The Role of Torque in Haptic Perception of Object Location in Virtual Environments

Steve Wang and Mandayam A. Srinivasan
The Touch Lab
Dept. of Mechanical Engineering and The Research Laboratory of Electronics
Massachusetts Institute of Technology, Cambridge, MA 02139.
srini@mit.edu

Abstract

An experimental study was performed with human subjects to determine the role of torque feedback in purely haptic perception of object location within virtual environments. The experimental hardware consisted of two Phantom haptic interface devices connected by a common stylus. Ray-based rendering technique that models the user-controlled stylus as a line was used for computing collision detection with a virtual object and its force as well as torque response. The subjects were trained with correct-answer feedback to obtain their best performance. Results demonstrate that the most significant improvement in perception occurred during the first training session. They also show that identification of object location by purely haptic cues through a tool can be accomplished in two distinct ways: (1) with full force and torque feedback, even when only tapping with a fixed orientation of the stylus is permitted; and (2) with only force at the stylus tip reflected back to the user, when multiple stylus orientations ('rocking') are allowed in contacting the object. Under these conditions, the estimated JND, which is expected to be an upper bound, ranged from about 22% for the nearest locations to 12% for the farthest locations.

1. Introduction

Successful implementation of haptics in Virtual Environments (VEs) requires a proper balance between hardware, software, and human perceptual abilities (see Basdogan and Srinivasan, 2002 and Biggs and Srinivasan, 2002 for recent reviews of haptic interfaces and haptic rendering). While haptic exploration simulated by most force-feedback devices in VEs today takes place between a hand-held probe and the environment, very little is known about haptic perception through a tool. Several studies have been conducted on haptic perception through probing of real objects in real environments (REs) (e.g., Chan and Turvey, 1991; Carello et al., 1992; Chan, 1995; Chan, 1996; Klatzky and Lederman, 1999; Lederman et al., 1999; O'Modhrain, 1999). In addition, investigations into haptic perception in VEs must be conducted to determine optimal hardware and software to simulate touch.

In REs, the interactions between objects are dictated by physics and cannot be altered. In VEs however, a programmer defines the physics of the environment. Therefore, many experiments that cannot be conducted in the RE can be conducted in a VE. For example, the programmer can choose to provide the user with only a limited set of haptic cues or create completely non-physical responses to interactions with the environment. Examination of the role of torque in object perception through a tool is an example of studies that when conducted in the virtual world may yield results not directly achievable in the real world.

From a designer's viewpoint, understanding the role of torque in object perception through a tool has important implications. In terms of hardware, for a given task, are 6 DOF force/torque feedback needed or are 3 DOF force feedback sufficient? For rendering software, is point-based rendering that can compute only the forces at the tool tip adequate? Or is a ray- or 3-D object-based collision detection and response algorithm needed? Although VEs provide the ability to separate the multiple cues present in the RE, to ensure the validity of the results, it is first necessary to establish that any deterioration in perception is not due to poor hardware or software design. Only then can perception with reduced haptic cues be tested. The goals of this study are to conduct experiments in:

1) RE to determine if haptic perception of object location through a tool is possible in general.

2) VE with complete haptics (force and torque) display to verify effectiveness of hardware and software.

3) VE with reduced haptics display to determine what cues are necessary and sufficient to identify object location.

In this work, the effects of training for all three cases were also investigated so that the differences in the three cases above are not due to the lack of familiarity with the task or the devices. To ensure that a reasonable
comparison could be made between perception in RE and VE, the highest possible level of consistency among the experimental set ups was maintained.

2. Experiment 1 – Real Object (7 object locations; unrestricted probing)

Experimental Setup: The end-effectors of two Phantom haptic interface devices (Model A1.0 and Model T1.0, each capable of 3 DOF force feedback) were connected as shown in Figure 1, except that the connecting hollow aluminum rod was much longer (54cm). A large cardboard box and cloth screen hid the apparatus from the subject’s view. A cut in the cloth screen allowed the subject to reach through and grasp the probe as in Figure 1. A triangular prism shaped rubber eraser was placed on a flat surface in front of the hand location such that its top edge was perpendicular to the rod. The subjects were instructed to identify the location of the edge by contacting it through the rod. A tape measure ran along the side of the cardboard box parallel to the rod for the subject to indicate, by pointing, where he/she perceived the object to be inside the box. Pointing to the perceived object location provided subjects with a physical reference point to judge the position of the stimulus rather than merely choosing a numerical abstraction. The subjects wore headphones to mask the sound of contact between probe and object as well as any sound from the motion of the Phantoms.

Software: In each experimental trial, the subject identified one edge location. Software written in C++ was used to record the results, and during training sessions, it also provided correct-answer feedback for each trial after the subject’s response. While data from all experimental trials were recorded, trials in which the probe was not moved and oriented in accordance with the experimenter’s instructions were tagged and not included in the analysis. The position and orientation of the probe was monitored through the position encoders of both Phantoms, but the force/torque feedback capabilities of the Phantoms were turned off for experiment 1.

Training: Training occurred in three stages. In the first stage, the subject’s initial performance was measured without correct-answer feedback. This stage provided a baseline performance with which the subsequent results could be compared. In the second stage of training, the same object and the same set of object positions were presented to the subject, this time with correct-answer feedback following each trial. Subjects completed a total of four sessions of these experiments. The final step in the training was a repetition of the first stage in order to measure the performance of these experienced subjects.

Procedure: Subjects were asked to sit in a chair facing the computer monitor such that he/she could comfortably grasp and manipulate the rod with the right hand. The rod was grasped as one would a pointer. A piece of tape on the rod indicated where the subject was to grip the probe so that grasp position was kept constant for all trials, experiments, and subjects. The subject was asked to grasp the probe, and, using any exploratory style they chose so long as the probe remained approximately in the vertical plane, make contact with the object. Subjects were permitted as much time as they needed to judge the position of the object. The position of the object was fixed in each trial, but was changed randomly among 7 different distances (ranging from 9 and 33 centimeters forward of the hand position at 4cm increments) from trial to trial. To simplify the problem, only the z-component of the object position was varied. (Axes orientations are such that the z-axis is the longitudinal axis of the probe when it is parallel to the ground, the y-axis is perpendicular to the z-axis and vertical, and the x-axis is perpendicular the z-axis and horizontal). There were a total of ten sets of trials for each experimental session, with a total of six sessions (see below). A total of twenty trials for each stimulus per experimental condition were presented to each subject. In addition, one practice set was given to the subjects at the beginning of each session.

Subjects: Eight naive subjects from the MIT community ranging in age from 18 to 27 were paid to participate in this experiment. All subjects were right-hand dominant.

Results: Figure 2 shows plots of subject response versus actual object position for the initial and final stages of training for Experiment 1. The object positions are measured according to their distance from the hand before the subject begins probing, which is constant for each trial. The dashed line represents perfect identification of each object position for every trial. It can be seen that the data points, which are the mean of all responses for a given stimulus (object distance) averaged across all subjects, are monotonically increasing with object distance and approximately linear (in general, the mean responses actually form an S-curve). The line segments
Figure 2: Results for the real object

represent best-fit straight lines (to be referred to as the mean subject response) through these data points. Whereas the slope of the mean subject response line gives an indication of how well subjects perceived object distances within a group of objects, bias shows us how well subjects perceived the position of the group of objects as a whole (i.e., how the entire set of subject responses is shifted with respect to the set of actual object locations). The bias is computed by taking the average of the difference between the mean subject response curve and the ideal performance curve. Mathematically, this value is equivalent to the vertical offset of the midpoint of the mean subject response line segment from the ideal performance line. We chose a sign convention such that a positive bias denotes an average underestimation of object distance on the part of the subject, while a negative bias indicates an average overestimation. Ideal performance can be characterized as having unity slope with no standard deviation and zero bias.

Initially, before any correct-answer feedback was given to the subjects, perception of object distance was not very accurate or consistent. While the monotonic increase in perceived position indicates the ability to properly order the stimuli by distance, the 0.63 slope of Line A in Figure 2 implies that, on average, perceived distance between objects was 2.52 cm for an actual separation of 4 cm. The average standard deviation for the seven object positions was 7.49 cm, which tells us there was a large degree of scatter for each stimulus. Training through correct-answer feedback proved to be an effective method for improving haptic perception of object distance as indicated by Line B in Figure 2, which represents the test stage, following the correct-answer feedback sessions. The 0.9 slope implies that the 4 cm distance between objects was perceived to be 3.6 cm and the small bias of 0.17 cm indicates there was less than a 2 mm underestimation on average. There was also more confidence in each response as seen by the average standard deviation of 3.23 cm. Although the mean response does not exactly match the actual position, it can be seen from Line B in Figure 2 that for each stimulus, actual object distance lies within one standard deviation of the mean.

Figures 3a and 3b show the effect of training on subject performance, in terms of slope and percentage bias, respectively, as a function of the phase of the training, from the initial stage through the four correct-answer feedback sessions, and finally the test stage. The slope increased considerably between stages one and two of training from 0.63 to 0.89. In the remaining training stages, the slope did not vary by more than 5%. Percentage bias is the offset of the mean subject response line normalized by the average stimulus distance from the hand (21 cm). As with the slope, the most significant effect of the training took place during the first correct-answer feedback session in which percentage bias decreased in magnitude from approximately 16% to 1%. The percentage bias remained under 1% for the other training stages.

Figure 3c shows how standard deviation varied by actual object distance. Within each group is the standard deviation for each training stage. The solid line joins the average value of each group, while the dotted line is the average value of the group, excluding the initial training stage. Average standard deviations in each grouping increases with object distance between the first object position until it reaches its peak at the middle object distance, and then decreases with object distance with the remaining stimuli. This trend exists regardless of whether

304
the initial training stage is included in the average. The effect of the baseline experiment merely increases the average by approximately 0.5cm.

The effects of training are more apparent in Figure 3d, which groups the standard deviations by object distance and is plotted as a function of training session. The average standard deviation starts high, approximately 7.5cm, dips considerably after the first feedback session to approximately 4.5cm, then settles at approximately 3.75cm for the feedback sessions. Based on these results, the training of subjects for the remaining experiments consisted of only a single correct-answer feedback session separating the initial and test phases. While the additional feedback sessions led to improvement in subject performance in terms of accuracy and consistency, the change was not significant enough to warrant the extra training.

3. Experiment 2 – Virtual Object (true force and torque feedback; unrestricted probing)

The same setup as Experiment 1 was used here with the exception that the object whose location was to be determined was a virtual thin plate whose top edge was in the same orientation as the real edge of Experiment 1.

Software: In addition to the functions performed in Experiment 1, the software program used to conduct Experiment 2 also provided control of the virtual environment. This includes collision detection between the probe and virtual object, control of force output for both Phantom haptic interfaces, and random positioning of the virtual objects. The collision detection model for this virtual environment, which consists of a single virtual object and the probe, is shown in Figure 4. The magnitude of the reaction force due to contact between the probe and virtual object is computed from Hooke’s Law, $R = ky$, where $k$ is the stiffness of the object and $y$ is the vertical depth of penetration of the line segment past the top of the plate. In Experiment 2, the two Phantoms were programmed to apply forces $F_1$ and $F_2$ such that the resultant force on the user’s hand was equal to $R$ and the

![Figure 4: True force and torque feedback](image)

![Figure 5: Results for the virtual object with true force and torque feedback](image)

resultant torque on the user’s hand was equal to the cross product of moment arm $z$ with $R$.

$$F_1 + F_2 = R$$

$$z \times (F_1) + (z \times F_2) = z \times R$$

The simplicity of this virtual environment enabled us to achieve high rendering rates of up to 15000 kHz. However, due to the length of the probe and limitations of the encoder resolutions in the haptic interface, some vibrations were felt through the probe during prolonged contact with the virtual object. To minimize these vibrations, the stiffness of the virtual plate ($k = 6.4N/mm$) was optimized according to the experimenter’s judgment.

Training: In this experiment, subjects were trained as in Experiment 1. However, subjects completed only one session with correct-answer feedback rather than four sessions as in Experiment 1, whose results indicated that the additional correct-answer feedback sessions produced minimal improvement.

Procedure: The same procedure as Experiment 1 was followed here. While subjects were permitted to use any exploration method(s) they chose, observations by the experimenter and comments from the subjects indicated that each of the subjects chose to focus only on one technique. Three of the subjects practiced a rocking method in which the object was used as a fulcrum (pivot point) and the remaining three subjects employed a tapping method. Performance in each subgroup was comparable.

Subjects: Six of the eight subjects who participated in Experiment 1 took part in this experiment.

Results: Figure 5 shows the plot of subject response versus actual position of the virtual object for the initial and final stages of training for Experiment 2, in which subjects were presented with both true force and true torque feedback. Again, perception of object distance in the initial set, prior to correct-answer feedback, was not
very accurate as seen in the 0.73 slope, 3.82cm average standard deviation, and 3.26cm bias of Line A. The final set following training resulted in a slope of .95 (line B), which implies that for objects placed 4cm apart, on the average, a 3.80cm separation was perceived. The standard deviation decreased to 3.50cm, suggesting that subjects were more consistent in their responses following the correct-answer feedback sessions. The bias of .09cm indicates that there was an average underestimation of less than a millimeter for each object distance. These results demonstrate the effectiveness of ray-based rendering and the parallel configuration of the two force-feedback devices as haptic perception in the virtual environment was at least as good as haptic perception in the physical world. From Figures 6a and 6b, it is seen that slope and percentage bias improved for each stage of training. Figures 6c and 6d show the same trend for the average standard deviation as seen in Experiment 1: minimal standard deviation at the extreme locations of the object and a maximum peak at an intermediate distance.

4. Experiment 3 – Virtual Object (force reflection at probe tip only; probing restricted to rocking)

The same setup as Experiment 2 was used here. **Software:** With the exception of the force display, the same software program was used for Experiment 3 as in Experiment 2. In this experiment, although the setup included both Phantoms connected by a probe, only the single Phantom at the front end of the probe provided force feedback. Thus, contact between the probe and the virtual object resulted in correct force reflection but the torque feedback experienced by the user was physically inaccurate.

5. Experiment 4 – Virtual Object (force reflection at probe tip only; probing restricted to tapping)

The same setup as in Experiments 2 and 3 was used here. **Software:** The same program used in Experiment 3 was used in this experiment.
Figure 8: Results for the virtual object with force reflection only at the probe tip and tapping

**Training:** The same three training stages used in Experiments 2 and 3 were used here. The initial training stage for this experiment was taken to be the same as that of Experiment 3.

**Procedure:** The procedure for Experiment 4 was the same as that for Experiment 3 with the following exception: in the correct-answer feedback and test sessions, the method of exploration for was restricted to tapping against the object such that the probe always struck the object with the same orientation.

**Subjects:** Five of the six subjects who participated in Experiments 2 and 3 took part in this experiment.

**Results:** The plot of the initial training stage for Experiment 3 is reproduced along with the test stage for this experiment in Figure 8. Unlike the previous training experiments, the correct-answer feedback session did not result in significant improvements in subject perception. The slope of 0.29 for the final set (Line B) is comparable to initial performance (Line A). The bias improved to 0.48 cm. On the average, responses were distributed by an even larger margin during this session than in the initial stage as indicated by the average standard deviation of 6.33 cm. Although the bias was largely corrected through training, the slope remained very poor and the standard deviation did not vary a great deal as the object distance increased. These results imply that when restricted to tapping, torque feedback is a necessary haptic cue for best performance, since in Experiment 2 (true force and torque feedback), subjects were able to accurately discern the different object positions when only tapping. The required haptic information to judge the distance of the objects was not adequately presented when tapping with force-feedback only at the tip. In spite of the wrong torque feedback, the presence of a slight slope in lines A and B of Figure 8 is possibly due to the presence of kinematic cues (see Discussion).

### 6. Measure of Resolution

The just noticeable difference (JND) between two object positions describes the smallest distance separating the two objects such that each one can be distinguished on a consistent basis, for example 70% correct discrimination may be used as the minimum requirement for consistent accurate perception. Typically, JND is computed from the results of pair-wise discrimination tests. While the results obtained in the previous section are taken from identification experiments, it is possible to compute a measure of haptic perceptual resolution of distance based on signal detection theory and a decision model for the one-interval, 2AFC (two alternative, forced choice) paradigm (Durlach, 1968). Due to the imprecision caused by attempting to measure perception of a continuous random variable (distance) with a finite number of discrete object positions, the values reported here are merely estimates of JND for the specific conditions examined. It is likely that these values can be taken as the upper bounds for JND. Perception of the different object positions is expected to be better for pair-wise discrimination experiments since the number of possible responses is reduced. In order to apply the decision model for a one-interval, 2AFC paradigm, each interval must be examined independently. Each of the 7 object positions can be viewed as a reference position with which the remaining 6 positions are compared. For this study, however, only neighboring object positions were analyzed. The sensitivity index d', measures the separation of the probability density functions for two stimuli (i.e. d' indicates how well the two different object positions could be distinguished). For this study, d' was estimated by the difference in the means for two neighboring object positions and divided by the average of the standard deviations. JND is defined as the distance between two stimuli at which d' has the value of 1.

The table below shows how %JND varies by mean stimulus distance relative to the hand. Weber’s Law states that percentage JND will remain constant and independent of variations in the reference stimuli (i.e. change in object distance in our case). For the real object, true force and torque feedback, and rocking under tip force conditions, we see that this is not the case. Instead, excluding the first stimulus interval distance, we have a fairly linear relation, in which the percentage JND is inversely related to the mean stimulus distance from the hand. Knowledge of the 7 discrete object distances, gained from the correct-answer feedback sessions, required subjects only to bin their judgments into one of seven categories rather than responding with any other distance in the continuous range of positions.
Table 1: Estimated %JNDS for each condition

<table>
<thead>
<tr>
<th>Force Display</th>
<th>Interval</th>
<th>Int. 2</th>
<th>Int. 3</th>
<th>Int. 4</th>
<th>Int. 5</th>
<th>Int. 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt. 1 Real Object</td>
<td>21.6</td>
<td>22.8</td>
<td>18.9</td>
<td>15.3</td>
<td>13.0</td>
<td>11.5</td>
</tr>
<tr>
<td>Expt. 2 True Force and Torque</td>
<td>17.7</td>
<td>21.2</td>
<td>20.5</td>
<td>17.7</td>
<td>13.9</td>
<td>11.9</td>
</tr>
<tr>
<td>Expt. 3 Tip Force: Rocking</td>
<td>21.4</td>
<td>22.7</td>
<td>18.1</td>
<td>14.3</td>
<td>13.5</td>
<td>11.5</td>
</tr>
<tr>
<td>Expt. 4 Tip Force: Tapping</td>
<td>36.3</td>
<td>31.1</td>
<td>26.9</td>
<td>26.2</td>
<td>29.4</td>
<td>27.3</td>
</tr>
</tbody>
</table>

Therefore, errors in perception will not result in slight variations in response, but correspond to immediate neighboring positions. For gross errors, the responses may not necessarily correspond to adjacent object positions to the actual stimulus but those farther away. This is likely the case in Experiment 4 for the tapping under tip force feedback condition. When the object to be located is closer to the subject’s hand, %JND is high (over 35%), and settles at approximately 27% for interval distances from 19cm to 31cm. Clearly, regardless of the distance of two stimuli from the hand, %JND is much higher for this experimental condition than in the others. The reason that they could discriminate at all with the wrong torque feedback may lie in the kinematic cues that are available in addition to force and torque cues (see below).

7. Discussion

One might expect torque information to play a vital role in the haptic perception of object location. Indeed, examining the static force diagram of a probe in contact with an object (Figure 9a), one can easily calculate the moment arm, d, if the reaction force, R, and the torque, T, are known.

\[ d = \frac{T}{R} \]

But the distance of the object from the hand, d, is equal to the magnitude of the moment arm, l d l, since the moment arm is the distance over which R acts to create a torque, T. For this approach to finding d, knowledge of the value of |T| is required. This may have been the method by which subjects in the tapping subgroup in Experiment 2 were able to accurately identify the various object positions. This approach, however, does not explain how subjects were able to perceive the different object positions in Experiment 3 when the magnitude of the torque did not agree with the position of the object.

Therefore, there must be another method by which object distance is perceived that is not dependent solely on force and torque feedback. If we take a kinematic rather than kinetic approach to finding d, knowing torque is not necessary. Below is one such computational theory for determining d based on geometry by examining the intersection of two or more probe orientations at time of contact with the object (Figure 9b). However, since the rod and object are hidden from view of the subject, this approach is dependent upon the subject’s haptic sense of probe orientation, which is possibly related to awareness of positioning of the arm and hand. The magnitude of the force felt at the moment of contact is not important. The force itself is only important in that it indicates contact has been made at that moment with the probe held in a particular orientation. If at the next moment of contact, the probe is in a different orientation, the position of the object can be determined by locating the point at which the two probe orientations intersect.

A special case exists when using the object as a fulcrum and rocking the probe against it (Experiment 3). This provides a continuous change in orientation for a fixed pivot point. Therefore, all the orientations would intersect at the same point. A second, but similar special case, requires knowledge of only a single orientation of the probe during contact with the object. The height of the object was constant throughout all trials and experiments, though subjects were not informed of this fact. If a subject hypothesized that the height of all objects was the same, then a horizontal line could be
projected on which the object must lie. By contacting the object with a single non-horizontal orientation of the probe, the location of the object can be determined from the intersection of the horizontal line with the orientation of the probe at the time of contact. The subject, then, needs to have a sense of the horizontal line and therefore, the height of the object, as well as the orientation of the probe. Geometrically, this is the same as the case presented above with one of the contact angles equal to zero. Following Experiment 4, several subjects commented that the height of the object seemed to be changing from trial to trial. So the subjects may have been attempting to use this approach in the absence of other cues, but were unable to perform as well as in other conditions, since they perceived the height of the object to be shifting between trials.

For any application in which the interaction of the entire stylus/tool with the environment is of interest, the collision detection algorithm must include contact between objects and any point along the stylus, such as in the ray-based rendering technique. If not, as in the point-based rendering employed by most of the haptic rendering today that detects only contact with the tool tip, the stylus can unrealistically slice through a virtual object without reflecting any forces to the user. Determining the hardware requirements for accurate haptic perception in virtual environments is not so straightforward. As one might expect, it will depend on several factors such as the goals of the task, the force and torque requirements, the kinematic degrees of freedom, and the workspace available to the user. The results described in this paper show that at least in probing for object location, there is a trade-off between whether torque reflection is present or whether rotational motion is allowed. In cases in which the user is free to move the stylus/tool through any motion and orientation, a single 3 DOF haptic interface is sufficient to present the haptic cues for locating the objects in the environment. This is because the information needed to locate the object can be obtained from the geometry of the situation. Torque feedback becomes a redundant cue provided that the user is permitted to change the orientation of the probe. Virtual sculpting is one such application. If the sculptor is to have the freedom to approach the piece from any direction, he must have the ability to manipulate a sculpting tool through a wide range of motions and orientations. This freedom, while possibly expensive in terms of manipulator workspace, eliminates the need for more than 3 DOF force feedback.

If, however, constraints are placed upon the motion of the stylus, such as in a laparoscopic surgical simulator, both force and torque feedback may be needed to obtain best judgments of the distances of the objects that come into contact along the length of the instrument. In minimally invasive procedures on real patients, the motion and orientation of the instrument is restricted by the trocar port through which it must pass. Because of the compliance of tissues around the trocar, the trocar does not act as a perfect hinge, but constraints the translation along and rotations about the axis perpendicular to the tool. The consequent reduction in the range over which the tool orientation can be varied would lead us to expect that torque feedback may be of importance in the real world surgical task. Therefore, if accurate judgment of distances from the hand to tool-tissue contacts along the length of the tool is important for the surgical training task in VE, a 6 DOF haptic interface, or a parallel configuration of two 3 DOF devices as described in this paper, would be required. This hypothesis needs to be tested.

8. References


