Application-based Evaluation of Haptic Interfaces

Arthur E. Kirkpatrick
Simon Fraser University
ted@sfu.ca

Sarah A. Douglas
University of Oregon
douglas@cs.uoregon.edu

Abstract

A taxonomy of haptic modes is proposed as a basis for evaluating haptic interfaces. Haptic modes are defined as distinct styles of using haptic perception for purposeful activity. An evaluation method that exercises a single haptic mode tests the hardware and software of the interface on a task that has a clear relationship to actual applications. To demonstrate this approach to evaluation, the mode of geometric perception was evaluated for a point force device. Twelve observers performed a shape recognition task using a PHANToM display. The task proved difficult, with a mean recognition time of 22 s. This class of devices does not appear to support adequate performance for applications that require geometric perception. The paper concludes with possible extensions to both the shape recognition protocol and the taxonomy of modes.

1. Introduction

While a wide diversity of haptic interfaces has been developed, a much smaller number of evaluation methods has been proposed. This restricts the effectiveness of our designs because evaluation coerces designs in positive ways. Evaluation methods separate good designs from bad. Evaluation theories can do even more, suggesting effective directions for new designs.

In this paper, we propose an application-driven approach to evaluation. This is a higher level of consideration than previous approaches, which have focused on specific device characteristics, such as stiffness [1] or psychophysics [2]. Application-driven methods evaluate the whole interface, including the software interaction technique controlling how the device is used. They also evaluate the interface at a level much closer to the level at which the user will use it. A haptic device ultimately only renders sensations—changes in force, vibrations, displacements of the fingerpad—whereas a user is concerned with the level of percepts—shapes, textures, objects, and relationships between objects. A focus on applications implies a focus on the percepts that are important in different applications.

As an example of the different percepts of interest in different applications, consider three applications that are frequently proposed for haptic interfaces: simulated mechanical disassembly, analysis of multivariate datasets, and surgical simulation. In each of these cases, the user is performing different styles of movement and attending to different percepts. For example, the engineer disassembling a mechanism begins by reaching for a simulated part, then moves it along a path, orienting it to avoid collision with other parts. The engineer is attending to the object only insofar as it collides with other objects and is unconcerned with its texture or compliance.

This is very different from a seismic engineer exploring an oilfield dataset. The data might be displayed as a single object, with different dimensions of the data mapped to different visual and haptic parameters. The engineer can rotate the object, view it, and explore the haptic properties using a haptic interface. In this case, the engineer may be paying detailed attention to the overall shape of the dataset, detailed features within that shape, the texture and compliance of regions of the data, and other properties that have been mapped to dimensions of the underlying data. Neither selecting an object nor orienting it relative to other objects will be considerations.

The surgeon practicing a surgical procedure will likely use a different constellation of skills. Initially exploring the body region, she may attend to the compliance of various tissues. Inserting the surgical implement requires precise control of force and direction. Reviewing the results of the procedure can require precise knowledge of specific features such as sutures and their location with respect to the overall organ. Of all these applications, surgery likely requires the most comprehensive set of haptic percepts.

It is generally agreed in the haptics community that different applications will require different devices. However, the link between applications and devices, describing why a device may or may not be suited to an application, has not to our knowledge been precisely articulated. We propose to specify this link in terms of the "haptic modes" each application uses. We can then evaluate the suitability of a haptic interface for an
application by evaluating the interface for each individual mode and combination of modes that the application requires.

We begin the paper by describing different ways in which the haptic system may be used. It is a central claim of this paper that there are a small number of these haptic modes, ways of using the haptic system for goal-oriented purposes. We describe five of these modes, deriving our definitions from evidence in the psychological literature. Haptic modes provide a parsimonious scheme for classifying evaluation methods according to the haptic modes that they exercise.

As an example of this approach and a contribution in its own right, in the second half of the paper we evaluate a haptic interface. We measure the performance of the predominant class of commercially produced haptic devices using a task that exercises the haptic mode of geometric perception. Basing the evaluation on a haptic mode allows us to generalize the results to which applications to which it will apply. Of the three applications considered in this paper, the results apply most strongly to multivariate data analysis, for which the geometric mode is a particularly strong component. The results will have fewer implications for surgical simulation and apply least of all to mechanical disassembly. We conclude the paper with suggestions for haptic interface designs that might offer improved performance for the geometric mode, possible extensions to our geometric mode evaluation method, and possible extensions to our taxonomy of modes.

2. Evaluating haptic interfaces

Most metrics of human performance with haptic interfaces have focused on the hardware, the haptic device. For example, Hayward and Astley [1] propose a standard set of device metrics such as degrees of freedom, motion range, peak force, inertia, acceleration, resolution, precision, and bandwidth. These metrics capture essential performance characteristics of the device.

However these metrics are at too low a level to characterize the performance of the interface. First of all, the experience of the human using the system is described in such terms as objects, space, locations in space, shapes, features on those shapes, parts, an object's function, its material, and its texture. Users are unconcerned with the device resolution or its degrees of freedom. For them, the usefulness of an interface is best measured in terms of how readily they can perceive and manipulate the objects simulated by the interface. Device metrics cannot capture these qualities.

Secondly, the interface comprises the device and the interaction technique, not just the device. Users execute interaction techniques, not devices, and the visual feedback provided by the interaction technique crucially determines the performance of the interface. Consider drawing a straight line with a mouse. While in principle it is possible to draw a perfectly straight line using an interaction technique for drawing an arbitrary curved path, the process is so awkward that drawing programs instead provide distinct interaction techniques specifically for lines. The interaction technique determines whether the mouse is a poor device or a good device for line drawing. Similarly, the haptic interaction technique will ultimately determine the performance of a haptic device. Device metrics thus cannot assess performance of an interface.

We believe that the best metrics for haptic interfaces measure the performance of an interaction technique in terms of percepts meaningful to the human user. This implies that the metrics are determined by the application, since both the interaction technique and the percept are potentially specific to that. The range of such applications is infinite. However, we propose that all applications are based upon a small set of distinct modes of using the haptic system.
2.1. Haptic modes

We define a haptic mode as a distinct style of using the haptic system, characterized by the nature of the user's attention, the path and duration of any movement, and the skin location contacting the object of interest. If these factors were fully independent of one another, they would produce a highly multidimensional space. However the psychological literature suggests that these factors are in fact closely linked. The styles in which humans use their haptic system cohere into a few distinct modes rather than spreading over a continuously varying space.

Our taxonomy of modes incorporates elements of previous taxonomies but is distinctly broader. We aim to characterize the entire range of possible applications of haptic interfaces. No existing taxonomy covers that range. For example, in one of the standard handbooks of applied psychology, discussions of kinesthesia [5] and tactual perception [6] are in completely separate chapters and do not establish a common taxonomy. However, any description of applications of haptic interfaces must incorporate both aspects.

Our taxonomy of haptic modes is diagrammed in Figure 1. For the sake of exposition, we assume that the user's hand or hands are being used to manipulate the object of interest. The taxonomy can be extended easily enough to account for the use of other limbs if necessary.

At the highest level, we distinguish between using the haptic system for perception or motor control. The discriminating characteristic at this level is whether the system is being used in the subjective or objective pole [6, p. 31-4]. When the observer is using the haptic system for perception they report their haptic experience in terms of the object: "The object feels smooth" (objective pole). By contrast, when the observer is using the haptic system for motor control their haptic experience is reported in terms of their own hand posture: "I have made contact with the object" (subjective pole). An equivalent distinction can be made in terms of the user's intention: For perceptual tasks the user intends to learn the properties of an object whereas for motor tasks the user intends to modify the object's location or structure.

The motor control mode is further subdivided into object manipulation and target acquisition submodes. These modes are distinguished by their duration and by the way in which haptic sensations feed back to the motor control system. In object manipulation, haptic perception guides motor skills used to transform an object. Object manipulation tasks have indefinite duration. For these tasks the haptic system is used to modulate the degree of grip force [7] to ensure the object is not dropped. A secondary use is to detect collisions of the held object with other objects. There are various characterizations of the movements [8] and hand postures [9] used to manipulate objects.

Target acquisition tasks are the classic action of rapid aimed movement whose movement time is described by Fitts' law. These tasks have a finite duration, completing as soon as the hand has successfully acquired the target. The main role of haptic perception is to provide confirmation of target acquisition.

We divide the perceptual mode into three submodes: geometry apprehension, material apprehension, and environmental monitoring. The distinction between the apprehension of geometric or material properties is due to Klatzky and Lederman [10]. Geometric properties are the arrangement of the object's parts in space. They include the size and dimensions of a bounding volume for the object, the geometry of localized features such as bumps or dents, and the arrangement of those local features in the overall structure of the object. Material properties are properties of the constituent substance of an object. They include compliance, texture, weight, and other properties. Klatzky and Lederman [11] demonstrated a close correspondence between the properties apprehended by the user and classes of stereotyped hand movements, called exploratory procedures (EPs), the user performed. Motoric constraints thus distinguish these two haptic modes, compelling the user to choose between precise apprehension of one or the other set of properties and even between specific properties (e.g., texture vs. compliance) within those broader categories.

The final submode, environmental monitoring, represents the ongoing background monitoring of the objects in contact with the skin. It is distinguished from the other modes by the lack of attention it requires. The observer allocates no attention to the monitoring but becomes aware of a change in the tactual sensations when it occurs. This is also the only mode in our taxonomy that can be performed with any portion of the skin surface. In fact, given the need to use the hand surface for other haptic modes, the monitoring mode is rarely performed with the hand.

The above taxonomy attempts to be complete for strongly goal-oriented uses of the haptic system. It does not cover less directed uses such as stroking a kitten's fur. Our goal is to facilitate the development of interface evaluation techniques. We expect that completely different evaluation procedures will be required for these less goal-directed uses and so do not include them in our taxonomy.

2.2. Haptic modes in applications

The above taxonomy of haptic modes distinguishes the ways in which the haptic system might be used. Different applications will require different haptic modes for their successful completion. Experimental tasks are more specialized, typically testing one or perhaps two modes. At the lowest level of Figure 1, we list some tasks that
directly exercise single modes of the taxonomy. Explicitly locating the common haptic modes of applications and experimental tasks establishes a basis upon which results of controlled evaluations can be generalized to applications.

This taxonomy provides an organizing structure for performance estimates for haptic interfaces. Most previous studies of haptic interface performance have used experimental tasks that evaluated haptic interface performance in motor control mode [e.g., 12, 13] while the perceptual mode of the taxonomy has remained relatively unexplored. Only three studies have measured perceptual mode performance of haptic interfaces. Tan [14] found that users of the PHANToM haptic device could haptically identify between three and four distinct sphere sizes. Ming [summarized in 15] found a 25% improvement (compared to a visual-only interface) in one phase of a molecular docking simulation. Lederman and Klatzky [2] measured a variety of psychophysical parameters under conditions approximating a haptic interface.

None of these studies has measured the performance of a haptic interface on a task requiring the geometric mode (Tan's [14] participants knew the shapes were all spheres). We have seen how two important applications, data analysis and surgical simulation, could benefit from effective use of the geometric mode of haptic perception. Shape recognition is an experimental task of pure geometric perception. It is therefore an appropriate task to assess the effectiveness of a haptic interface for the geometric mode.

3. Evaluating performance of haptic geometric perception

As a specific example of applying this taxonomy, we developed a method for evaluating haptic interface performance of the geometric mode using shape recognition. We applied this protocol to the SensAble PHANToM. This device exemplifies the class of point force devices, haptic displays that provide only a single point of contact. Nearly all commercially available haptic devices, from two-dimensional haptic mice and joysticks to three-dimensional devices such as the PHANToM, are point force devices. Point force designs are widespread because they offer an appealing balance between mechanical simplicity and high display resolution. However, this does not guarantee they support the same level of perceptual performance found in haptic exploration of real objects. To achieve that performance point force devices must support the same or similar perceptual processes as perception of real objects.

3.1. Haptic interface structure and human performance

The perceptual processes supported by a haptic interface are a consequence of the structure of both the device and the cues provided by the interaction technique. The device structure—whether force display, tactile display, or both; number of contact points; point versus area display; and so forth—sharply defines the stimuli that can be presented to the human user. The structure of the interaction technique—the user's limb movement, the changes to the haptic display in accordance with those movements, and any visual feedback—determines how the user will move and how effectively she or he can complete the task.

Of the device properties, the number of contact points has particular relevance to performance of geometric tasks. Klatzky, Loomis, Lederman, Wake, and Fujita [16] demonstrated that observers performing a (haptics only) shape recognition task with just their index finger required 45 s, whereas observers using their whole hands only needed 16 s. Klatzky et al. attribute the difference in performance between full-hand haptics and index-finger haptics to how information is presented to the observers in each condition. In full-hand haptics, the initial enclosure provides an overall map of the object and the more specialized hand movements that follow the enclosure afford multiple points of contact with specific local object features. Using these movements, the observer can construct a representation of the object from multiple data points at the same time, a process of spatial integration. By contrast, in index-finger haptics, the observer must perform temporal integration, constructing a representation of the object from a sequence of single data points presented over time and without an overall map from an initial enclosure.

The low performance of the observers in the index-finger condition demonstrates that temporal integration is a slow, error-prone process. Haptic displays that provide only a single point of contact compel the user to perform temporal integration when using the haptic mode of geometric perception.

Effects such as temporal integration are consequences of the entire interface, not just the device. While the structure of the physical device sets intrinsic limits on the haptic component of geometric perception, the interaction technique might provide visual feedback that averts or mitigates this effect.

For example, temporal integration in shape recognition can be trivially eliminated by graphically displaying the shape. In such a case, the user will simply use vision to determine the shape and not use haptics at all [16]. However, since our interest in shape recognition is as an experimental task that exercises the haptic mode of geometric perception, this "solution" is useless.

A more interesting possibility is to provide the
absolute minimum visual feedback and see if any mitigation occurs. Vision is arguably a more spatially oriented sense than haptics and it may be easier to construct a representation of the object's shape from seeing sequential locations of a visual cursor over time than from feeling sequential displacements of a surface. Displaying a fixed reference point in addition to the cursor provides even more cues, as it establishes an underlying coordinate system for assessing the cursor's movement over time. If temporal integration limits are specific to the haptic system, adding a visual cursor and a reference point will improve performance, whereas if temporal integration is limited by some more central bottleneck visual cues will provide no improvement.

3.2. Experimental design

The primary aim of our protocol was to evaluate how well point force devices support the geometric mode of haptics. To measure pure haptic performance, our first experimental condition was haptics-only, where participants could feel the shapes but had no graphical display whatsoever. To ensure that participants could not "cheat" by looking at their hand position, we concealed their hand behind a curtain.

As noted above, the presence of graphical display can dramatically change the way haptics is used. The nature of that change appears to be influenced by the extent of graphical display. A secondary aim of our study was to determine if a minimal amount of graphical display would improve performance. To test this, in the second experimental condition we provided a graphical cursor indicating the current location of the PHANToM tip. A reference dot was also displayed at the center point of the shape. Using the cursor and reference dot, participants could visually estimate their current location relative to the center of the shape. We were interested in measuring whether this extra information mitigated temporal integration and enhanced performance.

To test the effects of spatial resolution on temporal integration, we made shape size a second experimental factor. Participants were presented two sizes of stimuli, spanning distances of approximately 1.5 cm and 7 cm.

We drew our stimuli from the class of smooth-flowing three-dimensional shapes defined by Koenderink and used in the shape recognition task of Kappers et al. [17]. These shapes are constructed from two orthogonal parabolas. A shape scale, computed from the direction of curvature of the two parabolas identifies the shapes. There are five critical points on the scale, named Cup, Rut, Saddle, Ridge, and Cap (see Figure 2). The shapes lying at these points are distinguished from one another by the signs of the curvatures of their constituent parabolas. An observer can distinguish them simply by determining whether each parabola is curving up or down or (in the case of Groove and Ridge) is flat.

These five shapes represent an excellent base for a standard set. They are smoothly curving and can be readily distinguished by assessing direction of curvature without assessing its degree. They are a canonical set in the following sense: Any more complex solid shape can be constructed from a combination of these shapes. Performance of users recognizing these shapes constitutes a baseline performance for shape recognition.

To increase diversity of stimuli, three different rotations were presented for each one. All factors and stimuli were fully crossed, giving a 2 by 2 by 3 within-subjects design. There were a total of 30 trials per cursor condition and 60 total trials for the experiment.

3.3. Method

The protocol had a separate training phase and testing phase. In the training phase, participants were told the names of the five shapes and felt them with the PHANToM. Once they had felt every shape twice, they were asked to identify the five shapes when presented in random order. In this phase, the shapes were presented "head on" (unrotated) and in a size that spanned approximately 3 cm of movement of the PHANToM. When a participant could recognize the shapes perfectly for two consecutive blocks of five, the testing phase began.

The three rotations moved the shapes obliquely away from the head on configuration used in the training phase but were small enough that the front of the shape remained facing the user. The rotations were simply used to provide variety of stimuli and the effects of this factor were not analyzed.

The experiment was run on a 300-Mhz Pentium II under Windows NT 4.0. The haptic display was a PHANToM 1.0A and the visual display device was a color screen. The haptic rendering loop ran at 1000 Hz
and consumed approximately 30% of the processor time. For each trial, the environment recorded the time from first contact of the PHANToM tip with the shape until the participant ended the trial by pressing the space bar.

The participants were 12 unpaid graduate students from a computer science department. The choice of participants with strong technical backgrounds was deliberate. In our pilot studies and open houses we have observed that users with less technical experience often take time to learn such basic skills as indirect three-dimensional pointing and associating the force display with the visual cursor. We used computer science graduate students to minimize the confounding effects of learning these basic skills. Nine participants were male and three female. Their ages ranged from 22 to 42 with a median of 31. Ten were right handed and two left handed. All used the PHANToM with their dominant hand.

Six trials were observed to be incorrect while the experiment was proceeding. For example, participants sometimes inadvertently terminated the trial before they had determined the shape. These trials were deleted before the data analysis. No further editing was performed.

4. Results

Participants took quite a while to recognize the shapes. Arithmetic mean trial time was 28.6 s, with a geometric mean of 22.5 s and an interquartile range of 13.7 to 37.7 s. These long times indicate that temporal integration has as much of an effect in haptic interfaces as it does in working with physical objects. Point force devices do not appear to effectively support the geometric mode of haptics. Quantile-quantile plots showed that the distribution of times was lognormal, so the most representative estimate of trial times is the geometric mean.

There was also a large range of individual differences. The geometric means of the participant times ranged from 13.7 to 42.7 s, a factor of 3.1.

While participants took a long time on the task, they were able to recognize the shapes reasonably well. The mean score for participant accuracy was 84% (s.d. 12%), well above the chance performance level of 20%.

A two-way within-subjects ANOVA of main effects and interactions for log of time was computed for the mean time for each participant in each condition. An alpha level of .05 was used for all tests. Table 1 shows the ANOVA results and estimates of the effect size. The effect of cursor condition was both unreliable in direction and small. Providing a small amount of visual feedback did not appreciably change performance. The graphical cursor did not mitigate temporal integration.

Small shapes had significantly longer recognition times (24 s) than large shapes (21 s), although the size of the increase was only 12%. This suggests that the finer haptic spatial resolution required for smaller shapes modestly increased the difficulty of recognition. The interaction effect between size and cursor condition was not significant.

A similar ANOVA was performed for accuracy of recognition. No significant effects were found for cursor condition, shape size, or their interaction.

4.1. Analysis of trial paths

In addition to the total trial time, the experimental software recorded the current location of the PHANToM tip every time the graphics loop was executed. The resulting sample intervals ranged from 10 ms to 61 ms, with most in the 10-20 ms range.

Three-dimensional plots of the paths traced around the shapes revealed great diversity. They differed between individuals and between trials for the same individual. There was no straightforward correlation between the path and the shape being recognized. However, there was a striking similarity between large and small shapes: Participants traveled similar distances for both sizes of shapes. This provides further evidence that participants based their judgements upon absolute changes in distance rather than rate of curvature.

To reduce the dimensionality of the data to a manageable level, the 1-dimensional velocity (i.e., speed) was computed for each second of every trial. First, the Euclidean distance was computed between every pair of sample points. Then the data points for each trial were collected into the smallest groups spanning at least a second. The resulting groups ranged in duration from 1.00 s to 1.06 s. The velocity for each group was then computed by dividing the sum of the sample distances for each group by the duration. Finally, the normalized time profile of the velocity was computed by dividing the timestamp of each 1-second group by the total trial time. The result was a profile with velocity as the dependent variable and normalized time, ranging from 0 at trial start to 1 at trial end, as the independent variable.

The velocity profiles increased significantly as the trial proceeded. As with total trial times, velocities were distributed lognormally. For each user, the mean velocity

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>Effect size</th>
<th>Lower 95 % conf. int.</th>
<th>Upper 95% conf. int.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cursor</td>
<td>1</td>
<td>0.10</td>
<td>.76</td>
<td>N.S.</td>
<td>-14%</td>
<td>16%</td>
</tr>
<tr>
<td>Size</td>
<td>1</td>
<td>6.99</td>
<td>.02</td>
<td>12%</td>
<td>2%</td>
<td>22%</td>
</tr>
<tr>
<td>Cursor X size</td>
<td>1</td>
<td>2.79</td>
<td>.12</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
was computed for the first and second halves of each trial. The geometric mean of the velocity increased from 15.5 mm/s to 17.6 mm/s between the two halves of the trial \((t = -3.50, df = 11, p = .005,\) performed on the log of time). Users appear to have performed the task in three distinct phases. Figure 3 shows a loess plot of the velocity against normalized time, computed from all the 1-second velocities for all users. The first 25% of the trial shows the velocity sharply increasing, the middle period is characterized by an essentially constant velocity, and the final 25% shows the velocity increasing sharply.

There are many possible explanations for the shape of this curve. An intriguing interpretation is that the phases correspond to distinct phases of a generate-and-test strategy: First, generating a list of possible shapes (alternatively, ruling out some of the possible shapes), then selecting one shape from that list, and finally confirming that selection. We intend to investigate this explanation with further analysis of this data.

5. Discussion

The most striking result of this study is the difficulty of the task. Despite the simplicity of our stimuli and task, participants still had a geometric mean time of 22.5 seconds with a 15% error rate.

This level of performance is clearly unacceptable for many applications. The minimal visual feedback provided by our cursor condition did not provide nearly as large an effect as we need to reduce shape recognition times to a practical level. Temporal integration appears to be a problem for vision as well as haptics. It may not be possible to mitigate the effects of temporal integration even by using an interaction technique with more elaborate visual cues.

However, we note that all participants’ performance continued to improve up until the 60th trial. They do not appear to have reached skilled performance during the course of our experiment. Some participants reported that they found the visual cursor condition distracting. Many had their hands full merely attending to the haptic sensations. We speculate that the sensory overload might reduce with practice. The visual cursor might have a larger (and significant) effect when participants achieve practiced performance.

A second possible approach to improving performance is increasing the number of points of contact between the user and a displayed shape. As noted earlier, Klaczky et al. [16] found this produced a 180% improvement in performance recognizing physical shapes with the hands. If a comparable improvement could be produced by increasing the contact points of haptic display in our task, response times for our stimuli would drop to around 8 seconds—a gain of great practical consequence for those applications emphasizing the geometric mode. However, devices that can display shape information to that many fingers are inherently more complex and suffer from their own limitations. Adding more contact points is neither a cheap nor a guaranteed solution.

6. Conclusions and future work

We have argued that haptic interface evaluation should be driven by the haptic modes required by various applications. As an example of this approach, we developed a shape recognition protocol to test performance of haptic interfaces for the haptic mode of geometric perception.

The protocol can be used as a benchmark task to evaluate new haptic interface designs, in the same way that Fitts tasks are used to evaluate two-dimensional pointing devices. Furthermore, the results of our study suggest an important direction for future work. Understanding the various factors underlying the poor performance shown in this study is crucial to the usability of environments incorporating point force devices. In the future, we intend to study how long it takes individuals to reach skilled performance, what the level of that performance is, and what factors might facilitate the geometric mode with a point force device for different user populations.

This sample evaluation also demonstrates how task-based evaluation can produce useful, generalizable data on haptic interfaces. At each step, the taxonomy suggests where to look and how far the results generalize. The taxonomy of haptic modes identified geometric perception as a stable, distinct mode of the haptic system. Data analysis and surgical simulation use this mode extensively and hence the results of the evaluation apply to them.

The taxonomy also indicates which applications will not be affected by this performance. The shape recognition evaluation does not test the performance of point force devices for the environmental monitoring, material perception, object manipulation, or target
acquisition modes. These devices might provide excellent support for applied tasks that predominantly use these modes. We note that the most successful commercial application of point force devices, SensAble's FreeForm product, uses the PHANToM exclusively for object manipulation, "sculpting" a model with a "virtual knife".

There are many possible extensions to the taxonomy. The categories of modes and tasks are still a first cut and can be refined. In particular, the distinction between perception and motor control is not hard and fast. How much perception occurs when manipulating an object, for example? The taxonomy might also have to be extended for bimanual interfaces. Many interfaces (including FreeForm) allow the user to orient the object of interest with the nondominant hand and alter the object with the dominant hand holding a haptic device. We are not sure if the current taxonomy is capable of describing these interfaces.

The taxonomy allows the development of a systematic body of results and open questions. If new devices or interaction techniques are designed to provide better support for geometric perception, their effectiveness can be evaluated using the shape recognition protocol. Above all, the taxonomy focuses attention on testing interfaces against tasks, a more comprehensive metric than electromechanical properties of the device. Device-oriented metrics indicate that the PHANToM and other point force devices can capably render a large range of haptic sensations. However, performance at the device level does not guarantee adequate performance at rendering all haptic percepts. As with any tool, these devices are well suited to some tasks but not others. Evaluations must be done at the application level.

7. Acknowledgments

We gratefully acknowledge the donation of a PHANToM and Pentium computer by Intel Corporation. Susan Lederman, Gary Meyer, Michal Young, and Paul van Donkelaar provided detailed comments on the first author's doctoral dissertation, from which this work is derived. Christine MacKenzie suggested simplifying the path data by collapsing it to 1-dimensional velocity.

8. References
