035–2057, Dec. 2006.

[4] G. Sankaranarayanan, L. Potter, and B. Hannaford, "Measurement and emulation of time varying packet delay with applications to networked haptic virtual environments," in *Proc. 1st Int. Conf. Robot Communication and Coordination (RoboComm'07)*. Piscataway, NJ: *IEEE Press*, 2007, pp. 1–8.

[5] V. Hayward and K. MacLean, "Do it yourself haptics, Part I," *IEEE Robot. Automat. Mag.*, vol. 3, no. 2, pp. 88–104, 2008.

[6] 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 and Expo (ICME'01),* Los Alamitos, CA, Aug. 2001, pp. 563–566.

[7] 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 Perform. Eval. Rev.*, vol. 29, no. 1, pp. 334–335, 2001.

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

[9] A. Ortega and Y. Liu, *Lossy Compression of Haptic Data*. Englewood Cliffs, NJ: Prentice-Hall, 2002, ch. 6, pp. 119–136.

[10] C. Borst, "Predictive coding for efficient host-device communication in a pneumatic force-feedback display," in *Proc. 1st Joint Eurohaptics Conf. Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, World Haptics 2005, Pisa, Italy, Mar. 2005, pp. 596–599.

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

[12] 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. Int. Conf. Acoustics, Speech, and Signal Processing,* Philadelphia, PA, vol. 2, Mar. 2005, pp. ii/1097–ii/1100.

[13] 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 Processing*, vol. 56, no. 2, pp. 588–597, Feb. 2008.

[14] M. 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.

[15] S. Hirche, M. Buss, P. Hinterseer, and E. Steinbach, "Network traffic reduction in haptic telepresence systems by deadband control," in *Proc. Int. Federation of Automatic Control (IFAC) World Congress,* Prague, Czech Republic, July 2005.

[16] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss, "Towards deadband control in networked teleoperation systems," in *Proc. Int. Federation of Automatic Control (IFAC) World Congress,* Prague, Czech Republic, 2005.

[17] S. Hirche, P. Hinterseer, E. Steinbach, and M. Buss, "Transparent data reduction in networked telepresence and teleaction systems. Part 1: Communication without time delay," *Presence: Teleoper. Virtual Environ.*, vol. 16, no. 5, pp. 523–531, 2007.

[18] S. Hirche and M. Buss, "Transparent data reduction in networked telepresence and teleaction systems. Part 2: Time-delayed communication," *Presence: Teleoper. Virtual Environ.*, vol. 16, no. 5, pp. 532–542, 2007.

[19] J. Kammerl, P. Hinterseer, S. Chaudhuri, and E. Steinbach, "A theoretical analysis of data reduction using the Weber quantizer," in *Proc. Int. Data Compression Conf. (DCC),* Snowbird, UT, Mar. 2008, pp. 524–524.

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

[21] N. Sakr, J. Zhou, N. Georganas, J. Zhao, and X. Shen, "Prediction-based haptic data reduction and compression in tele-mentoring systems," in *Instrumentation and Measurement Technology Conf. Proc. (IMTC'08). IEEE,* Austin, TX, May 2008, pp. 1828–1832.

[22] 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, 2008.

[23] Y. You and M. Y. Sung, "Haptic data transmission based on the prediction and compression," in *Proc. IEEE Int. Conf. Communications,* May 2008, pp. 1824–1828.

[24] J. Kammerl, I. Vittorias, V. Nitsch, B. Faerber, E. Steinbach, and S. Hirche, "Perception-based data reduction for haptic force-feedback signals using velocity-adaptive deadbands," *Presence: Teleoper. Virtual Environ.,* submitted for publication.

[25] P. Liu, M. Meng, and S. Yang, "Data communications for internet robots," *Auton. Robots*, vol. 15, no. 3, pp. 213–223, 2003.

[26] S. Lee and J. Kim, "Priority-based haptic event filtering for transmission and error control in networked virtual environments," *Multimedia Syst.*, vol. 15, no. 6, pp. 355–367, 2009.

[27] M. Fujimoto and Y. Ishibashi, "Packetization interval of haptic media in networked virtual environments," in *Proc. 4th ACM SIGCOMM Workshop Network and System Support for Games*, Portland, OR, 2005, p. 6.

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

[29] E. Weber, *Die Lehre vom Tastsinn und Gemeingefuehl, auf Versuche gegruendet.* Braunschweig, Germany: Vieweg, 1851.

[30] G. A. Gescheider, *Psychophysics.* Lawrence Erlbaum, 1985.

[31] H. Tan, M. Srinivasan, B. Eberman, and B. Cheng, "Human factors for the design of force-reflecting haptic interfaces," in *Proc. ASME WAM*, 1994, pp. 1–11.

[32] G. C. Burdea, *Force and Touch Feedback for Virtual Reality*. New York: Wiley, 1996.

[33] J. Greenspan and S. Bolanowski, "The psychophysics of tactile perception and its peripheral physiological basis," in *Pain and Touch*, L. Kruger, Ed. San Diego, CA: Academic, 1996, pp. 25–104.

[34] S. Allin, Y. Matsuoka, and R. Klatzky, "Measuring just noticeable differences for haptic force feedback: Implications for rehabilitation," in *Proc. 10th Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS*, Orlando, FL, 2002, pp. 299–302.

[35] P. Hinterseer, E. Steinbach, and S. Chaudhuri, "Perception-based compression of haptic data streams using Kalman filters," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing,* Toulouse, France, May 2006.

[36] N. Sakr, N. Georganas, and J. Zhao, "A perceptual quality metric for haptic signals," in *Proc. IEEE Int. Workshop on Haptic, Audio and Visual Environments and Games (HAVE)*, Ottawa, Canada, 2007, pp. 27–32.

[37] P. Hinterseer and E. Steinbach, "A psychophysically motivated compression approach for 3D haptic data," in *Proc. 14th Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*, Arlington, VA, Mar. 2006, pp. 35–41.

[38] H. Pongrac, B. Färber, P. Hinterseer, J. Kammerl, and E. Steinbach, "Limitations of human 3d force discrimination," in *Human-Centered Robotics Systems, Munich*, Germany, Oct. 2006.

[39] P. Mitra and G. Niemeyer, "Model-mediated telemanipulation," *Int. J. Robot. Res.,* vol. 27, no. 2, p. 253, 2008.

[40] J. D. Hwang, M. D. Williams, and G. Niemeyer, "Toward event-based haptics: Rendering contact using open-loop force pulses," in *Proc. Int. Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems.* Los Alamitos, CA: IEEE Computer Society, Apr. 2004, pp. 24–31.

[41] K. J. Kuchenbecker, J. Fiene, and G. Niemeyer, "Improving contact realism through event-based haptic feedback," *IEEE Trans. Visual. Comput. Graphics*, vol. 12, no. 2, pp. 219–230, 2006.

[42] P. Wellman and R. D. Howe, "Towards realistic vibrotactile display in virtual environments," in *Proc. ASME Int. Mechanical Engineering Congress and Exposition Symp. Haptic Interfaces for Virtual Environment and Teleoperator Systems*. San Francisco, CA, Nov. 1995, vol. 57, no. 2, pp. 713–718.

[43] J. Kammerl, R. Chaudhari, and E. Steinbach, "Haptic contact models for improved contact rendering and perceptual coding in networked virtual environments," *IEEE Trans. Instrum. Meas.,* submitted for publication.

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

[45] G. Niemeyer and J. Slotine, "Stable adaptive teleoperation," *Int. J. Oceanic Eng.*, vol. 16, no. 1, pp. 152–162, 1991.

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

[47] S. Hirche and M. Buss, "Packet loss effects in passive telepresence systems," *in Proc. 43rd IEEE Conf. Decision and Control,* Paradise Island, Bahamas, 2004, pp. 4010–4015.

[48] H. Feichtinger, "Coherent frames and irregular sampling," in *Recent Advances in Fourier Analysis and Its Application (NATO ASI Series C)*, 1989, vol. 315, Pisa, Italy, pp. 427–440.

[49] 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. 1, no. 1, pp. 28–36, 2009.

[50] M. Eid, J. Cha, and A. E. Saddik, "An adaptive multiplexer for multi-modal data communication," in *Proc. Int. Workshop on Haptic Audio-Visual Environments and Games,* Lecco, Italy, Nov. 2009, pp. 111–116.

[51] 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. World Haptics Conf.*, 2007, pp. 274–279.