Influence of Visual and Haptic Delays on Stiffness Perception in Augmented Reality

B. Knörlein*
Computer Vision Laboratory
ETH Zurich

M. Di Luca†
Max Planck Institute for
Biological Cybernetics
MPI Tübingen

M. Harders‡
Computer Vision Laboratory
ETH Zurich

Abstract

Visual delays are unavoidable in augmented reality setups and occur in different steps of the rendering pipeline. In the context of haptic interaction with virtual objects, it has been shown that delayed force feedback can alter the perception of object stiffness. We hypothesize that delays in augmented reality systems can have similar consequences. To test this, we carried out a user study to investigate the effect of visual and haptic delays on the perception of stiffness. The experiment has been performed in an optimized visuo-haptic augmented reality setup, which allows to artificially manipulate delays during visual and haptic rendering.

In line with previous results, delays for haptic feedback resulted in decreased perceived stiffness. In contrast, visual delays caused an increase in perceived stiffness. However, the simultaneous occurrence of delays in both sensory channels led to a partial compensation of these effects. This could potentially help to correct stiffness perception of virtual objects in visuo-haptic augmented reality systems.

Keywords: augmented reality, multimodal, user evaluation, delay, haptic

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Haptic I/O

1 Introduction

Our research aims at the application of visuo-haptic augmented reality (AR) in medical education. In this field, the faithful display of visual and haptic properties of deformable objects is a key element. Accurate calibration, system stability, and low latency are indispensable prerequisites in such a setting. Therefore, we have developed a specialized, visuo-haptic AR system (see e.g. [3]). Rendering delays, however, are often unavoidable in a complex AR setup and can occur in different steps in the rendering pipeline.

It has been shown that delays can alter a user's perception of virtual objects, e.g. the perception of stiffness. Recent studies by Pressman et al. [8] and Mochizuki et al. [7] revealed that objects are perceived softer when force feedback is delayed. However, perception of objects experienced in AR is not only due to haptic information. The human perceptual system automatically integrates information available in all sense modalities [14]. For example, it has been shown that stiffness perception is also influenced by visual information [16]. For this reason we hypothesize that delays in any of the sensory channels can potentially lead to an alteration of a user's perception.

*e-mail:knoerlein@vision.ee.ethz.ch
†e-mail:max@tuebingen.mpg.de
‡e-mail:harders@vision.ee.ethz.ch

Simultaneous delays in the visual and haptic channel have been studied focusing for instance on collaborative interaction [4], perceived synchrony [5], or delay detection [15]. Only few investigations on the influence of these phenomena in AR setups were performed. Moreover, AR is generally only verified by subjective user feedback or by measuring user performance during specific tasks [2]. Given the importance of obtaining faithful display of material properties in our application area, we carried out a user study to investigate the effects of multimodal delays on stiffness perception.

In the following we first describe our augmented reality setup and then provide quantitative measurements of the delays occurring in the system. Thereafter, the experiment for investigating the perceptual effects is presented. Finally, we conclude with a discussion of the results and an outlook on future work.

2 System Overview

In earlier work we have developed a specialized visuo-haptic AR environment for medical training [3]. This system has been further optimized to minimize delays. An overview of the system is shown in Figure 1. In the current environment, we use a stereo video see-through setup consisting of a head mounted display (HMD) (trivisio 3scope) and two dragonly2 firewireA cameras (point grey). The HMD provides a 40 degree FOV, a resolution of 800x600, and an update rate of 60 Hz. The cameras are equipped with 4.4 mm lenses, concurring with the resolution and FOV of the HMD. This yields appropriate depth perception without folding the optical path [11] or loss of accuracy due to interpolation. An external infrared tracking system (NDI OPTOTRAK 3020) is used, providing an accuracy of 0.2 mm at an optimal distance of 2.25 m. A custom marker, providing high accuracy and large visibility, is attached to the HMD. The cameras are calibrated and the marker is registered using Hand-Eye calibration. Haptic feedback is rendered on a PHANToM 1.5 haptic device (Sensable). The device is controlled with a Real Time Linux system (RTAI) via custom drivers, providing tight timing for haptic rendering and precise feedback.

To control the devices, four PCs (each dedicated either to tracking, haptics, graphics, and simulation) are connected in a distributed framework using a client-server architecture. Every PC runs a multi-threaded application for data processing and acquisition. Tracking and visual rendering are performed at 30 Hz while haptic feedback and simulation are performed at 1 kHz. The distributed framework is connected via two 1Mbps networks, separating the vision client from the haptic and tracking client. Round-trip time of the networks is measured to be below 0.1 ms.

A hardware-triggered setup with time-stamped data packets is used for synchronization. In order to obtain a uniform time scheme on all modules, the clock difference between the clients and the server is determined during initialization of a connection. For the soft synchronization, data are acquired and stored in a ring-buffer together with the acquisition time. The acquisition time is adjusted according to measured hardware delays (see Section 3). The determined clock-difference is used to adjust the timestamps when data are received on or sent from a client. The hardware trigger signal is emitted by the tracker and used to ensure synchronized data acquisition of all devices. It also controls the data exchange and helps to refine the measured clock difference.

To accelerate the visual rendering, a GPU-based image processing pipeline is set up. Delays occurring from image transfer as well as camera on-chip color conversion are reduced by acquiring and only sending unconverted Bayer images. GPU shaders are later applied for Bayer conversion, image undistortion, and visualization of the augmented scene. The virtual scene is predrawn when tracking and simulation data are available, while the graphics loop waits for camera image data. This avoids additional delays from rendering the virtual scene.

3 PERFORMANCE MEASUREMENT

In order to characterize the delays occurring in the setup described above, we carried out several performance measurements. Some methods for measuring delays have been discussed, e.g. in [10, 12]. Nevertheless, our interest not only lied in the achieved end-to-end latency, but also in the contribution of individual steps in the pipeline. This allows highly accurate synchronization and identifies the major origins of delays.

Delays can occur in three steps of the pipeline: The data acquisition $d_{acquisition}$, the data processing $d_{processing}$ and during the display of the feedback $d_{display}$. The overall delay of a system is therefore given by $d_{overall} = d_{acquisition} + d_{processing} + d_{display}$. For haptic rendering delay mainly occurs during processing, while for acquisition and display it can be assumed to be zero. In the employed setup, processing of the data is performed in a 1 kHz loop and therefore negligible. In the following, we therefore only analyze the visual pipeline and report averages for each of the steps.

In the visual case, the data acquisition time is given by the maximum of the delays required for acquisition of the images d_{camera} and the tracking data $d_{tracker}$. We measured these delays as the time difference between reception of the trigger signal and reception of the data. This revealed an overall delay for the camera d_{camera} of 35.5 ms (including a 5 ms shutter and a firewireA frame transmission time of 28.75 ms). For the tracking system, the measurements have been performed for different frame rates and numbers of infrared LEDs. Analysis of the data revealed a linear relationship between delay and the number of LEDs. A first-order polynomial fit showed a base delay for the tracking device of 4.48 ms with additional 0.46 ms per LED. In the current setup 14 LEDs are used, resulting in a delay $d_{tracker}$ of 10.92 ms. The data acquisition time $d_{acquisition}$ is therefore limited by the camera to 35.5 ms.

For the image processing pipeline the computation times affecting $d_{processing}$ have been determined on a Nvidia GT285 graphics card as: image upload 0.9 ms, bayer conversion 0.8 ms, undistortion 0.4 ms and visualization 0.4 ms, resulting in an overall image processing delay $d_{processing}$ of 2.5 ms.

Finally, we considered the delay due to the HMD display $d_{display}$. It was determined by measuring the achieved end-to-end latency $d_{overall}$ and subtracting the previously obtained delays for processing and acquisition. The former was acquired by using two photodiodes. One attached next to the camera and a second one onto the display of the HMD. The photodiode attached to the camera responded directly to incoming light, while response of the second diode was subject to the systems end-to-end delay. We reciprocated a black and white gradient pattern in front of the camera to modulate the light intensities. This resulted in two correlatable 1D signals, with the shift of the HMD diode corresponding to the end-to-end latency. The delay of the HMD was then determined by cross-correlation of the two signals. In this way we measured the overall end-to-end delay to be 66 ms. By subtracting the previously acquired times for $d_{acquisition}$ and $d_{processing}$, the delay for the HMD $d_{display}$ was estimated to be 28 ms.

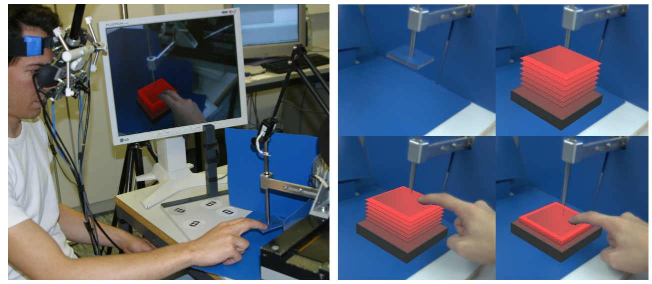

Figure 1: Visuo-haptic augmented reality setup and the compression of the virtual spring

4 USER STUDY

We conjectured that additional delays in the visual or haptic rendering can change stiffness perception. Delays occurring during haptic rendering result in the position, sensed through proprioception, leading the exerted force. In contrast, visual delay results in the position, sensed through vision, trailing the rendered force. Taking the results reported in [7, 8] into account, we hypothesized that haptic delay leads to a decrease of perceived stiffness, and visual delay to an increase. In addition to this, we also investigated the effects of simultaneous presence of delays in both channels, which we expected to lead to a partial cancellation of the changes in stiffness perception.

A two interval forced choice (2IFC) design was employed in the experiment where the main task was the compression and comparison of virtual springs in the visuo-haptic AR environment. Experimental conditions were created by introducing artificial delays in the rendering pipeline.

4.1 Participants

Fourteen participants (7 male, 7 female, 19-38 years) took part in the experiment. None of them reported any haptic deficits. Eleven of the participants had never used a haptic interface before and all were naive about the goal of the study. Participants have been recruited by public announcement. A consent form was signed and a reimbursement of 30 CHF was paid. The study has been approved by the ETH ethics committee.

4.2 Experimental Setup

The study setup consisted of a support base on which a virtual 10 cm spring-like object was displayed (see Figure 1 top). For direct interaction with the spring, a plate was mounted to the end of the haptic device to simulate the top surface of the spring. Motion was restricted to vertical movement by a guide rail. The weight of the construction was compensated by rendering a predetermined offset force. The virtual spring forces were linearly proportional to the distance to the table, representing a simple hookian spring. The HMD was supported by a flexible arm and not worn directly by the participants. This setup permitted small head movements during the experiment, while it increased user comfort. Visual occlusion of virtual objects was obtained through blue color keying. A typical interaction in this setup is depicted in the bottom of Figure 1.

4.3 Experimental Stimuli

In one trial, a pair of two virtual linear elastic springs was presented to the subject in random order; one being the test (standard stimulus) and the other being the reference sample (comparison stimulus). The former had a constant stiffness, but varied across trials with regard to visual and/or haptic delay. The latter only varied with regard to stiffness.

The conditions for the standard stimuli (independent variables) were *no delay*, *force delay*, *visual delay*, and *visual+force delay*.

Note that in all these conditions (including *no delay*) the intrinsic visual end-to-end delay of 66 ms was present. For the visual delay condition, rendering was shifted by three frames, resulting in an additional 100 ms delay. For the force delay condition, haptic rendering of the spring force was delayed by 20 ms while the intrinsic delay of the haptic feedback can be assumed to be zero. The spring stiffnesses of the standard stimuli were set to 0.03 N/mm while the comparison stimuli were taken from the set R={0.023, 0.0245, 0.026, 0.0275, 0.029, 0.031, 0.0325, 0.034, 0.0355, 0.037} N/mm. The range of these values was chosen according to JNDs for stiffness discrimination, as reported in literature [13], while taking the expected strength of the effects into account, as well as the capabilities of the haptic device. Again, note that also for the reference samples the intrinsic visual delay of 66 ms was present.

The values for the haptic and visual delay were selected according to the results of a pilot experiment involving five subjects. We measured detection thresholds for delay using the same experimental setup but employing a three interval forced choice design. Here, three samples were presented consisting of two delayed and one undelayed stimuli or vice versa. The subjects performed the same movements as for the main experiment, but they were asked to identify the stimuli perceived different from the other two. For visual feedback, additional delays of 33, 66, 100, 133 and 166 ms were displayed, while for haptic feedback, delays of 10, 20, 30, 40 and 50 ms were used. The detection of additional delays was identified as the value corresponding to a 66% chance for correct detection. The results showed that additional visual delay was on average detected for a delay higher than 109.4 ms, while haptic delay was noticed after 29.6 ms. Thus, by choosing smaller values, delays during the experiment should barely be noticeable by the participants.

4.4 Experimental Procedure

Pairs of springs, each rendered in a different color, were consecutively presented to the participants. Subjects were told to probe each spring with a steady in- and outward movement by pushing the plate attached to the haptic device to a depth of about 7 cm. After indentation of both springs, subjects were asked to select the one which they perceived as stiffer. The answer (1^{st} or 2^{nd}) was given by pressing one of two correspondingly colored buttons located next to the setup. Thereafter, the next pair was presented. Black frames were rendered after each trial, in order to hide any noticeable changes in the visual delay.

Before the actual experiment a training phase took place. First, each subject was introduced to the procedure, while feedback about the correct answer was given. Thereafter, a preparatory test was performed, only involving the 10 reference samples and the *no delay* sample. Each combination was randomly presented five times, resulting in 50 test-reference pairs. These data were used to examine the accuracy of the subjects' haptic perception and to verify that the task had been fully understood. After this, the actual experiment was performed. Each test-reference pair was randomly presented ten times resulting in 400 comparisons. Breaks of 2-10 min were enforced to reduce stress and fatigue after 100 trials. One phase took approximately 10 min. Finally, after the procedure a questionnaire was filled out by the subjects. Including training and introduction the experiment lasted on average 120 min. The answers as well as the movement during the trials were anonymized and stored for later analysis.

5 Results

The probabilities of perceiving the comparison stimulus stiffer than the standard stimulus were fitted with a cumulative Gaussian function. From these psychometric functions we obtain our dependent variables. We determine the Points of Subjective Equality (PSE) as the stiffness corresponding to a perception probability of 50%. The PSE identifies the unconditioned comparison stiffness which was perceived equal to the stiffness of a conditioned standard stimulus. Furthermore, the Just Noticeable Differences (JND) were determined based on the differences in stiffness between the probabilities of 50% and 84%. The JND describes the minimal difference in stiffness between reference and test sample, for which accurate classification as being stiffer, or softer, could be made.

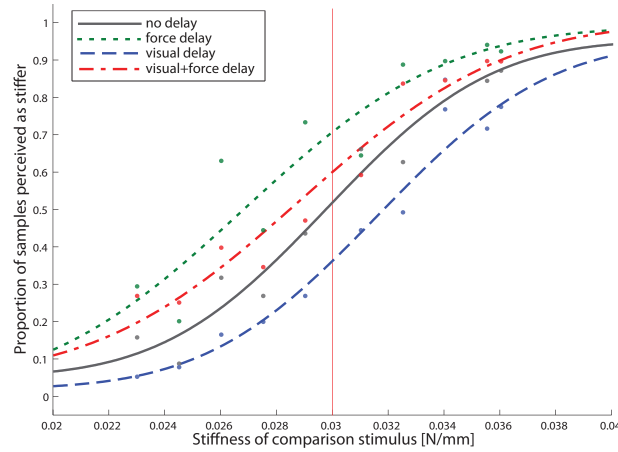

Figure 2: Experimental data averaged across participants. Curves are cumulative normal distribution fitted to the responses. The vertical line represents the simulated value of stiffness for the standard stimulus.

The data of each participant from the training phase and the experiment were analyzed separately. Two subjects had to be excluded from the analysis due to their JND percentage (Weber fraction) being above 50% in the training phase or the experiment (since they were not able to accurately discriminate the range of comparison stimuli). In addition, one participant decided to stop the experiment after 320 of the 400 trials due to the sensation of nausea. After careful examination of the data, we decided to still include this reduced dataset.

Figure 2 shows the psychometric functions acquired from the responses of all subjects. It can be seen that the curve is shifted in the visual and force delay conditions and thus a change in perception of stiffness occurred. In the visual delay condition, the curve shifted to the right, indicating increased perceived stiffness for the standard stimulus, while in the force delay condition the curve shifted to the left, indicating reduced perceived stiffness. Moreover, the simultaneous presence of visual and force delay appears to partly cancel out the effects.

Figure 3 shows the average PSEs and JNDs for the four conditions and the standard errors across participants. Analysis of the PSEs demonstrated that the changes were significant under visual (2-way repeated-measures ANOVA: $F(1,11)=66.775$, $p=0.000005$) and under force delay ($F(1,11)=35.752$, $p=0.00009$). Interaction was not significant ($F(1,11)=0.196$, $p=0.666$). In line with our hypothesis, stiffness was in general perceived to be higher under the influence of visual delay (one-sample t-test: $t(11)=7.677$, $p=0.00001$), and lower for force delay ($t(11)=6.349$, $p=0.00006$). For simultaneous visual and force delay, the perceived stiffness was in between that of the separate delays (paired-sample t-test, Bonferroni corrected, visual: $t(11)=5.266$, $p=0.0005$; haptic: $t(11)=4.966$, $p=0.0009$).

The effect of altered perception occurred for every subject, however, the strength of the effects changed slightly among them. This can be attributed to subjects integrating proprioceptive and visual information about their hand position with different weights, as described for instance in [6].

Further, the resulting JNDs show that delay in general worsens

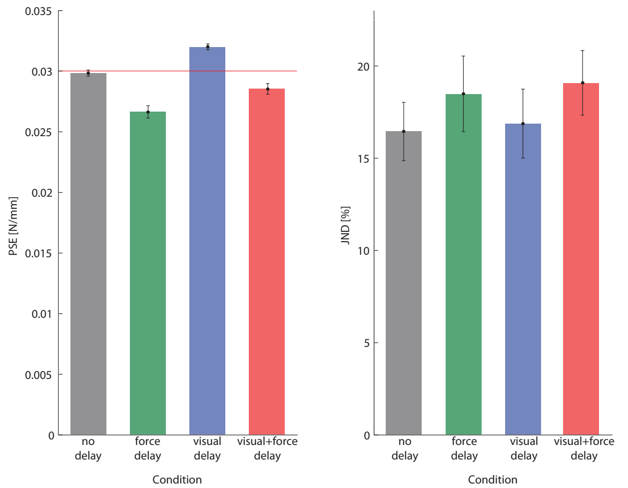

Figure 3: Perceived stiffness (simulated value indicated as red line) and discrimination thresholds under influence of delays (+/- 1 s.e.m.)

the ability to discriminate stiffness. For force delay a significant change could be detected (2-way r.m. ANOVA: F(1,11)=5.411, p=0.04) while for visual delay the data were less conclusive (F(1,11)=0.233, p=0.638).

Finally, in the questionnaires only one subject mentioned the potential presence of any visual delays. This supports the detection thresholds obtained in our pilot experiment.

6 Discussion

In our performance measurements we have shown that delays in AR are often unavoidable and occur mainly due to data acquisition and display. However, visual rendering can be optimized for processing times close to zero. Further reduction of the visual delay would only be possible by using improved hardware components, e.g. firewireB. An alternative AR setup could be optical see-through [9], however, in this case the visual rendering of the virtual objects would lag behind the direct view of the real environment. It is most likely that users will notice this kind of visual asynchrony. Any effects on perception in such a setting still need to be investigated. In this light, having a slightly, but consistently, delayed visual augmentation might be preferable.

Our study demonstrated the influence of delays on a user's stiffness perception, supporting the necessity of our developed low latency framework, especially in the area of medical training. In line with previous results [7, 8], delays in force feedback resulted in a decrease of perceived stiffness. Since stiffness can be expressed as force over position, it appears that delaying the force decreases the perceived stiffness. On the other hand, delaying the position should increase the perceived stiffness. Since it is impossible to delay the proprioceptively-sensed position, we showed that delaying the visually-sensed position creates an increase in stiffness (in [1] we provide a full explanation for the perceptual mechanism underlying this phenomenon).

The changes in JND indicated that the discrimination of stimulus stiffness under force feedback delay degrades. For visual delay the change was not significant. However, it has to be kept in mind that comparisons were only performed relative to the intrinsic visual delay of 66 ms. Therefore, a degradation could still take place compared to a true zero delay situation. In addition, other factors potentially affecting perception in AR also need to be considered in system design, e.g. misalignments, reduced depth perception, or synchronization errors.

In addition our results showed,that the effects of visual and haptic delay were additive and counteract each other under simultaneous occurrence. This could be used to correct haptic feedback under visual delays. However it has to be taken into account that the perceived force over position does not follow a linear relation, which results in an increase in the discrimination threshold, as can be seen in figure 3 (right). Further studies are therefore required to assess an acccurate correction of the feedback.

7 Conclusion

We have carried out a user study in which the effect of delays in visual and haptic rendering on the perception of stiffness was shown. In addition, the end-to-end delay already present in our optimized augmented reality setup has been measured. Additional studies will be carried out in the future to give further insight into haptic perception and multimodal integration in a visuo-haptic augmented reality environment. This will include the adjustment of force feedback in order to balance the effects of visual delay.

Acknowledgements

The authors wish to thank Habib C. Bou Sleiman for his help with the experiments. This work has been performed within the frame of NCCR Co-Me supported by the SNSF and of the EU project Immersence IST-2006-27141.

References

[1] M. Di Luca, B. Knörlein, M. Ernst, and M. Harders. Effects of visual-haptic asynchronies and loading-unloading movements on compliance perception. *(submitted)*, 2009.

[2] A. Dünser, R. Grasset, and M. Billinghurst. A survey of evaluation techniques used in augmented reality studies. In *SIGGRAPH Asia courses*, pages 1–27. ACM, 2008.

[3] M. Harders, G. Bianchi, B. Knörlein, and G. Székely. Calibration, registration, and synchronization for high precision augmented reality haptics. *TVCG*, 15(1):138–149, 2009.

[4] C. Jay, M. Glencross, and R. Hubbold. Modeling the effects of delayed haptic and visual feedback in a collaborative virtual environment. *ACM Trans. Comput.-Hum. Interact.*, 14(2):8, 2007.

[5] C. Jay and R. Hubbold. Delayed visual and haptic feedback in a reciprocal tapping task. In *Worldhaptics*, pages 655–656, 2005.

[6] A. Lecuyer, J.-M. Burkhardt, S. Coquillart, and P. Coiffet. "boundary of illusion": an experiment of sensory integration with a pseudo-haptic system. In *Virtual Reality*, pages 115–122, 2001.

[7] H. Ohnishi and K. Mochizuki. Effect of Delay of Feedback Force on Perception of Elastic Force: A Psychophysical Approach. *IEICE Trans Commun*, E90-B(1):12–20, 2007.

[8] A. Pressman, L. Welty, A. Karniel, and F. Mussa-Ivaldi. Perception of delayed stiffness. *Int. J. Rob. Res.*, 26(11-12):1191–1203, 2007.

[9] J. Rolland and H. Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoper. Virtual Environ.*, 9(3):287–309, 2000.

[10] T. Sielhorst, W. Sa, A. Khamene, F. Sauer, and N. Navab. Measurement of absolute latency for video see through augmented reality. In *ISMAR*, pages 215–220, 2007.

[11] A. State, K. Keller, and H. Fuchs. Simulation-based design and rapid prototyping of a parallax-free, orthoscopic video see-through head-mounted display. In *ISMAR*, pages 28–31, 2005.

[12] C. Swindells, J. Dill, and K. Booth. System lag tests for augmented and virtual environments. *UIST'00*, pages 161–170, 2000.

[13] H. Tan, N. Durlach, G. Beauregard, and M. Srinivasan. Manual discrimination of compliance using active pinch grasp: the roles of force and work cues. *Perception & psychophysics*, 57(4):495–510, 1995.

[14] V. Varadharajan, R. Klatzky, B. Unger, R. Swendsen, and R. Hollis. Haptic rendering and psychophysical evaluation of a virtual three-dimensional helical spring. *Haptic Symposium*, pages 57–64, 2008.

[15] I. M. Vogels. Detection of temporal delays in visual-haptic interfaces. *Hum Factors*, 46(1):118–134, 2004.

[16] W.-C. Wu, C. Basdogan, and M. A. Srinivasan. Visual, haptic and bimodal perception of size and stiffness in virtual environments. *ASME/IMECE*, pages 19–26, 1999.