Opening the Haptic Loop: Network Degradation Limits for Haptic Task Performance

Rahul Chaudhari*, Clemens Schuwerk*, Verena Nitsch†, Eckehard Steinbach* and Berthold Färber†

*Institute for Media Technology, Technische Universität München, Germany
†Human Factors Institute, Universität der Bundeswehr, Munich, Germany
Email: {rahul.chaudhari, clemens.schuwerk, eckehard.steinbach}@tum.de
Email: {verena.nitsch, berthold.faeber}@unibw.de

Abstract—For realtime teleoperation with haptic feedback, network-induced artifacts like delay, packet loss and lossy coding of haptic data deteriorate the operability of the system. We provide a systematic quantification of the limits within which these network-induced degradations are tolerable for human task performance in executing a given haptic task. These limits are conservative in the sense that we do not account for any stabilizing approaches from control engineering to counter the impact of the network-induced degradations. Through experimental evaluation with a pursuit tracking task, we show that task performance is most affected by delay on the network and a strong lossy coding scheme. With statistical analysis we show that task performance is significantly decreased for one-way delay higher than 14 ms and strong lossy coding with a deadband parameter higher than 27%. For the given task, packet loss is found not to affect task performance significantly.

I. INTRODUCTION

Haptic teleoperation in remote environments is made feasible by existing communications infrastructure like the Internet. This extensive gain in outreach, nevertheless, is plagued with its own set of problems. Realtime teleoperation (see Figure 1 for a conceptual overview) imposes exacting demands on the communication network, in terms of volume and rates of data traffic generated and strong delay constraints. Efficient haptic data communication and coding schemes have been introduced in the past to address these concerns [1]. Subsequently, they have repeatedly been proven to be very successful at limiting the network data transfer rates, making realtime haptic interactions tractable.

Like every real physical system, communication networks introduce finite artifacts into the haptic interaction. On the other hand, the human user can also tolerate some disturbances arising from the network and the communication schemes. In other words, in presence of degrading artifacts, a haptic task can still be executed without significant deterioration in performance [2], [3], [4], [5], [6]. In this paper, we aim to provide a systematic quantification of the limits within which the network induced degradations are tolerable to the human user in executing a given haptic task. We achieve this by performing extensive experiments with subjects interacting with a dynamic virtual environment (VE). The haptic task and the VE have been chosen to be representative of typical teleoperation applications, and comprehensive objective analyses of the haptic interactions are presented. At this point, to place this paper in a larger context, two fundamental motivations for this work meriting attention are explained next.

Objective quality evaluation (OQE) of haptic communication schemes requires not only a perception model of the human haptic system, but also a human action model, on account of bidirectional physical interaction with the system [7]. In order to simplify the complexity of such an undertaking, we intend to separate the perception modeling from the action modeling. Obviously, such a decomposition is only reasonable if the interaction between these two components is small. In [8], studies performed by Witmer and Singer led to the conclusion that although there does exist a coupling between presence and task performance for teleoperation, this link is indeed very weak. The results from the present work can be used to determine the network degradation limits, within which, haptic tasks can be executed satisfactorily. If we stay within these limits, OQE for haptic interaction can ignore the forward (action) channel and concentrate on the reverse (perception) channel.

Besides OQE, knowing the degradation limits on the task performance is also essential for designing shared haptic virtual environments (SHVE). If multiple users can explore a virtual environment simultaneously, several challenges like scalability and synchronization between the users arise. Network impairments and their effect on the overall task performance become even more important in this case. Two common communication architectures, client-server and Peer-to-Peer (P2P), are deployable for an SHVE and both have their own pros and cons [9]. But both approaches struggle with the same network-induced degradations - namely delay, jitter and high packet rates.

We state that there are different types of haptic teleoperation tasks depending more either on the user experience of the haptic feedback or on the haptic task performance. A perfect scenario would be where the user could not feel any degradation; this would then lead to the best task performance. But as network-induced degradations cannot be avoided, we think that it is reasonable to reduce the demands on the communication channel as long as a given task can still be
completed satisfactorily. This would be of course at the cost of degradations in perceived haptic feedback.

The organization of this paper is as follows. In Section II, we review related work done in the area of task performance evaluation for haptic teleoperation. We also summarize Quality-of-Service (QoS) determination and provisioning approaches for networked haptics. Section III describes typical network impairments that occur in haptic data transmission over packet-switched communication networks. Section IV describes the design of the experimental evaluation and procedure. In Section V, we analyze the experimental data obtained, test it against objective performance measures and summarize the results. Finally in Section VI, we conclude with some limitations of this work and how it can be improved and expanded upon in the future.

II. RELATED WORK

A number of researchers have studied the effect of network delays on haptic task performance previously. For example, in experiments conducted in [4], [5], subjects performed target acquisition tasks with a haptic device in the presence of delays on the haptic and/or visual channels. In [6], for a cube stacking task, effects of network delay on objective task performance measures like task completion time, failed attempts at lifting the cube up and the quality of stacking were tested. For haptic network traffic, Marshall et al. [10] establish Quality-of-Service (QoS) requirements and propose architectures for QoS provisioning. In [11], the authors investigate the influence of constant network delay on haptic collaboration with subjective tests and objective measurements. All of these works, however, consider haptic tasks in which the start- and end-points of motion are specified, but not the intermediate trajectory. As a result, the haptic signals generated out of the haptic interaction can vary widely across two different evaluation runs. This makes it impossible to compare task performance between the two runs based on sample-by-sample comparisons of signals. We perform our study around a controlled pursuit tracking task, the rationale behind which is explained in Section IV-A.

The authors of [12] investigate the effects of packet loss and time-varying delay in haptic telepresence systems operating over packet-switched networks. This work studies the impact of these network-induced degradations on stability of the deployed robot control architectures. Additionally, system performance criteria like the magnitude of maximum displayable force feedback as well as fidelity of the teleoperator robot’s motion with operator-commanded motion are also studied. In [13], the same authors further analyze the influence of packet-loss and the corresponding compensating packet-processing strategies on stability and system performance. Similar to [12], human operator performance in haptic tasks was irrelevant to the scope of [13].

Other researchers have studied the influence of haptic data reduction (lossy coding) on perceived haptic quality and human task performance in virtual simulations of telepresence systems [3] as well as real-world systems [2]. Prompted by safety-critical operations, where there is little margin for error, performance criteria like surface penetration depth and its variance are considered for virtual haptic simulations. Beyond lossy data reduction dealt with in [2] and [3], we also include delay and packet loss in our analyses. Moreover, our motivations for doing so are fundamentally different, as described in the introduction. We conduct our experiments in a more realistic virtual environment with dynamically moving objects. Therefore, the results we present are also interesting for the design of shared haptic environments operating over a network, where implementation of static environments would otherwise be trivial.

III. TYPICAL NETWORK COMMUNICATION ARTIFACTS

Delay/jitter

Delay and its time variance (jitter) in packet-switched networks originates from factors like signal processing, dynamic multi-path routing of packets, background traffic, congestion at intermediate routers, etc. In this work, we deal with delay on the haptic channel only, with the haptic feedback always arriving later than the visual one. Nevertheless, if the delay is known (or measured) and constant, the two modalities can be synchronized by simply delaying the visual feedback to match the haptic one. But for time-varying delay, this would not be as easy. In any case, we presently neglect these additional complexities, and view the delay as being a discrepancy between the two feedbacks, without pinning it to a specific cause.

Packet loss

When the Internet is used as a communication channel between the human operator and the remote teleoperator, packet loss becomes a pivotal factor influencing the quality-of-service (QoS). Researchers have emphasized previously on this fact for other delay-sensitive multimedia applications [14]. Several models accounting for packet losses in the Internet have been employed in research on networked multimedia. Correlated packet-losses (commonly called bursty losses) have been modeled successfully using the Gilbert-Elliot model, which we also use for our simulations.

Lossy coding

The state-of-the-art approach for online lossy haptic coding reduces the number of transmitted haptic samples using an irregular subsampling process [1]. This process exploits limitations of human haptic perception, represented by Weber’s
law of psychophysics. According to this law, a stimulus \( I \) is compared with a reference stimulus \( I_{\text{ref}} \), \( \Delta I \) being the difference between them. This difference is only perceivable if it exceeds the Just Noticeable Difference (JND). Such a relationship can be mathematically described as \( \Delta I / I_{\text{ref}} = k \) where \( k \) denotes the deadband parameter and can be empirically determined for several types of stimuli such as force, velocity, pain and temperature [16].

We translate this to haptic data reduction in the following manner. Whenever a haptic sample \( f(n) \) falls within the so-called deadband thresholds \( f(n) \pm \Delta f(n) \), defined by the last transmitted haptic sample \( f(n) \), it is dropped from transmission. At the receiver side, the dropped sample is estimated by holding the last received haptic sample \( f(n) \). On the other hand, if the haptic sample \( f(n+1) \) violates the perception thresholds, the inter-sample difference from \( f(n) \) to \( f(n+1) \) is said to be perceivable and \( f(n+1) \) needs to be transmitted. Following this transmission, the last transmitted sample is updated to its new value as \( f(n) = f(n+1) \), and the process repeats.

**Stabilizing approaches for network artifacts**

In the control engineering literature, a number of passivity-based stabilization methods have been proposed to counter potential instabilities in teleoperation arising from network-induced artifacts. For example, transmission of wave-variables instead of haptic signals for countering constant time-delay [16], filtering or rateBounding for stable compensation of packet-loss [13] and energy-supervised reconstruction for lossy deadband coding [17]. For the present work, however, we do not implement any of these in determining network degradation limits on task performance. Assuming that such stabilizing strategies will only relax the limits we determine, we can safely operate inside them for our purposes.

**IV. EXPERIMENTS**

**A. Task description**

In this work, we design our evaluation methodology around a pursuit tracking task. A number of factors drive this choice. First and foremost, it is a good abstraction of a variety of elementary tasks commonly performed in haptic teleoperation, wherein an object is moved from a start-location to an end-location following a certain path.

Secondly, being able to specify a reference signal to follow brings along a number of advantages. Haptic tasks, on account of their bidirectional nature may lead to widely varying haptic interactions across subjects. Tracking tasks, for a given reference signal, significantly reduce this variability and establish a common basis for evaluation across subjects [7]. In addition, such a design furnishes us with the possibility to devise convenient objective performance measures (elaborated upon in Section V-A) based on sample-by-sample comparisons of clean signals and signals distorted by network-induced degradations. Moreover, tracking tasks have been widely studied in manual control to propose and identify human control action models [18], [19].

**B. Experienced human operators**

At this juncture, it should be noted that we study human performance for the purposes of engineering design. As analyzed in [20], general human behavior is too complex and variable to lend itself willingly to quantitative modeling. In the context of man-machine interaction, prediction of human performance is desired for situations in which the human operator behaves rationally, skillfully and in a goal-oriented manner. Hence we distinguish between an operator and a layperson, which is reflected in our choice of subjects for the experimental evaluation.

Thirteen subjects participated in the evaluation, 2 of them female and 11 male. The participants’ age was in the range of 23-31 with a mean of 26.8 years and standard deviation of 2.4 years. All of them were right-handed. Six of them are researchers working with haptic data compression and are daily users of desktop haptic devices. The rest have frequently participated in subjective tests for evaluation of haptic compression schemes in the past and were therefore familiar with haptic devices. Nevertheless, before the actual experiment, they were trained in a training session to ensure they were at par with the more experienced participants for the given task.

**C. Experimental procedure**

As mentioned before, to eliminate individual differences related to experience with the experimental setup, or differences in understanding the instructions, a training session was conducted at the beginning of the experiment. Herein, the task to be performed was explained to the subjects and they had the chance to get familiar with the task execution, the experimental setting and the parameter variations. We used a within-subject experimental design, with all participants executing the same haptic task under the same parameter settings.

The pursuit tracking task consisted of a one-DoF dynamic virtual environment (VE) (refer to Figure 2). In this VE, the “reference” cube (white) was to be pursued as closely as possible with the “pursuit” (green) cube. The reference cube was only a visual reference, and could not be haptically touched. The subjects could, on the other hand, haptically interact with the pursuit cube and push it to follow the white cube. The start- and the end-points of the pursuit task were fixed. The velocity of the reference cube was controlled to have a Gaussian function profile with the Gaussian mean set to 2.0 \( \text{cm/s} \) and the standard deviation to 0.85 \( \text{cm/s} \). This afforded a reasonable degree of control over the reference
input, while ensuring smooth variation in the pursuit velocity. The choice of the velocity profile is also motivated with the shape of a typical velocity profile that a human arm follows in reaching from one point to another in task space [21]. This task to be executed remained the same throughout the experiment.

The main experiment, after the training session, contained four sub-experiments, characterized by which independent variable \( D = \) one-way network delay, \( DBK = \) deadband parameter \( k \) and \( PL = \) packet loss) was controlled, and which ones were kept constant:

1) variation of \( D \) in the range 0 to 20 ms, with \( DBK = 0 \) and \( PL = 0 \% \),
2) variation of \( DBK \) in the range 0 to 80 \%, with \( D = 0 \) and \( PL = 0 \% \),
3) variation of \( PL \) in the range 0 to 75 \%, with \( D = 0 \) and \( DBK = 0 \% \),
4) variation of \( PL \) in the range 0 to 75 \%, with \( D = 0 \) and \( DBK = 10 \% \).

The range of this independent variables was determined in pilot tests, where we found these values to be reasonable for our experiment setup. The choice of \( DBK = 10 \% \) in sub-experiment 4 is motivated by the results in [1].

Each sub-experiment had 7 runs. In a random manner, the order of sub-experiments changed from one subject to another, along with the order of the controlled variable values. This was done to prevent prediction of trends by the subjects and to avoid any systematic manifestation of fatigue in the results of the evaluation. The seating posture and hand-device configuration was standardized across subjects. In order to prevent distractions due to noise emanating from the haptic device, they wore headphones playing music. In addition, distractions due to visual observation of the haptic device were avoided by blocking the view to the haptic device by a cardboard. The PHANToM Omni haptic device (Sensible Technologies Corp., Woburn, MA) was used for the experiments.

V. ANALYSIS OF COLLECTED DATA

A. Objective task performance measures

We describe here various measurements used in the experiment in order to score the task performance objectively. As we concentrate on a pursuit tracking task, an obvious measurement of task performance would be the error between the reference and the actual velocity profiles. We also measure the variance of force displayed to the user. Ideally, the force profile would be similar to the reference velocity profile (up to user variability), but artifacts in haptic display due to network impairments may lead to disturbances from the nominal at the haptic device. This is why the user cannot control the exerted force as accurately as would be necessary to achieve perfect task performance. Another measurement is the contact time of the haptic interface point (HIP) with the cube relative to the total task execution time. Ideally, the user pushes the cube and adjusts his velocity according to the reference smoothly by increasing or decreasing the exerted force. Disturbances from the network impairments, however, may prevent continuous contact. Relative contact time is also a measurement for instability in the haptic loop, which leads to sudden high forces being displayed. All the objective measurements are calculated in real world units, representing the actual velocity or position displacement of the stylus of the haptic device.

B. Results

Aiming to ascertain whether any measured impairments in task performance could be confidently attributed to the experimental manipulation rather than chance, statistical analyses were conducted with the obtained data. Due to the small sample size, non-parametric tests were chosen as a normal distribution could not be assumed [22]. Consequently, main effects were investigated with Friedman’s Analysis of Variance (ANOVA). Significant findings were followed up with Wilcoxon tests, comparing each measurement in the experimental conditions to those in the respective baseline conditions. In order to minimize the inflated risk of Type I errors (false positive findings) due to multiple comparisons, a Bonferroni correction was applied to the accepted \( \alpha \)-level. Consequently, observed differences in task performance were only accepted as significant if \( p < .008 \). For significant findings, effect sizes were also calculated as Pearson correlation coefficients (\( r \)). According to [23], a value of \( r = .10 \) signifies a small effect, a value of \( r = .30 \) a medium effect. A value of \( r \geq .50 \) denotes a large effect, as it indicates that the experimental manipulation accounts for more than 25% of variance in the observed scores.

a) The effect of time delay on task performance: In Figure 4a, the mean square error (MSE) between the reference velocity and the traced velocities can be seen to increase...
with increasing delay. This suggests that increasing delay interferes with the task performance. Friedman’s ANOVA confirmed that the difference in reference and tracing velocities was significantly affected by the introduction of time delay ($\chi^2(6) = 64.45, p < .001$). Follow-up Wilcoxon tests showed a significant deterioration in this performance measure with a time delay of 14 ms. ($z = -3.04, p < .008, r = -.84$) and higher. The same is suggested by profiles of the relative contact time and the force variance. Relative contact time can be seen to decrease while force variance increases with increasing delay. Physically, this delay interference manifests itself by causing instability of haptic interaction with the virtual object (see Figure 3). The ANOVA confirmed this finding with respect to relative contact time ($\chi^2(6) = 69.00, p < .001$). Here, Wilcoxon tests showed a significant deterioration in relative contact time with a time delay of 17 ms. ($z = -2.62, p < .008, r = -.73$) and higher. The statistical analyses also confirmed a significant influence of time delay on the force variance ($\chi^2(6) = 62.04, p < .001$). A significant performance impairment was observed with a time delay of 14 ms. ($z = -3.04, p < .008, r = -.84$) and higher. These findings are in accordance with the results presented in [11], which found that the mean-opinion-score (MOS) for task performance significantly decreases for a complete round trip delay of about 30 milliseconds.

b) The effect of deadband coding on task performance:
Similar conclusions can be drawn from the sub-experiment comprising of the strength of the deadband coding scheme (Figure 4b). The highly non-linear distortion introduced into the haptic interaction on account of the lossy coding scheme, increasingly degrades the task performance according to the criteria considered here. The user’s ability to match their velocity to that of the target object was significantly affected by the deadband coding scheme ($\chi^2(6) = 57.30, p < .001$), and was significantly impaired with deadband parameters of $k = 67$ ($z = -3.11, p < .008, r = -.86$) and higher. The force variance also indicated an effect of deadband coding on task performance ($\chi^2(6) = 66.33, p < .001$). This aspect of task performance was particularly sensitive to its effects, as it was significantly affected with a deadband parameter as low as $k = 27$ ($z = -2.97, p < .008, r = -.82$). Although relative contact time varied significantly with the different deadband parameters ($\chi^2(6) = 18.30, p < .05$), none of the individual follow-up comparisons to the baseline conditions reached significance at $p < .008$.

c) The effect of packet loss on task performance:
Finally, the interference of packet-loss in the haptic interaction is not found to be as strong as for the previous two sub-experiments with delay and deadband coding. Figures 4c and 4d show more or less random variations in the values of the objective measures, without any evidence of a conclusive pattern. The statistical analysis confirmed this observation for each performance measure (velocity MSE: $\chi^2(6) = 8.80, p = .19$; contact time: $\chi^2 = 4.44, p = .06$; force variance: $\chi^2 = 12.33, p = .06$). This phenomenon can be explained as follows. A haptic signal affected by packet-loss is reconstructed according to the hold-last-sample (HLS) method. The selection of the smooth Gaussian velocity profile for the motion of the reference cube inherently limits the pursuit tracking task to the low-frequency region. Therefore, the HLS reconstruction of the haptic feedback signal works very well, which explains minimal interference of packet-loss with the haptic interaction and therefore the task performance. The sub-experiment in which packet loss was combined with a deadband coding scheme ($k = 10$) led to similar (non-significant) results.

The results of the delay sub-experiment (Figure 4a) for all three performance criteria are seen to consistently have larger scales as compared to those of the deadband coding sub-experiment (Figure 4b) for the respective ranges. The velocity MSE for example is increasing up to 3 cm$^2$/s$^2$ for the delay sub-experiment, while it increases up to 1 cm$^2$/s$^2$ for the deadband sub-experiment. Additionally, the amplitude...
scales for the task performance criteria in case of the packet-loss sub-experiments are considerably smaller than those of the previous two sub-experiments. The velocity MSE stays below 0.5 cm²/s² for the complete range of the independent variables. This corroborates the previously stated observation with respect to packet loss.

VI. CONCLUSIONS AND FUTURE WORK

This work establishes limits on the network-induced degradations, beyond which task performance starts getting affected unacceptably, for the given pursuit tracking task. We found that delay as well as lossy coding affect task performance adversely, with a monotonic degradation for the task performance measurement criteria we use. For the given task, packet-loss was found to not affect task performance significantly. Since we do not account for any stabilizing approaches from control engineering to counter the impact of the network-induced degradations, the limits that we determine are conservative.

As mentioned before, in the future, we plan to concentrate on perception modeling for haptic communication, while remaining within these degradation limits. We plan to conduct subjective psychophysical tests for our experimental setup followed by a multi-dimensional analysis of the results. This will help us to extract the perceptual dimensions of the subjective Quality-of-Experience when dealing with haptic feedback in networked haptic applications, and their relative importance scales. Knowledge of the underlying perceptual dimensions will contribute towards arriving at a more elaborate model for objective evaluation of haptic perceptual quality in networked haptics.

For the future we also plan to investigate haptic tasks where two or more users collaboratively contribute to task execution. Performance limits for asymmetric variations of independent variables for different collaborators should be investigated for multi-user haptic tasks. With the knowledge of these limits and requirements, multi-user haptic communications schemes can be designed efficiently.

ACKNOWLEDGMENT

This work has been supported, in part, by the German Research Foundation (DFG) under the project STE 1093/4-1 and, in part, by the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement no. 258941.

REFERENCES


