A shape recognition benchmark for evaluating usability of a haptic environment

Arthur E. Kirkpatrick and Sarah A. Douglas
Department of Computer and Information Science
University of Oregon
Eugene, OR, 97403 USA
{ted,douglas}@cs.uoregon.edu

ABSTRACT

This paper describes a benchmark task for evaluating the usability of haptic environments for a shape perception task. The task measures the ease with which observers can recognize members of a standard set of five shapes defined by Koenderink. Median time for 12 participants to recognize these shapes with the PHANToM was 23 seconds. This recognition time is within the range for shape recognition of physical objects using 1 finger but far slower than recognition using the whole hand. The results suggest haptic environments must provide multiple points of contact with an object for rapid performance of shape recognition.

Keywords

Haptic interfaces, benchmarks, PHANToM

INTRODUCTION

Haptic interfaces are frequently claimed to permit “more natural” (and hence more usable) interactions than current WIMP interface styles. These claims are based upon analogies with physical environments, where vision and touch have complementary roles and touch is both a familiar and a necessary part of interaction. But do these analogies hold? Existing performance metrics for haptic interfaces summarize the mechanical and electrical characteristics of the interface hardware [1]. These are useful for comparisons of the hardware but provide only a limited indication of the performance of the overall system on actual tasks. We believe that the usability of a haptic interface is best answered by considering all aspects of the interface: the hardware, software interaction techniques, and the task to which they are applied. We call such combinations *haptic environments* and contrast them to our daily interactions with physical objects, which we call *physical environments*.

We have developed a benchmark task for evaluating one aspect of usability of haptic environments. We believe that shape perception is a fundamental component of many haptic perceptual tasks. Consequently, we have designed a task that measures the ease with which observers can recognize members of a standard set of five shapes. In this paper we describe our benchmark and use it to characterize performance of a common haptic environment.

THE SHAPE RECOGNITION TASK AND STIMULI

Many potential applications of haptic environments include haptic shape perception as a component task. An obvious example is analysis of scientific data, such as a geophysicist searching isosurfaces of a volumetric dataset for the contours of an oil field. However, there are many other tasks that implicitly require the assessment of shape, even though the primary attribute extracted may not itself be inherently geometric. In fact, we believe that any task that involves haptic exploration of the properties of an unknown object has an underlying shape component. To make sense of a stream of haptic sensations, an observer must be able to relate them to some form of spatial representation of the object, both for the haptic description itself and possibly to correlate that description with visual sensations of the object. Thus even the perception of such non-shape properties as texture or hardness will frequently be described with respect to regions of shape—“it’s smooth *on the side*”, or “it’s hard *at the protrusion*”.

Many studies of human haptic performance in physical environments have used either shape recognition or the related task of object recognition. Shape recognition tasks use stimuli which can only be discriminated based upon the geometric arrangement of their features, whereas object recognition tasks feature stimuli such as familiar household objects for which material properties are also a strongly diagnostic attribute. Object recognition was used in the classic series of studies by Klatzky and Lederman where observers were asked to haptically identify common objects [3]. In another study, the same authors used a restricted set of objects which could not be readily distinguished by material properties, producing a shape recognition task [4]. Other researchers have used abstract stimuli for shape recognition tasks [2].

The central role of shape perception makes it an excellent benchmark task for evaluating the usability of haptic environments. The large body of shape recognition results for physical environments provides a base for comparison. The selection of a set of standard shapes is crucial to the success of such a benchmark.

We began with a class of smooth-flowing three-dimensional shapes defined by Koenderink and van Doorn [5] and used in the shape recognition task of Kappers et al. [2]. These shapes are constructed from two orthogonal parabolas and are a canonical set in the following sense: Any more complex solid shape can be constructed from a combination of these shapes. The shapes are identified by a *shape* scale, computed from the direction of curvature of the two parabolas. There are five critical points on the scale, named Cup, Rut, Saddle, Ridge, and Cap (see Figure 1). 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.

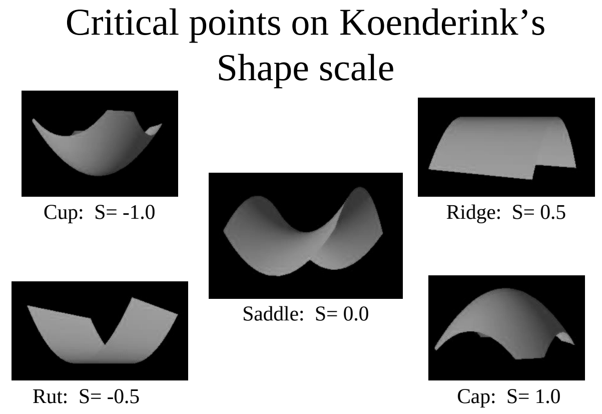

Figure 1: The Koenderink shape scale.

These five shapes represent an excellent base for a standard set. They are smoothly curving, can be used to construct larger shapes of arbitrary complexity, and can be readily distinguished by assessing direction of curvature without assessing its degree. Performance of observers recognizing these shapes constitutes a baseline performance for shape recognition in a haptic environment.

Shape recognition is potentially based upon both abstract knowledge of the shape (in the case of our shapes, the directions of curvature for the two parabolas) and hand movements specific to an exact configuration. The balance between these two will depend upon the frequency of contact with a specific object. For example, you might identify an arbitrary cup shape using one sequence of hand movements but identify the coffee cup you drink from every day using movements specific to its geometry. This presented us with a dilemma: We needed to teach our participants to haptically recognize the five shapes but did not want them to use recognition methods that relied upon specific geometries of a given stimulus set. We resolved this by training our participants with one set of stimuli and testing them with another. The training phase used medium-sized shapes presented in "head on" configuration whereas the testing phase used small- and large-sized shapes presented in different rotations.

Using these stimuli, we developed a benchmark task, summarized in the next section, for evaluating haptic environments. In addition to evaluating the benchmark, we were interested in the relationship between haptics and vision in shape recognition. The perception of shape using a point force haptic device alone is somewhat difficult. We wondered whether the visual proprioception offered by a screen cursor might improve the time and accuracy of performance by providing a superior representation of the spatial location of the current haptic sensation. To test this, we had our participants perform our benchmark shape recognition task with both a visual cursor display and no visual feedback whatsoever.

METHOD

The benchmark was implemented for SensAble's PHANToM device. The participants never saw a visual representation of the shapes at any time during the protocol—the shapes were only rendered haptically. The protocol had two separate phases, a training phase and a 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" 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 testing phase used a 2 by 2 by 3 within-subjects design, with cursor condition, stimulus size, and rotation as the independent variables and time and accuracy as the dependent variables. The two levels of the cursor condition were cursor present (a visual cursor was displayed on the screen corresponding to the location of the haptic device in the virtual environment) or absent. A curtain prevented the participant from seeing the location of their hand and ensured that the only visual cue was the cursor, if present.

The small and large sized objects spanned distances of approximately 1.5 cm and 7 cm, respectively. 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. Shapes, sizes, and rotations were fully crossed, for a total of thirty combinations in each cursor condition. The complete testing phase consisted of 60 trials.

The haptic environment was a 300-Mhz Pentium II running Windows NT 4.0. The visual display device was a monoscopic color screen and the haptic display device was a PHANToM model 1.5. 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 with the stimulus until the participant ended the trial by pressing the space bar.

The participants were 12 unpaid graduate students from the computer science department. Nine were male and three female. Their ages ranged from 22 to 42 with a median of 30.6. Ten were right handed and two left handed. All used the PHANToM with their dominant hand.

RESULTS

Recognition times and accuracy

Participants took quite a while to recognize the shapes. Arithmetic mean, geometric mean, and median time for a trial were 28.6, 22.5, and 23.8 seconds, respectively, with 50% of the values between 13.7 and 37.7 seconds. The distribution of times was clearly 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 seconds, a factor of 3.11. The mean score for participant accuracy was 84.5% (s.d. 12.0%).

Effects of cursor condition and size

We computed the mean time and accuracy for each participant under each experimental condition and calculated a two-way within-subjects ANOVA of main effects and interactions for accuracy and log of time. No significant effects were found for accuracy. For time, the effect of cursor condition was both unreliable in direction ($F_{1,11} = .098$, $p = .760$) and small (95% confidence interval [-14%, 16%]). No effect was found for order of presentation of cursor condition. Smaller sizes had 12% longer recognition times ($F_{1,11} = 6.986$, $p = .023$, 95% CI [2%, 22%]). The interaction effect between size and cursor condition was not significant.

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 mean time of 22.5 seconds with a 15% error rate. How does performance in this environment compare with human performance at recognizing physical objects? Using a superset of the five shapes we selected for our benchmark, Kappers et al. [2] found accuracy rates close to 100%. They do not report response times.

Lederman, Klatzky, and Reed [6] devised stimuli simpler than ours, three ellipsoids of revolution that differed only in their height to width ratio. Using both hands, observers could distinguish these objects in 1.0 seconds. For common household objects, whose shapes are more complex than those we used, observers were able to perform haptic object recognition in under 2 seconds with a 4% error rate [3]. With a set of common objects that did not appreciably differ in material properties (a shape recognition task), the mean time was 6.2 seconds with a 5 % error rate [4].

These times are far faster than the performance in our haptic environment, but comparisons must be made cautiously. Participants in all these studies were able to use their full hand. When Klatzky et al. [4] required their participants to wear gloves, the mean response time rose to about 16 seconds. When they further restricted their participants to using a single gloved index finger, mean response time leapt 2.8 times to 45 seconds with an error rate of 25%. Restricting the haptic flow to a single point, requiring the observer to induce object shape over time, dramatically limits performance in physical environments. Note that this last condition corresponds most closely to using a point force device in a haptic environment.

These comparisons suggest that we have a considerable room for improvement of shape display in haptic environments. If we can increase the number of points of contact between the user and a displayed shape, we might get a two- to three-fold improvement in performance—a gain of great practical consequence, given the underlying importance of shape recognition and how long it currently takes.

These comparisons also provide a useful validation of our evaluation benchmark itself. The response times in our task are well within the range that would be predicted from data on a comparable task with physical objects. Our task appears to measure the determining factors in performance of shape recognition.

Finally, we consider the non-significance of the cursor condition. We note that all participants continued to improve performance up until the 60th trial. They do not appear to have reached skilled performance during the course of our benchmark. 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 may have reduced with practice. The visual cursor might have proved significant when participants had achieved practiced performance.

CONCLUSIONS AND FUTURE WORK

Many users clearly have difficulty performing this basic task of shape recognition with a point force device. Given that perceiving shapes is a fundamental component of many tasks for which we might wish to use such devices, this difficulty represents a significant barrier to their usability

Understanding the various factors underlying this low rate of performance is crucial to the usability of environments incorporating point force haptic devices. In the future, we intend to study how long it takes individuals to reach skilled performance, what that level of that performance is, and what factors might facilitate spatial perception with a point force device for different user populations.

We also would like to extend our set of reference shapes with edges, textures, and shapes that are between our five basic shapes. This larger reference set will allow us to explore the effect of multiple factors on shape perception performance.

ACKNOWLEDGEMENTS

We gratefully acknowledge Intel Corporation for the donation of a Pentium computer and PHANToM. Susan Lederman gave insightful comments on point force haptics and Roberta Klatzky graciously provided the precise numerical values for Fig. 2 of [4]. Reference [1] was suggested by an anonymous referee of this paper.

REFERENCES

1. Hayward, V. and Astley, O.R. Performance measures for haptic interfaces. In *Proceedings of Robotics Research: The 7th International Symposium*. 1996. Springer Verlag, Berlin, pp. 195-207.
2. Kappers, A.M.L., Koenderink, J.J., and Lichtenegger, I. Haptic identification of curved surfaces. *Perception and Psychophysics 56*, (1994), 53-61.
3. Klatzky, R.L., Lederman, S.J., and Metzger, V.A. Identifying objects by touch: An "expert system". *Perception and Psychophysics 37*, (1985), 299-302.
4. Klatzky, R.L., Loomis, J.M., Lederman, S.J., Wake, H., and Fujita, N. Haptic identification of objects and their depictions. *Perception and Psychophysics 54*, (1993), 170-178.
5. Koenderink, J.J. and van Doorn, A.J. Surface shape and curvature scales. *Image and Vision Computing 10*, (1992), 557-565.
6. Lederman, S.J., Klatzky, R.L., and Reed, C.L. Constraints on haptic integration of spatially shared object dimensions. *Perception 22*, (1993), 723-743.