User-Based Evaluation of Data-Driven Haptic Rendering

RAPHAEL HÖVER, ETHZ
MASSIMILIANO DI LUCA, Max Planck Institute for Biological Cybernetics
MATTHIAS HARDERS, ETHZ

In this article, the data-driven haptic rendering approach presented in our earlier work is assessed. The approach relies on recordings from real objects from which a data-driven model is derived that captures the haptic properties of the object. We conducted two studies. In the first study, the Just Noticeable Difference (JND) for small forces, as encountered in our set-up, was determined. JNDs were obtained both for active and passive user interaction. A conservative threshold curve was derived that was then used to guide the model generation in the second study. The second study examined the achievable rendering fidelity for two objects with different stiffnesses. Subjects directly compared data-driven virtual feedback with the real objects. Results indicated that it is crucial to include dynamic material effects to achieve haptic feedback that cannot be distinguished from real objects. Results also showed that the fidelity is considerably decreased for stiffer objects due to limits of the display hardware.

Categories and Subject Descriptors: H.5.2 [**Information Interfaces and Presentation**]: User Interfaces—*Haptic I/O*

General Terms: Algorithms, Human Factors, Measurement, Verification

Additional Key Words and Phrases: Discrimination study, force discrimination, haptic rendering, deformable models

ACM Reference Format:

Höver R., Di Luca, M., and Harders, M. 2010. User-based evaluation of data-driven haptic rendering. ACM Trans. Appl. Percept. 8, 1, Article 7 (October 2010), 23 pages.
DOI = 10.1145/1857893.1857900 http://doi.acm.org/10.1145/1857893.1857900

1. INTRODUCTION

Data-driven haptic rendering is an emerging technique for generating haptic feedback of high fidelity. In contrast to well-established rendering methods that employ physical or heuristic models, this approach completely relies on data recorded from interactions with real objects. Based on these data, a data-driven model is generated, capturing the haptic properties of the objects. No explicit knowledge about or setting of material properties is necessary.

In our approach, we considered continuous recordings of a hand-actuated tool indenting visco-elastic objects along one degree of freedom. To this end, we extended a PHANToM 1.5 (SensAble) with a force sensor and used it as a recording as well as rendering device. During the interaction with the real object, the tool movement and contact forces were captured. These data were used to derive a data-driven model of the recorded object.

This work was supported by the EU project Immersence IST-2006-27141.
Author's address: R. Höver, Computer Vision Laboratory, ETH Zurich, Virtual Reality in Medicine Group, Sternwartstr. 7, 8092 Zurich, Switzerland; email: hoever@vision.ee.ethz.ch.

DOI 10.1145/1857893.1857900 http://doi.acm.org/10.1145/1857893.1857900

The central goal of this work was the assessment of the rendering fidelity of our data-driven model. Two experiments were carried out to this end.

In the first study, the Just Noticeable Difference (JND) for force levels encountered in our set-up was determined. It should be noted that some JND values have already been reported in the literature; however, these were usually only determined for rather large force levels and/or for different interaction paradigms. Hence, we conducted an experiment to estimate the complete threshold curve, covering the full range of force magnitudes in our system.

JNDs were obtained both for active and passive user interaction. Obtaining the JND curves allows us to determine a theoretical threshold below which force errors should not be perceivable anymore by a user. The purpose of this experiment was twofold. The threshold was firstly used for controlling an optimization of the interpolation used for data-driven haptic rendering. This technique is described in more detail in Hoever et al. [2009a] and will briefly be summarized in Section 3. Secondly, the threshold allows the theoretical evaluation of the rendering algorithm by examining force errors for prerecorded interaction trajectories.

In the second experiment, the achievable rendering fidelity was examined in a discrimination study where users directly compared real and virtual objects. Subjects penetrated the visco-elastic test objects with movements of their choice along 1 degree of freedom. In contrast to the first study, this was a holistic assessment considering all system components—the rendering algorithm, the display hardware, as well as user performance. In the study, the importance of incorporating dynamic material effects into the rendering became apparent. Results indicated that it is crucial to include such effects to achieve haptic feedback that cannot be distinguished from real objects. Results also showed that the fidelity is considerably decreased for stiffer objects due to the limited capabilities of the display hardware.

In the following, we first give an overview of related work in Section 2. Then in Section 3, the algorithm for the generation of data-driven haptic feedback is outlined. Section 4 presents the JND study from which the perceptual threshold is derived that guides the model generation in the discrimination study in Section 5. In this second study, the fidelity of our data-driven approach is scrutinized. Section 6 discusses the results of both studies, and Section 7 concludes the article.

2. RELATED WORK

2.1 Data-Driven Haptic Rendering

The concept of using recordings from real objects for generating haptic feedback has been addressed in a few projects. However, most of these are model-based and only use the recorded data to determine model parameters for the underlying simulation.

For example, in MacLean [1996], a piecewise linear model is fit to recorded force data. Material parameters like stiffness, damping, and inertia are derived. In Okamura et al. [2003], haptic feedback during virtual surgical cutting is rendered using a piecewise linear model that is fit to recordings obtained from real scissors and tissue. In Colton and Hollerbach [2007b, 2007a], haptic feedback of push buttons was considered. A nonlinear model was used to approximate captured force and acceleration data. Other work like Andrews and Lang [2007] and Richard et al. [1999] focused on surface properties like texture and friction.

A different approach is undertaken when no model is used to reproduce the interaction. Captured data are only stored and replayed on demand. This approach is followed by Kuchenbecker et al. [2005, 2006], where an event-based algorithm is used to render haptic feedback of tapping on rigid objects. Transient acceleration signals of collisions with real objects are recorded and stored in a library. During rendering, these transients are then overlaid to enhance the realism of the feedback. In Hoever et al.

[2008, 2009], we presented a data-driven haptic rendering approach for visco-elastic objects and viscous fluids. We do not estimate any underlying parameters. Instead, a generic interpolation technique was used to learn the material properties and provide virtual feedback. The fidelity of this approach is improved in this article by incorporating human perceptual capabilities to tune the data-driven model. This technique is then verified in a separate user study by showing that the obtained model is perceptually indistinguishable from real objects.

2.2 Human Force Sensitivity

Various approaches are available in the literature that quantify the human capability to discriminate forces. A widely used concept in this context is the JND, the smallest increment or decrement of force ΔF that can still be perceived for a specific stimulus. It has been shown that this quantity depends on the stimulus intensity F, according to the so-called Weber law [Weber 1978]. This law states that the Weber fraction $W = \Delta F/F$ is constant for each stimulus type across a wide range of intensities (see Luce et al. [1963] and Gescheider [1985]). The original value of the Weber fraction was measured in experiments involving active lifting of 32oz weights (corresponding to 8.9N of force), resulting in roughly 10%. Although this kind of task involves more sensory information than simple force, similar values of sensitivity have been measured in different studies related to force. In a force-matching task for the elbow, the Weber fraction is found to be between 5% and 9% [Jones 1989]. Similarly, for pinching between index finger and thumb with a constant forces between 2.5N and 10N, the Weber range is between 5% and 10% [Pang et al. 1991]. Also, when subjects moved their index fingers against a resistive force of 2.25N, a Weber fraction of about 10% was found [Allin et al. 2002].

The results of such studies have for instance been exploited in data compression of haptic signals [Hinterseer et al. 2005; Hinterseer and Steinbach 2006]. Hirche et al. [2007] limited the network traffic by transmitting only changes in the force signal that exceeded the perceptual threshold according to the Weber fraction. We also employed the discrimination threshold for force in our earlier work to assess offline different rendering strategies [Hoever et al. 2008, 2009]. However, the forces that were displayed in our set-up were rather small, while the previously mentioned works on JNDs employ forces larger than 2N. It is generally expected that the Weber fraction increases significantly for smaller stimulus intensities [Luce et al. 1963; Gescheider 1985; Pang et al. 1991; Norwich and Wong 1997].

Research exists that looked into this matter from the perspective of weight discrimination. Studies examined discrimination performance in the range below 50g. This corresponds to 0.5N if the weight is unaccelerated. It was found that below this value, the resulting Weber fraction increased as the weight decreased [Engen 1971; Ross and Brodie 1987]. However, the property being judged during weight perception is different to force discrimination. Hence, the sensitivity can significantly deviate [Kawai 2003; Jones and Hunter 1990; Ross and Brodie 1987]. This factor needs to be considered when applying such thresholds in the proposed modeling technique.

From the previously summarized literature, it is evident that there is no single study that investigates sensitivity to force difference over a wide range of force magnitudes, especially for forces close to detection threshold. It is also not clear whether the particular way of grasping the stylus typically employed in haptic interfaces would lead to different thresholds than the ones obtained in the studies cited. For these reasons, we conducted an experiment that covers a range of forces that extends to very low magnitudes, thus complementing the results of other JND studies. Moreover, we employed an experimental setting that is similar to the typical interaction with force feedback devices, where the user is holding the stylus between thumb, index, and middle fingers similar to a writing grip. In this way, the thresholds obtained with this experiment characterize human force sensitivity over a large range of forces while holding a stylus. Furthermore, the experimental conditions are very similar to the ones intended for typical utilization of the rendering algorithm, so they can be employed in its tuning.

2.3 Assessment of Haptic Feedback Systems

In order to measure the accuracy of a haptic feedback system, it is necessary to incorporate all different aspects such as the rendering algorithm, the limitations of the display hardware, and the perceptual capabilities of the user.

An offline technique for testing haptic rendering algorithms was presented in Ruffaldi et al. [2006], where recorded force signals along trajectories were compared with the virtual feedback using the same trajectory. This method leads to error measures that directly rate the accuracy of the rendering approach without taking the display hardware or the human perception into account.

In Okamura et al. [2003], a method to render surgical cutting was proposed. The system was evaluated by subjects comparing the haptic recordings from real cutting tasks against an empirical model. This comparison takes perceptual capabilities into account. However, there is no comparison of the model with cutting real tissue.

Such direct comparisons were performed in other studies. For example, in Kuchenbecker et al. [2005], an event-based rendering approach was assessed by a user study where subjects tapped on real and virtual samples. The similarity ratings that the users gave revealed a high realism of the event-based rendering technique.

Similarly, Leskovsky et al. [2006] evaluated the fidelity of haptic rendering using mass-spring models. Multidimensional scaling was applied to identify different perceptual cues used to discriminate real and virtual objects. The results showed that users were able to perceive differences between real and virtual objects of similar stiffness, when a rendering algorithm of reduced fidelity was applied. This effect could be reduced if a more sophisticated rendering algorithm was used.

In Swindells et al. [2009], haptic feedback of different turning knobs was modeled using recorded data and a predefined model structure. Automatically estimated model parameters were compared to manually tuned parameters from humans who matched the rendering to the feedback of the real knobs. Based on the similarity of both parameter settings, it is inferred that the automatic procedure is able to capture knob dynamics as perceived by a human.

In our earlier work [Hoever et al. 2008, 2009], we combined the concept from Ruffaldi et al. [2006] with perceptual thresholds. The approach used human sensory limitations to derive a first quality measure of our proposed rendering algorithm. We used test-trajectories for which real contact forces were recorded. The relative force error of the rendered forces was compared with the corresponding Weber fraction reported in Pang et al. [1991]. Besides the limitation of this approach to rather large force magnitudes, as mentioned in Section 2.2, the test set-up and the user task in Pang et al. [1991] was different to ours. Hence, the applied threshold only provided a rough estimate of user performance for our set-up. In Hoever et al. [2009a], we extended results from the literature using a perceptual threshold over the whole range of force magnitudes. In addition, a strategy was presented to incorporate the threshold into the computation of the data-driven rendering.

Nevertheless, the JNDs that were used for this offline evaluation only captured the perception thresholds for constant force signals, with no changes over time. Active user interaction was also not taken into account.

Therefore, we extended the assessment of our data-driven rendering approach in this article with a user study. Subjects were asked to discriminate between real and virtual objects, in order to perform a comprehensive test of the devised method.

3. RENDERING ALGORITHM

In this section, a concise summary of our data-driven haptic rendering approach is given. For more details refer to Hoever et al. [2009b, 2009b].

In order to generate the data-driven haptic feedback, recordings are obtained from real objects with a sensorized probing tool attached to a PHANToM haptic device. Using position encoders and an additional acceleration sensor, position, velocity, and acceleration of the probing tool are estimated by a Kalman filter. Also, the contact forces at the tip of the tool are recorded by a force sensor.

These raw data are then interpolated by a generic interpolation technique. To this end, the independent variables of the interpolation function need to be defined beforehand. For instance, to interpolate elastic properties of a material, recorded forces need to be interpolated over the tool position. If also dynamic effects need to be considered, additional input dimensions—tool velocity or acceleration—need to be incorporated. For visco-elastic materials considered in this article, tool position, velocity, as well as a low-pass–filtered version of tool velocity are used. The latter is important for transient material effects like force relaxation. We use a first order low-pass filter given by

$$\mathbf{v}_{out}[nT] = a \cdot \mathbf{v}_{out}[(n-1)T] + (1-a) \cdot \mathbf{v}_{in}[nT] \tag{1}$$

$$\text{with} \quad a = \exp\left[-\frac{T}{\tau}\right], \tag{2}$$

where $\mathbf{v}_{in}$ is the measured and $\mathbf{v}_{out}$ the filtered tool velocity, T is the sample period and τ is the time constant of the filter.

For one-dimensional tool interaction, the input space of the interpolator can be represented by a three-dimensional vector $\mathbf{x}$, comprising of the position and the two velocity estimates. For the generic interpolation, a sum of radial basis functions (RBF) is fit to the data. This sum can be expressed as

$$h(\mathbf{x}) = \sum_{j=1}^{M} w_j \phi(\parallel \mathbf{x} - \mathbf{c}_j \parallel) + \sum_{k=1}^{L} d_k g_k(\mathbf{x}) \qquad \mathbf{x} \in \mathbb{R}^n, \tag{3}$$

with radial basis functions ϕ, RBF centers $\mathbf{c}_j$, and linear weights w_j. The functions g_k form a basis of an additional polynomial term with coefficients d_k.

Usually, for each recorded data sample, one RBF center $\mathbf{c}_j$, located at the corresponding interaction vector in the interpolation domain is used. However, in Hoever et al. [2009a], we presented an approximation technique based on a kD-tree subdivision where only a small subset of centers was required to obtain interpolants with high accuracy, thus allowing to speed up the computation considerably.

Data reduction in the kD-tree is carried out by first forming approximation functions using the barycenters in the tree leaves. The approximation error at each recorded data site is then evaluated. The leaf that contains the site at which the error deviates the most from the perception threshold derived in Section 4 is then further subdivided into two new leaves. This process is repeated until either all approximation errors at the recorded interaction vectors are below the threshold curve or until a maximum number of RBF centers is reached. If the kD-tree was fully expanded, each leaf contained only one data point and the normal interpolation problem is restored.

We set the maximum number of RBF centers to 50 to limit the complexity of the generated models. If this maximum is reached and not all errors are below the threshold, then the approximation with the smallest maximum deviation from the threshold curve is used for the rendering. In addition to the algorithm as it was presented in Hoever et al. [2009a], we only considered recorded data samples for the subdivision where the force magnitude was larger than 0.1N. This avoids overfitting in regions with a low signal-to-noise ratio. However, these samples were only neglected during the search for the next subdivison. For the determination of the approximation function, these samples were included in the regression.

Fig. 1. General set-up of the JND study. The user grasped the stylus of the PHANToM with his right hand while his left hand was on a keyboard to select his responses.

4. FIRST STUDY: JND FOR SMALL FORCES

The purpose of this study was to determine a perceptual threshold based on JNDs that could be used for the refinement of the data-driven model, as described in Section 3. Our study particularly addressed small force magnitudes, as the JNDs for such low stimulus intensities have not yet been investigated. Details on the haptic device and the arm configuration are given in Section 5. The study followed a two-alternative forced choice (2AFC) design. Force pairs were rendered to subjects; for each pair, a subject had to decide which of the two forces was the larger one.

The study consisted of two independent subparts. In the first part, subjects remained a constant arm pose and forces were rendered as temporal cues. In the second part, subjects had to actively move and enter a force field in order to receive feedback. Note, that both parts were independent from each other, and each subject either performed the first or the second. In the following, we term these parts *passive condition* and *active condition*.

4.1 Participants

Eight male and four female right-handed college students participated in the first part of the study. A different group of again eight male and four female students took part in the second. All subjects were naïve with respect to haptic displays and did not have any haptic deficit due to accidents or illness. Subjects were compensated with 20 Euros for their participation. Ethical approval for the study was obtained from ETH ethics board.

4.2 Set-up

The PHANToM Premium 1.5 (SensAble) was used as the haptic display. The choice of the device was based on availability in the lab. In order to guarantee a high accuracy of the rendered forces, the motor gains of the PHANToM were calibrated using a force sensor (ATI Nano17). The device was put on a box, as shown in Figure 1. For the desired arm configurations, the haptic interaction point was located closely to the center of the device workspace. In this area, the spatial resolution of the device is 0.03mm and force artifacts due to the spindle and wire mechanism of the PHANToM are minimized.

The set-up was placed behind a curtain. Users received instructions given on a computer screen visible through a hole in the curtain. In addition, users were asked to wear headphones to avoid acoustic cues from the PHANToM motors.

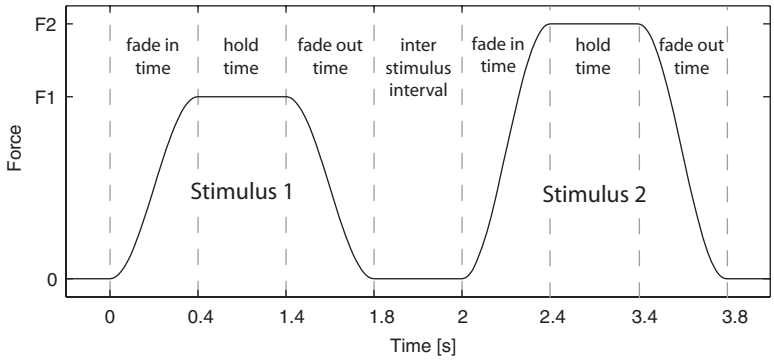

Fig. 2. Temporal profile of a force pair for the passive condition in the JND study. The two forces are faded in and out using a cosine-ramping function.

4.3 Stimuli

In both the active and passive condition, we examined four different force levels, namely 0.3N, 0.5N, 1.0N, and 2.5N, in the following referred to as "standards." In the desired configuration of the participants' arms, forces were roughly directed toward a user, oriented along the stylus of the PHANToM. Standards larger than 2.5N were not considered since the corresponding Weber fraction has already been investigated in similar experiments [Jones 1989; Pang et al. 1991], consistently being reported as about 10%.

For each of the four standards, eight comparison forces were chosen in order to set up pairs of forces. The comparison forces deviated from the standards by $\pm 5\%$, $\pm 10\%$, $\pm 20\%$, and $\pm 40\%$. For instance, for the standard 1N, the comparison forces 0.6N, 0.8N, 0.9N, 0.95N, 1.05N, 1.1N, 1.2N, and 1.4N were used. Each combination of the four standards and their eight comparisons was repeated 10 times during the study, thus resulting in 320 force pairs for each subject. The order of standard and comparison force within one trial was set randomly. The method of presenting the forces differed in the passive and the active part.

Passive Condition. Forces were rendered consecutively over time to subjects passively gripping the stylus. To avoid high-frequency components in the force signal, each force was faded in and out using a cosine-ramping function. The temporal profile of one force pair is depicted in Figure 2. The fading time was set to 0.4s, the hold time to 1s and the time between two forces in one pair to 0.2s. In preliminary tests, we also tried out different fading times and conditions. However, this did not show any significant difference in the results.

Active Condition. Forces were rendered when a subject penetrated a force field. A virtual halfspace was defined by a tilted plane located closely to the workspace center. While the stylus was inside the force field, a constant force was rendered pushing the user toward the plane. Note that force magnitude was independent of penetration depth. Moreover, the force field was tilted such that with the desired arm movement, the plane was contacted at roughly normal orientation. To avoid high-frequency components in the force signal, the force was again smoothly faded in at the boundary of the halfspace. The spatial profile of this fading zone was also a cosine-ramping function as in the passive condition. The maximum slope within the fading zone was fixed. Hence, the width of the zone depended on the magnitude of the reference force. In preliminary tests, we examined different slopes for the fading. In these tests, some subjects performed slightly better for smaller, others for larger slopes. A fading

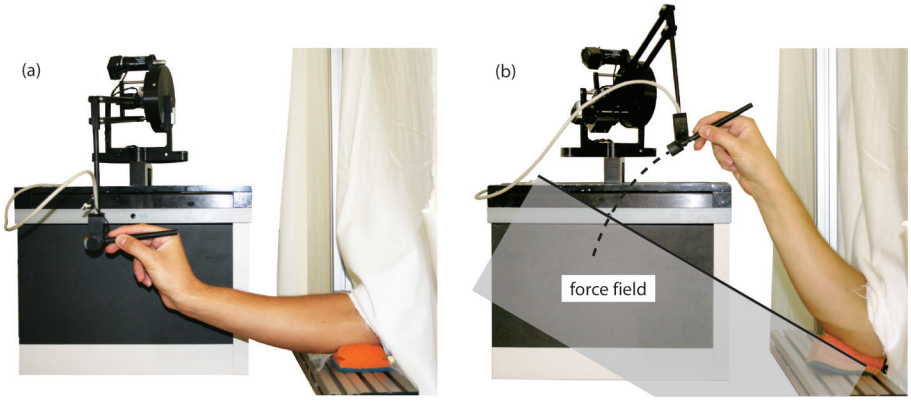

Fig. 3. (a) Arm configuration for the passive condition in the JND study. The subject maintained a fixed position while the forces where faded in and out over time pointing toward the user. (b) Arm configuration for the active condition in the JND study. The subject moved the stylus using his elbow joint and penetrated a force field. The virtual wall was tilted so that the force directly opposed the user movement.

width of 1.5cm for a standard of 1N resulted in the smallest JNDs accross subjects. For this reason, this setting was used in the study.

4.4 Procedure

Subjects were asked to put their right arm through a small opening in the curtain and grasp the PHANToM stylus using a specific grip. Responses could be given with the left hand by pushing either 1 or 2 on a keyboard.

In the passive condition, users were instructed to hold the stylus at a fixed position with an arm configuration as depicted in Figure 3(a). The elbow joint was supported by a soft sponge to increase user comfort. After both forces had been rendered and subjects had given their responses, a random break of 1 to 2s was inserted before the next force pair was presented.

In the active condition, participants had to move the PHANToM stylus and penetrate the force field, closely following a predefined trajectory. The corresponding arm configuration is depicted in Figure 3(b). Subjects only used their elbow joint to move the stylus. The position of the elbow was also indicated by a support plate covered with a soft sponge. At the beginning of each trial, the stylus was positioned outside of the force field. Then, the stylus was moved into the field and the first force of the current pair displayed. Moving back and entering the field a second time caused the display of the second force. In the end, the subjects again moved out and gave their response.

4.5 Training Phase

Before the study started, a preparation phase took place. For the passive condition, a comfortable arm pose was found. Participants were asked to maintain this pose throughout the complete study. Moreover, they were instructed to use a grasp force, which was just high enough to keep the PHANToM from slipping between their fingers. This was particularly important for the small force standards, where large grasping forces could potentially cause perceptual masking effects.

For the active condition, it was ensured that all subjects moved approximately with the same trajectory speed. Moreover, subjects were instructed to limit the dwell time within the force field to about 1s. This movement was learned by the subjects before the study started.

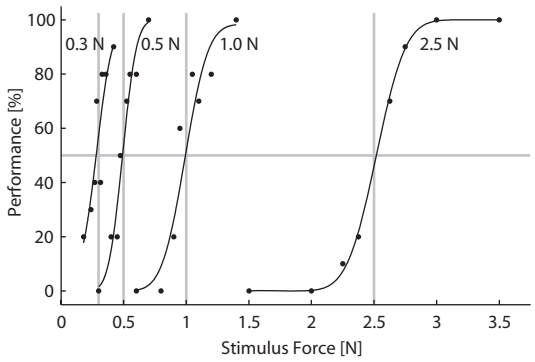

Fig. 4. Psychometric functions from subject 3. Next to each curve, the corresponding standard is printed. The JND in Newton is given by the force value at 84% performance minus the force value at 50% performance. The vertical gray lines mark the standards. The horizontal gray line marks the 50% level.

In a first training session, 12 predefined force pairs were rendered to all subjects. After each response, direct feedback was given whether their answer was correct or not. In a second training session, subjects ran through five random blocks without any feedback under actual study conditions. Subsequently, subjects had a 10m, break before the actual experiment commenced.

4.6 Experiment

The 320 stimuli were divided into 20 blocks each containing 16 pairs of one specific standard. Thus, within a single block the force magnitudes of different pairs were within a similar range (between 60% and 140% of the corresponding standard). The order of pairs within a block and the order of blocks was randomized. In addition, before each block started, the subject was given a sample pair that consisted of the minimum force (60% of the standard) and the maximum force (140% of the standard) of the following block. The sample could be repeated as often as necessary to adjust the grasp force on the stylus and get accustomed to the new force level. No answers were given for these calibration pairs, and thus they were not considered in the results. No repetitions of force pairs were allowed, except from rare cases where subjects were clearly distracted during a trial. This was for instance the case when the subject still talked to the experimenter when a trial had started.

After every five blocks, a break of at least 10min took place in order to avoid fatigue. At the end of the study, participants filled out a questionnaire, providing information on their strategies used for the study.

On average, the participants needed 80min in case of the first and 70min in case of the second experiment to finish the study including three 10min breaks after every five blocks.

4.7 Results

Passive Condition. A cumulative Gaussian was fitted to the experimental data obtaining psychometric functions for each subject and reference force. Figure 4 shows the resulting psychometric functions for an average example subject. The JNDs for the different standards were derived from these functions. The resulting Weber fractions for all 12 subjects are plotted in Figure 5 over the different standards. The diagram shows that for the smallest standard, the JND from subjects 8 and 11 deviated considerably from the rest of the group. In those cases, the fit of the cumulative Gaussian was very poor. For subject 8, we can only assume that this force level was already too close to the detection threshold to perform reasonable discriminations.

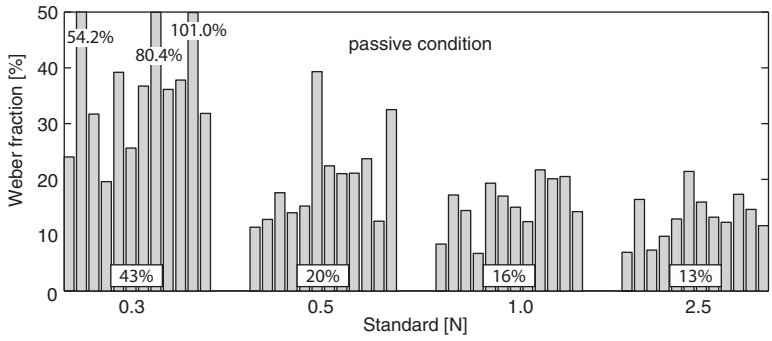

Fig. 5. Weber fractions for all subjects and the different standards in the passive condition. For the 2nd, 8th, and 11th subject, the bars for the smallest standard (0.3N) are clipped and the value is written onto the bar. For each standard, the average Weber fraction is given in the box.

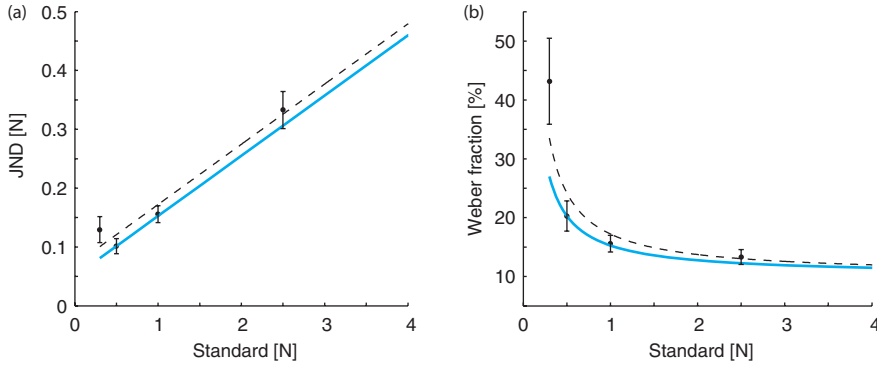

Fig. 6. (a) Mean JNDs over the reference force in the passive condition. The black dots are the average values from the experiment, and the error bars indicate the standard error of the mean. The dashed line is the linear fit to the data. The cyan line is the shifted linear fit that provides a conservative estimate of the threshold. (b) Mean Weber fractions over the reference force. The plotted curves are obtained by dividing the linear functions in (a) through the reference force.

For Subject 11, a training effect took place at this force level. For the first half of the trials, the subject's responses were almost at chance level, thus he was not able to distinguish between such small forces. Nevertheless, in the second half, the subject clearly improved his discrimination capability. This was also confirmed by comments in the questionnaire. In general, it should be noted that values as high as 100% or even larger are reasonable if the intensity of the stimulus is in the order of the detection threshold.

The questionnaire did not reveal any untypical strategies. Three subjects concentrated on the skin stretch between the stylus and their fingers and used this to discriminate small forces.

In Figure 6, the obtained JND and Weber fraction values are shown. The data in Figure 6(a) can be approximated by a linear function (dashed line). However, recall that the intention to use this threshold is the refinement of the data-driven models discussed in Section 5. Since they should be used to express the capability of a user to perceive force errors, we decided to be even more conservative when estimating the perceptual threshold.

Thus, we shifted the fitted line so that all averaged JNDs lie above the linear function. This shifted function is depicted as a solid cyan line in Figure 6(a). It has the intercept $JND(F = 0) = 0.05$N and

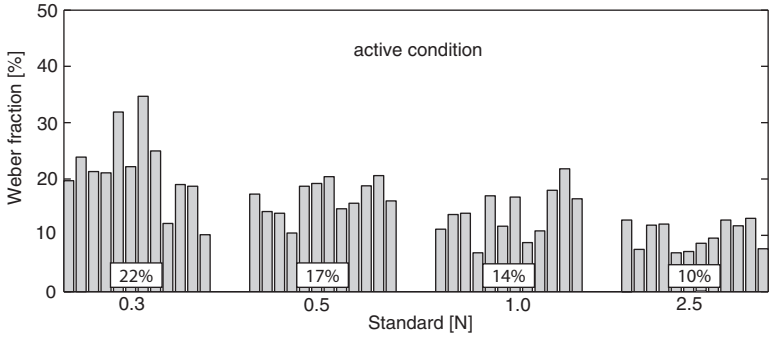

Fig. 7. Weber fractions for all subjects and the different standards in the active condition. For each reference force, the average Weber fraction is given in the box.

the slope 10.23%. This value of the Weber fraction is in line with other results in the literature for the discrimination of forces of 2.5N and above [Jones 1989; Pang et al. 1991]. However, as expected, sensitivity greatly decreased at lower forces reaching an average value of the Weber fraction of 43% at 0.3N. This value is much higher than the one obtained in this range of forces when addressing weight discrimination. In that case, the JND could be as low as 12% with weights of 50g roughly equivalent to 0.5N [Ross and Brodie 1987]. The difference is likely due to the more informative sensory signals available while actively lifting weights. Also, the different experimental conditions, like the configuration of the hand holding the stylus could have caused different results.

Active Condition. The JNDs are obtained from the experimental data in the same way as in the passive condition.

The resulting Weber fractions for all 12 subjects over the different standards are depicted in Figure 7. All average Weber fractions were lower than for the passive user condition. Hence, the discrimination task was, on average, easier in case of the active movement, into the force field. Also, the distribution of the Weber fractions for the smallest reference force had a much smaller variance and no subjects had problems in detecting the stimulus. Thus, the detection threshold for forces seemed to be lower in the active condition. Due to the active movement, users might have gathered an additional work cue, which could have helped in the discrimination task. Moreover, the questionnaire revealed that three users focused their attention on the transition region where the force faded in.

In Figure 8, the obtained JND and Weber fraction values are shown. Again, the average JND values can be well approximated by a linear function (dashed line). To obtain a conservative estimate to the perceptual threshold, this fit was again shifted so that all averaged JNDs were above the linear function (red line). This shifted function has the intercept $JND(F = 0) = 0.04$N and the slope 8.4%. This value is somewhat lower than the one obtained in the passive condition, indicating higher sensitivity, but it is still in line with results in the literature [Allin et al. 2002]. For small forces, sensitivity decreased causing the Weber fraction to increase to 22% for the smallest force utilized. Even for the smallest force magnitude, however, sensitivity was higher than in the passive condition, showing that more sensory information was available for the discrimination of forces while participants were moving. This result is coherent with what was found in the passive condition.

Note, that neither the fitted line in Figure 6 nor this fit should be considered as true JND curves. As mentioned earlier, they only served as rough estimates of the perceptual thresholds. When comparing the cyan and the red line in Figure 8, it became clear that the fit of the second scenario provided the more conservative estimate. Moreover, the task in the second scenario was more comparable to a

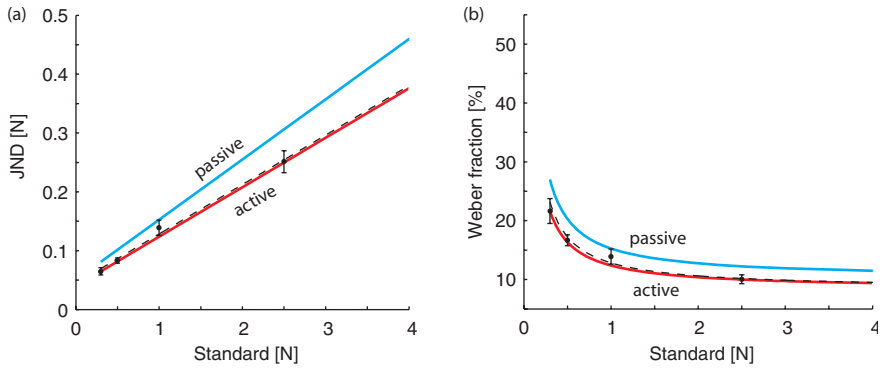

Fig. 8. (a) Mean JNDs over the standards in the active condition. The black dots are the average values from the experiment and the error bars indicate the standard error of the mean. The dashed line is a linear fit to the data. The red line is the shifted linear fit that provides a conservative estimate of the threshold. The cyan line is the shifted linear fit from Figure 6. (b) Mean Weber fractions over the standards. The plotted curves are obtained by dividing the linear functions in (a) through the standards.

typical interaction of a user with a haptic display device. Hence, we used this estimated perception threshold to guide the computation of the data-driven models in Section 5. As it was explained in Section 3, during the construction of the data-driven model, the algorithm iteratively evaluates the rendering error with respect to the threshold curve and refines the model in regions where the error exceeds the threshold. The rendering fidelity of the resulting algorithm was thereafter tested in a second user study.

5. STUDY 2: USER-BASED EVALUATION OF HAPTIC FEEDBACK

In order to evaluate the quality of the data-driven haptic feedback, a discrimination study was performed. The experiment was designed to only detect the presence of perceivable differences between real and virtual objects. However, the amount of difference was not quantified. This choice was based on the expected high accuracy of the rendering. The study focused on an "odd ball" stimulus detection task following a three interval forced choice (3IFC) design. Subjects had to discriminate between real objects of different stiffnesses and their virtual counterparts. A possible alternative would have been the use of multidimensional scaling (MDS) for quantifying the differences, as for instance employed in one of our earlier studies [Leskovsky et al. 2006].

Interaction between the tool and the real and virtual objects was restricted to one degree of freedom. Users could only indent the objects along the surface normal direction. No lateral movements or sliding contacts were considered. In the following, we first describe the generation of the virtual models and their relation to the considered real objects.

5.1 Generation of Virtual Objects

The real objects were two silicone cylinders, depicted in Figure 9(a). The haptic properties of these objects were recorded using the device presented in Hoever et al. [2009b]. Each cylinder was manually indented using the probe at different velocities, recording tool trajectory, and contact forces. In addition, in order to capture transient effects in the force signal, indentation was also halted at different penetration depths. The recorded forces are plotted against the penetration depth for both cylinders in Figure 9(b). As can be seen from the curves, cylinder 1 was softer than cylinder 2. Also, both objects exhibited visco-elastic behavior resulting in hystereses. Thus, the contact force is not uniquely defined as a function of the penetration depth. Moreover, both cylinders showed nonlinear behavior in their

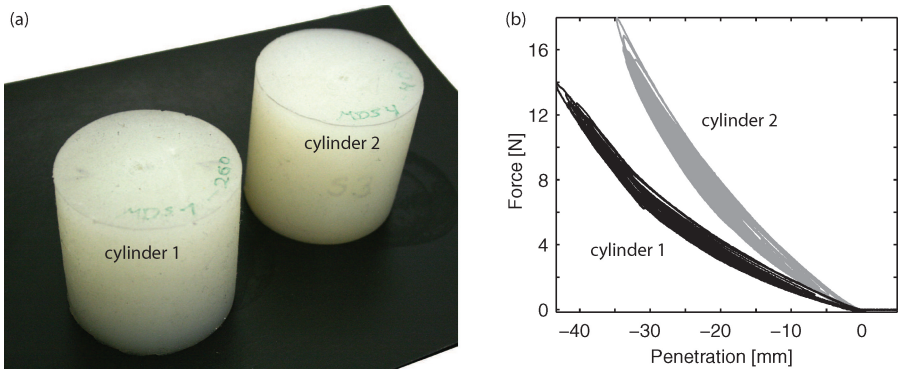

Fig. 9. (a) The two silicone cylinders that were used for the discrimination study. (b) Measured contact forces from a poking test where a user penetrated the two silicone cylinders with different velocities.

force response. Although the nonlinearity was rather smooth as opposed to more complex systems [Swindells and MacLean 2007], recovering this behavior was not trivial as we chose a generic model structure that does not explicitly address any particular material properties.

These recordings were used to create corresponding virtual objects employing our data-driven haptic rendering approach. For each of the real cylinders, we generated two virtual counterparts. The first only accounted for the elastic behavior (in the following referred to as static model V_{stat}), while the second also incorporated the dynamic material effects (denoted as dynamic model V_{dyn}).

Thus, six objects (two real and four virtual) were used in the study. While the static model only considered the tool position to calculate the contact force, the dynamic model employed two additional input dimensions, namely tool velocity as well as lowpass-filtered velocity. For the latter, a time constant of $\tau = 0.49$s was used according to Equation (2).

The derivation of the static and dynamic representations of the cylinders was conducted using the technique described in Section 3. The two static representations of the cylinders only used two radial basis functions, since more refined approximations did not improve the quality of the fit with respect to the previously determined threshold curve. The dynamic representation of the soft cylinder employed 20 RBFs. For this, the error for all training data was below the threshold curve. The dynamic representation of the stiff cylinder used the maximum number of 50 RBFs. A small number of force errors of the training data lied above the threshold.

For a first offline assessment of the virtual cylinders, we performed additional recordings with both real cylinders and used these recordings as test data. The test data were created using the same interaction approach as before, yielding tool trajectories and contact forces. The test data from the soft cylinder contained 53.580 samples, the one from the stiff cylinder 42.047 samples. The recorded trajectories were fed into the data-driven representations of the objects for off-line calculations. In case of the static models, only the position data was used, while in case of the dynamic model, also the recorded tool velocity was included. The resulting forces were then compared to the measured ones individually for each sample. In Figure 10, the relative error of each computed force sample is plotted against the magnitude of the corresponding measured force value for the different virtual cylinders. Since each force magnitude occurs several times in the test data, there are multiple data points for each force magnitude in the plot. The perceptual threshold obtained in Section 4 is depicted in red.

The plots in Figures 10(a) and 10(b) show that the rendering error of the static models was considerably above the threshold curve for most of the samples. Hence, this quality measure suggests that the static virtual models provided a perceivably different force feedback compared to the real objects. The

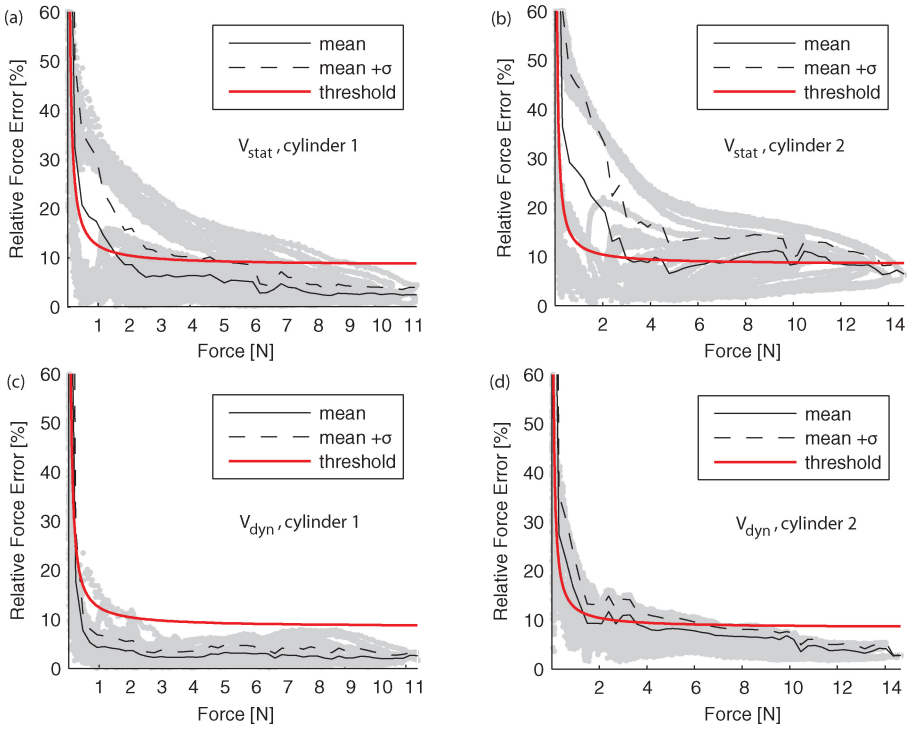

Fig. 10. Relative force errors over the force magnitude for the test trajectories using the four different virtual objects. One test trajectory was used for each of the two real cylinders. At the top, the results for the first (a) and second (b) cylinder are shown using the corresponding static models V_{stat}. At the bottom, the results for the first (c) and second (d) cylinder are shown using the corresponding dynamic models V_{dyn}. In red, the threshold curve from Figure 8(b) is shown.

forces computed by the dynamic models exhibited a considerably smaller relative error. In case of the soft cylinder (Figure 10(c)), the relative force error almost always lied beneath the threshold. For the stiffer cylinder, the error was slightly above the threshold for some samples, as shown in Figure 10(d). However, the individual, theoretical examination of the force error at each test sample only provides a first estimate of the fidelity of the virtual models. Therefore, we assessed the rendering accuracy in the user study described in the following text.

5.2 Hardware Set-up

The study set-up included a PHANToM haptic interface that was mounted on a frame. A probing rod attached to the PHANToM arm was guided by a bearing mechanism only allowing vertical movements. A turntable was put underneath the tool for quick switching of objects in the study. An overview of the set-up is depicted in Figure 11.

The turntable was divided into six equal slices. The two real cylinders were located at positions 1 and 4 on the turntable. At positions 2 and 3, the virtual counterparts V_{stat} and V_{dyn} of the soft cylinder were placed, and at positions 5 and 6, the counterparts of the stiffer cylinder. An encoder was used to read the turntable orientation in order to automatically switch the rendering according to the current position. If a real object was underneath the tool, no virtual feedback was rendered.

An important limitation of the PHANToM Premium 1.5 is the maximum force that can be continuously rendered. According to the manufacturer, this force is 1.4N, while a force of 8.5N can be

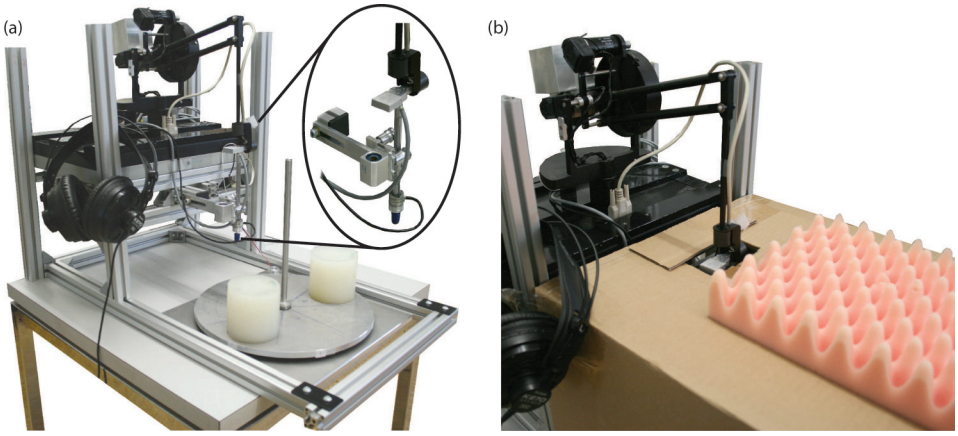

Fig. 11. (a) Experimental set-up for the discrimination study. The PHANToM was mounted on a frame and the tool was guided by a bearing allowing only vertical movements. A turntable was put underneath the tool to quickly switch objects. (b) Covered set-up as it was used during the study. Users did not see the objects they were touching.

rendered for short periods [SensAble 2009]. High forces can lead to overheating of the motors. The manufacturer's software drivers use a built-in security mechanism to avoid overheating. However, in our system, we employ proprietary drivers to operate the PHANToM. Therefore, we conducted some experiments to measure the actual motor temperature for different continuous force levels. Only the actuator for the vertical tool movement was measured, since in our configuration, the other motors were not activated. At a force level of 2N, the motor body heated up to 75°C within 15min. At 4N, the temperature rose to 80°C after 2.5min. For larger forces, the temperature increase was very rapid. Therefore, we decided to not use such large forces in the study.

Furthermore, in order to increase the force range that could be displayed, we added a counterweight at the back of the PHANToM, as shown in Figure 11(b). A similar configuration was also used before in Leskovsky et al. [2006]. This counterweight caused a permanent force of 2N pulling the tool upwards. Hence, during the study, we offset our forces by −2N. With this modification, it was possible to display forces up to 6N. Moreover, we limited the study sessions to 15min so that the motor temperatures stayed in the range of 75°C to 80°C. During the breaks, the motors were switched off so they could cool down again. Nevertheless, it should be noted that due to the counterweight, the freespace inertia of the device increased.

5.3 Participants

Eight participants, seven male and one female, took part in the discrimination study. The subjects were recruited from lab members. All subjects were naive to the purpose of the experiment. The age ranged from 24 to 28. One of the subjects was left-handed, and all others were right-handed.

5.4 Stimuli

In each trial, the subject was presented with three objects out of which two were identical and one was different. These were selected from the two real and four virtual objects on the turntable. No comparisons were performed across the two stiffness groups. This results in three different comparisons for each stiffness group: R versus V_{stat}, R versus V_{dyn}, and V_{stat} versus V_{dyn}, where R stands for the real object. Each of these combinations was presented 18 times resulting in an overall number of 3(comparisons) × 2(stiffnesses) × 18(repetitions) = 108 trials for the complete study. The order of

these trials was randomized and grouped into four blocks each containing 27 trials. For each object in a trial, the subject was given 5s to explore the object with the tool. Beginning and end of this phase were marked by acoustic signals.

5.5 Procedure

Subjects were asked to sit next to the experimental set-up and to adjust the height of the seat and the backrest so they could comfortably put their dominant arm on a foam cushion, seen in Figure 11(b). They grasped the PHANToM at the center of the vertical arm with their thumb, index, and middle finger and were told to only use their wrist joint to move the probing rod up and down. Based on the force range of the device and the force responses shown in Figure 9(b), users were asked to limit their maximum penetration depth to 20mm. This depth was visually marked on the PHANToM arm so the subjects could learn this range during the training session. However, we did not use any hard stops to enforce this maximum depth. Moreover, users were asked to always stay in contact and in command of the tool movement. Thus, it was not allowed to throw the tool at the object or let it slip between the fingers. This ensured a minimum human impedance during the interaction with the object. No further instructions for the tool movement were given so participants could intuitively explore the objects.

Subjects had to wear headphones in order to mask any noise generated by the motors of the PHANToM, which could have helped distinguishing real from virtual objects. Also, sound beeps were provided to indicate the start of the exploration as well as the end, when the tool ought to be moved back. This allowed the experimenter to move the turntable to the next position. If in one trial the identical objects were rendered in immediate succession, the turntable was still moved by the experimenter in order to avoid giving clues about the objects.

Participants were asked to detect which of the three stimuli was the “odd ball.” At the end of each trial, the subject had to tell which of the three objects felt “different.” However, this difference was not specified or quantified in any way as subjects were told that every detected difference could serve as an indication of the odd ball.

5.6 Training Phase

Each subject started with a training session in which at least 10 test trials were performed and feedback about the correctness of the answer was given. During the training, participants were allowed to look at their hand movement in order to learn the desired indentation range. The training lasted about 10min.

5.7 Experiment

During the real study, the subjects were asked to close their eyes or to look into another direction. This was to ensure that they did not see their movements and thus completely relied on the haptic feedback. No information about the correctness of the answers was given. Subjects needed on average 15min to complete one block and had breaks in between, lasting at least 10min.

5.8 Results

Figure 12 shows the resulting probabilities with which the individual subjects correctly selected the deviating object. The chance level for this task is marked as a dashed line. It indicates when the odd ball stimulus in the triplet could not be discriminated from the others.

In case of the soft objects (Figure 12(a)), the static virtual model V_{stat} could usually be distinguished from the corresponding real object. In contrast, the dynamic virtual model V_{dyn} could mostly not be discriminated from the real object. The probability for all subjects was around chance level and always

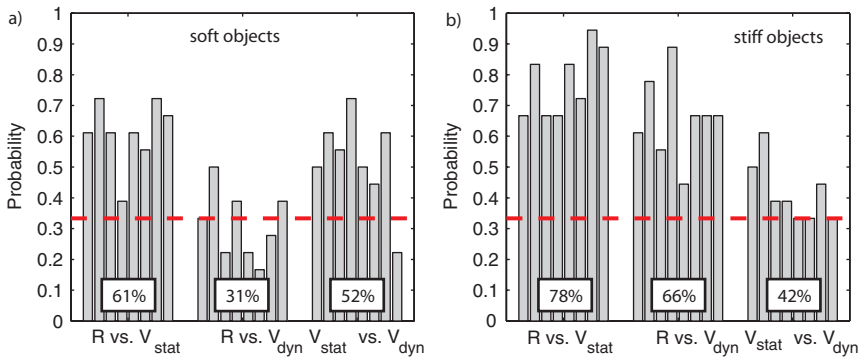

Fig. 12. Probablities with which the subjects correctly selected the deviating object, if (a) the soft objects and (b) the stiff objects were compared. The chance level of 33% is marked as a dashed line. The mean probabilities across subjects are printed for each combination of objects in the boxes.

lied in the 95% confidence interval for each subject. This indicates that for soft objects, the dynamic virtual model V_{dyn} and the real object were perceptively indistinguishable. Finally, most subjects could discriminate the two virtual models if both were directly compared.

The results in Figure 12(b) reveal that subjects could discriminate real and virtual objects in case of the stiff stimuli, regardless if the model incorporated the dynamic properties or just rendered the static material behavior. In contrast to soft objects, most of the subjects could not distinguish the two virtual models if both were directly compared.

The main finding emerging from this data is that the static model V_{stat} was too weak to provide high fidelity feedback. The dynamic material properties were essential for achieving high accuracy. Especially for soft objects, the dynamic model gave very good results.

Nevertheless, in case of the stiff objects, the dynamic model could be discriminated from the real object. On average, the discrimination probability was slightly lower than for the corresponding static model, but at 66%, it was clearly above chance level.

One possible reason for this result could be the slightly less accurate rendering of the stiff dynamic model, as seen in Figure 10(d). Some of the test samples lied above the threshold curve, however, the deviations from the threshold were rather small when compared to the static model in Figure 10(b). In contrast to this, the detection probabilities for R versus V_{stat} and R versus V_{dyn} in Figure 12(b) suggest that both models provided a similar accuracy. Moreover, for the usual interactions that subjects performed in the study, the relative force error lied well below the threshold curve used in Figure 10. Thus, this did not provide a reasonable explanation for the rather weak performance of V_{dyn} in case of the stiff objects. Therefore, this finding was examined further, as will be discussed in the next section.

5.9 Limitations of the Hardware Set-up

In order to explain the weak performance of V_{dyn} for the stiff objects, we first investigated the limited interaction range that was provided in the study. Subjects were only allowed to penetrate the objects by 20mm. However, during the recording of the real silicone cylinders, the objects were actually indented deeper. Thus, the training data, from which the virtual model was derived, covered a different range of interactions.

To examine possible issues caused by this difference, we recorded both silicone cylinders in two different ways. In the first recording session, we penetrated the objects as before using a maximum penetration of 40mm for the soft silicone and 35mm for the stiff silicone. For each cylinder, we recorded two

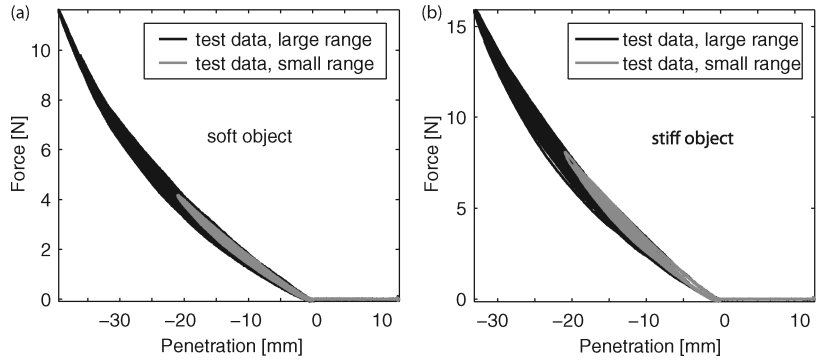

Fig. 13. Recorded test data from the (a) soft and (b) stiff silicone cylinder. The black curves show the measurements using large penetration depths. The gray curves show the measurements using the small penetration depths as they were used in the study.

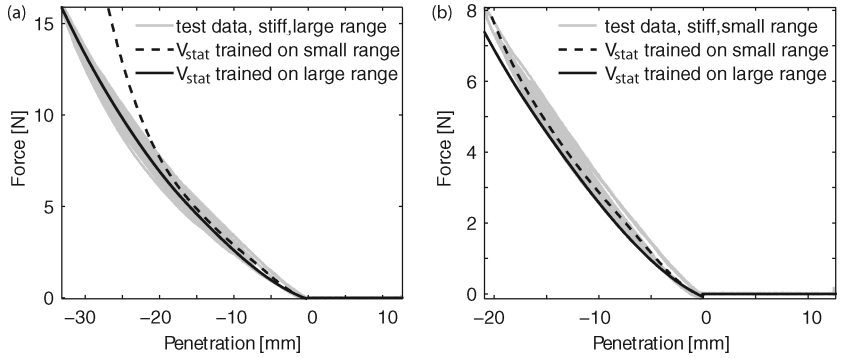

Fig. 14. Static virtual models based on the recordings of the stiff silicone cylinder. The solid black curves show the resulting interpolant, if the model was trained on the large-range data. The dashed black curves show the resulting interpolant if the model was trained on the small-range data. In (a), the interpolants are tested on the large-range test data; in (b), they are tested using the small-range test data.

data sets, one for the training of the data-driven model to determine the coefficients of the interpolation function and one data set for testing. In the second recording session, we penetrated the objects only up to 20mm as the movement required in the experiment. Again, for both silicones, we recorded two data sets, one for training and one for testing.

In Figure 13, the test dataset for each of the different recordings is shown. The resulting static models, which extracted the mean curve through the hysteresis, revealed that the average stiffness was somewhat increased for the small indentation ranges. These effects on V_{stat} can clearly be seen in Figure 14 for the stiff silicone cylinder. If V_{stat} was trained on the large range-data (solid black curves), it gave accurate results for the large-range test data as can be seen in Figure 14(a). If the same interpolant was tested on a small interaction range, it rendered a feedback that was slightly too soft (Figure 14(b)). A corresponding behavior could be observed if V_{stat} was trained on the small range data (dashed black curves). Here, it performed well on the small-range test data in Figure 14(b), but rendered too high stiffnesses for large-range test data in Figure 14(b). Hence, the static models were sensitive to the range of movements that were performed in the training.

The dynamic model did not exhibit this sensitivity. For both types of training data covering small

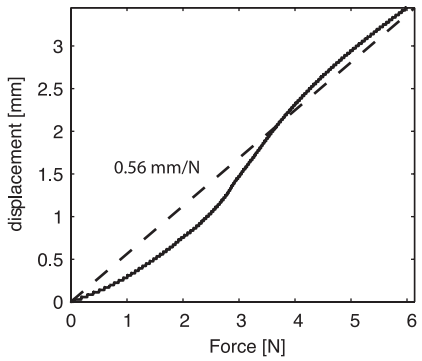

Fig. 15. The solid curve shows the measured displacement of the PHANToM arm over the applied force created by the motor. The measurements are based on the encoder readings. The gimbal of the PHANToM was locked to a fixed position so no real displacement of the tool occured. The dashed line shows the linear fit through the data and has a slope of 0.56 mm/N.

and large ranges of penetration depths, the dynamic model was capable of rendering feedback with a similar high accuracy. This demonstrated the capability of this model in extrapolating the force feedback even beyond interactions that were covered in the training.

Therefore, the varying average stiffnesses that were observed for different penetration ranges could not fully explain the weak performance of V_{dyn} in Figure 12(b). Nevertheless, this effect showed the importance of incorporating the dynamic effects into the data-driven model in order to provide high fidelity feedback.

Despite the fact that we have chosen small force ranges to be rendered in order to limit motor temperatures, we investigated whether a decrease in the rendered force due to heating could be responsible for the low performance in case of stiff objects. During the study, the housing of the motor reached a maximum temperature of 75°C to 80°C. This increase in temperature caused the resistance of the motor winding to grow, which reduced the current through the winding and, therefore, the motor torque and rendered force. To analyze this decrease of the rendered force, we applied a constant current to the PHANToM motor that generated a force of 6N at the tool. This was the maximum force during the study. After 2 minutes, the motor housing had a temperature of 71°C, and the force magnitude had decreased by 0.25N. This is less than 5% of the correct value. Considering the perceptual thresholds depicted in Figure 8(b) as well as the Weber fractions reported in Pang et al. [1991], the caused change in the force signal was well below the JND. Thus, the motor heating by itself should only have a minor effect on the rendered forces. Nevertheless, its impact added up to other effects and decreased the quality of the rendered feedback.

Finally, we questioned whether the compliance of the haptic device could play a role in the low performance of V_{dyn} for the stiff stimuli. The PHANToM utilizes encoders that are located at the motors at the back of its mechanical structure. Therefore, the tool at the gimbal with which the user interacts is not collocated with the position sensors. So, the compliance of the mechanical set-up causes the encoders to provide erroneous position readings if forces are applied to the PHANToM arm. A measurement of this effect is presented in Figure 15. In this test, the PHANToM arm was configured as in the study and the gimbal at the end effector was locked in a fixed position using a screw clamp. Using this set-up, a linear force ramp over time was sent to the PHANToM. The measured tool position based on the encoder readings is plotted as a solid curve in the figure. The dashed line is a linear fit to the data. In this force range and configuration, the device mechanics had an average compliance of 0.56mm/N. This compliance of the device acted in series with the rendered virtual objects. Hence, the displayed

compliance was always larger than the compliance of the model itself. The average compliance of the stiff silicone cylinder is 2.5mm/N. This value was increased by 22% due to the measured device compliance. In case of the soft silicone, the average compliance was increased by 11%. Considering, that subjects were trained to utilize a penetration depth of 20mm they were able to incorporate terminal force and work cues to discriminate compliances. According to Tan et al. [1995], the Weber fraction for compliances under these conditions is between 5% and 15% with an average of 8%. Thus, in the case of the stiffer silicone cylinder, the deviation of the average compliance was clearly larger than the corresponding compliance JND, while for the softer sample object the deviation was in the range of the JND.

In other words, if a virtual model of the stiff silicone was rendered, the subject could perceive a difference in compliance compared to the real silicone object. In case of the soft material, the difference was close to the perceptual threshold and thus, possibly harder to detect.

6. DISCUSSION

6.1 JND Study

We examined the JND for forces in two conditions: one in which the user was passive and one in which the user actively moved. The resulting values of the JND from both experiments were approximated by linear functions over the range of standard force magnitudes. The linear fit was shifted to obtain a conservative estimate of the perceptual threshold. For the active condition, in which the user moved the tool, we obtained lower JNDs for all examined standards. Moreover, the Weber fractions for the smallest tested force magnitude are much more consistent across subjects than for the passive condition. Thus, we concluded that the detection as well as the discrimination threshold of users was, on average, lower if the user actively moved her hand. A possible explanation for this difference in sensitivity could be an additional work cue during the interaction, as it was also proposed by Tan et al. [1995] for the discrimination of compliances.

The values obtained in this study are one of the few characterizations of force sensitivity across a wide range of pedestal forces while holding a stylus, as it is typical in most commercially available haptic interfaces. Results indicated a gradual decrease of sensitivity with deviation from the Weber fraction reported in the literature even for forces of 1N. The decrease in sensitivity was more pronounced with static posture of the hand holding the stylus. It was not clear from the results how the difference in sensitivity in the active and passive conditions was related to a difference in force detection. It must be noted that the fitted lines in Figures 6 and 8 do not represent psychometric functions capturing the user sensitivity to force difference in the considered force range. As mentioned earlier, the fits performed here followed our purpose of determining the most conservative estimate of sensitivity, which ensured that no artifacts of the optimized data-driven model could be detected by the user. Due to the higher sensitivity to force difference in the active condition, this curve provided the most conservative estimate. In addition, the task in the active condition closely resembled the intended utilization of the haptic device and the rendering algorithm developed. Hence, we used the active estimate to guide the computation of the data-driven models in Section 5.

6.2 Discrimination Study

The results of the discrimination study showed that, for soft objects, the dynamic virtual model could not be discriminated from the real object. If the dynamic material effects were neglected in the rendering, however, users could clearly discriminate the virtual and real feedback. Hence, the study confirmed our earlier results with offline tests [Hoever et al. 2009b] that the incorporation of the dynamic material behavior is crucial for high-fidelity feedback.

In case of the stiff object, both, the static and dynamic models could be discriminated from the real object. An exhaustive analysis of the study set-up revealed that this result could be explained by the limited stiffness of the haptic display hardware. For our configuration of the hardware, the average device compliance was 0.56mm/N. This limitation caused the compliance of the stiffer virtual objects to increase by 22%, which is clearly above the corresponding perceptual threshold. Thus, both stiff virtual objects felt considerably softer than their real counterpart and could be easily discriminated by the subjects. This is also confirmed by the lack of discrimination if both virtual models were directly compared with each other.

Considering that the stiffness of all tested virtual objects were in the range of typical virtual renderings and did not exceed the stability boundary of the device, these findings are generally relevant for studies, where virtual and real objects are compared. For instance, the results in Leskovsky et al. [2006], revealed a compression of the perceived virtual stiffnesses when compared to real objects of similar stiffness in an MDS analysis. The authors employed the same device as in our study. Hence, this effect can be directly explained by our findings about the device compliance. Indeed, the amount of compression that was observed in Leskovsky et al. [2006] corresponds well to our compliance estimate.

It should be noted that the previous considerations only serve as an approximation. On the one hand, the JND for compliances that we used from Tan et al. [1995] was only obtained for a standard compliance of 4mm/N, while our sample objects had an average compliance of 5mm/N and 2.5mm/N. On the other hand, we just examined the change in compliance regarding the average compliances of the sample objects. According to our measurements, both objects exhibited a nonlinear behavior for which the compliance changes with the penetration depth. Nevertheless, the impact of the device compliance on the display of the stiff virtual objects was very large. Hence, we can conclude that this effect was the main reason why the stiff virtual objects could be easily discriminated from the real ones.

In the analysis of our test set-up, we also found that the measured average stiffness of the sample objects depended on the maximum penetration depth that was used in the recording. This caused an inconsistency in the derived static models that were based on recordings with large penetrations but were tested in the study using only small penetrations. This inconsistency was resolved if the data-driven models also considered the dynamic material properties. Therefore, this finding confirms the importance of the dynamic material properties. Their extrapolation capabilities allow to use training data that only cover small penetration depths. If the user then uses larger penetrations that are not covered by the training data, the model still provides highly accurate feedback.

7. CONCLUSION

This article presents two studies. The first study examined just noticeable differences of forces for small force magnitudes and was motivated by the need for a perceptual threshold for the data-driven rendering technique presented in Hoever et al. [2009a]. Two different conditions were investigated. In the first condition, the user was completely passive while in the second he moved his arm in order to compare forces. The results from both conditions show that the Weber fraction severely increases for smaller forces and strongly deviates from the literature value of approximately 10%, which applies for larger forces. The determined JND values could be approximated by a linear function.

For the data-driven haptic rendering, we only used the results from the second scenario since it provided the lower threshold curve and thus led to a more conservative error measure during the training of the virtual model. Also, the task in the second scenario better resembled the typical interaction with the employed haptic display.

The application of the perceptual threshold in the training phase of the data-driven algorithm assures a perception-based fitting to the measured force data. To assess the quality of the resulting

virtual models, we performed a discrimination study in which users directly compared the virtual objects with their real counterparts. Regarding soft materials, the results of this second study showed that the data-driven haptic feedback is indistinguishable from the real feedback if the dynamic material effects are taken into account. If only the static material properties (i.e., the elasticity) was considered, the virtual model could be discriminated from the real one. For the stiffer sample material, the compliance of the employed haptic device cannot be neglected. It limits the achievable fidelity by rendering the stiff virtual objects considerably softer. Hence, these objects can still be discriminated from their real counterparts. A stiffer display device is required in order to maintain the high quality of the rendering as it was observed for the softer sample object.

In addition, we showed that the consideration of dynamic material effects is crucial in order to obtain an interpolation that is independent on the range of the recorded training data. If the model assumes a purely elastic material behavior, the resulting interpolation largely varies with different ranges for the penetration depth. This sensitivity to the training data is eliminated if the interpolation also allows for dynamic effects like viscosity.

REFERENCES

ALLIN, S., MATSUOKA, Y., AND KLATZKY, R. 2002. Measuring just noticeable difference for haptic force feedback: Implications for rehabilitation. In *Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 299–302.

ANDREWS, S. AND LANG, J. 2007. Interactive scanning of haptic textures and surface compliance. In *Proceedings of the International Conference on 3D Digital Imaging and Modeling*. IEEE, Los Alamitos, CA, 99–106.

COLTON, M. AND HOLLERBACH, J. 2007a. Haptic models of an automotive turn-signal switch: Identification and playback results. In *Proceedings of the 2nd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 243–248.

COLTON, M. AND HOLLERBACH, J. 2007b. Reality-based haptic force models of buttons and switches. In *Proceedings of the International Conference on Robotics and Automation*. IEEE, Los Alamitos, CA, 497–502.

ENGEN, T. 1971. *Woodworth and Schlosberg's Experimental Psychology*. Holt, Rinehart and Winston, Geneva, IL, 11–46.

GESCHEIDER, G. A. 1985. *Psychophysics: Method, Theory, and Application*. Lawrence Erlbaum Associates, New York, 1–36.

HINTERSEER, P. AND STEINBACH, E. 2006. A psychophysically motivated compression approach for 3D haptic data. In *Proceedings of the 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 35–41.

HINTERSEER, P., STEINBACH, E., HIRCHE, S., AND BUSS, M. 2005. A novel, psychophysically motivated transmission approach for haptic data streams in telepresence and teleaction systems. In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*. IEEE, Los Alamitos, CA, 1097–1100.

HIRCHE, S., HINTERSEER, P., STEINBACH, E., AND BUSS, M. 2007. Transparent data reduction in networked telepresence and teleaction systems. *Presence Teleoperator. Virtual Environ. 16*, 523–531.

HOEVER, R., DI LUCA, M., SZEKELY, G., AND HARDERS, M. 2009a. Computationally efficient techniques for data-driven haptic rendering. In *Proceedings of the 3rd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 39–44.

HOEVER, R., HARDERS, M., AND SZEKELY, G. 2008. Data-driven haptic rendering of visco-elastic effects. In *Proceedings of the 16th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 201–208.

HOEVER, R., KOSA, G., SZEKELY, G., AND HARDERS, M. 2009b. Data-driven haptic rendering-from viscous fluids to visco-elastic solids. *IEEE Trans. Haptics. 2*, 15–27.

JONES, L. AND HUNTER, I. 1990. A perceptual analysis of stiffness. *Exp. Brain Res. 79*, 150–156.

JONES, L. A. 1989. Matching forces: Constant errors and differential thresholds. *Perception. 18*, 681–687.

KAWAI, S. 2003. Heaviness perception. *Exp. Brain Res. 153*, 297–301.

KUCHENBECKER, K., FIENE, J., AND NIEMEYER, G. 2005. Event-based haptics and acceleration matching: Portraying and assessing the realism of contact. In *Proceedings of the 1st Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 381–387.

KUCHENBECKER, K., FIENE, J., AND NIEMEYER, G. 2006. Improving contact realism through event-based haptic feedback. *IEEE Trans. Visual. Comput. Graphics 12*, 219–230.

LESKOVSKY, P., COOKE, T., ERNST, M., AND HARDERS, M. 2006. Using multidimensional scaling to quantify the fidelity of haptic rendering of deformable objects. In *Proceedings of the 1st Joint EuroHaptic Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Sysetems*. IEEE, Los Alamitos, CA, 289–295.

LUCE, R., BUSH, R., AND GALANTER, E. 1963. *Handbook of Mathematical Psychology*. Wiley and Sons, Inc., New York.

MACLEAN, K. 1996. The "haptic camera": A technique for characterizing and playing back haptic properties of real environments. In *Proceedings of Dynamic Systems and Control Division*. ASME, New York, 459–467.

NORWICH AND WONG. 1997. Unification of psychophysical phenomena the complete form of fechner's law. *Percept. Psychophys 59*, 929–940.

OKAMURA, A., WEBSTER, R., NOLIN, J., JOHNSON, K., AND JAFRY, H. 2003. The haptic scissors: Cutting in virtual environments. In *Proceedings of the 2003 International Conference on Robotics and Automation*. IEEE, Los Alamitos, CA, 828–833.

PANG, X., TAN, H., AND DURLACH, N. 1991. Manual discrimination of force using active finger motion. *Percept. Psychophys 49*, 531–540.

RICHARD, C., CUTKOSKY, M., AND MACLEAN, K. 1999. Friction identification for haptic display. In *Proceedings of the ASME Dynamic Systems and Control Division*, vol. 67. ASME, New York, 327–334.

ROSS, H. E. AND BRODIE, E. E. 1987. Weber fractions for weight and mass as a function of stimulus intensity. *Exp. Psych. 39*, 77–88.

RUFFALDI, E., MORRIS, D., EDMUNDS, T., BARBAGLI, F., AND PAI, D. 2006. Standardized evaluation of haptic rendering systems. In *Proceedings of the 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. IEEE, Los Alamitos, CA, 225–232.

SENSABLE. 2009. http://www.sensable.com/, visited 09/2009.

SWINDELLS, C. AND MACLEAN, K. E. 2007. Capturing the dynamics of mechanical knobs. In *Proceedings of the 2nd Joint EuroHaptic Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Sysetems*. IEEE, Los Alamitos, CA, 194–199.

SWINDELLS, C., MACLEAN, K. E., AND BOOTH, K. S. 2009. Designing for feel: Contrasts between human and automated parametric capture of knob physics. *IEEE Trans. Haptics 2*, 200–211.

TAN, H. Z., DURLACH, N. I., BEAUREGARD, G. L., AND SRINIVASAN, M. A. 1995. Manual discrimination of compliance using active pinch grasp: The roles of force and work cues. *Perception and Psychophysics*. Vol. 57. 495–510.

WEBER, E. H. 1978. *De subtilitate tactus - The sense of touch (translated original from 1834)*. Academic Press, London.

Received September 2009; revised January 2010; accepted February 2010