On the Ability of Humans to Haptically Identify and Discriminate Real and Simulated Objects

Marcia K. O'Malley
Mechanical Engineering and Materials Science
Rice University
Houston, TX 77005-1892 USA
omalleym@rice.edu
Michael Goldfarb
Mechanical Engineering
Vanderbilt University
Nashville, TN USA

Abstract

The ability of human subjects to identify and discriminate between different-sized real objects was compared with their ability to identify and discriminate between different-sized simulated objects generated by a haptic interface. This comparison was additionally performed for cases of limited force and limited stiffness output from the haptic device, which in effect decrease the fidelity of the haptic simulation. Results indicate that performance of size-identification tasks with haptic-interface hardware capable of a minimum of 3 N of maximum force output can approach performance in real environments, but falls short when virtual surface stiffness is limited. For size-discrimination tasks, performance in simulated environments was consistently lower than performance in a comparable real environment. Interestingly, significant variations in the fidelity of the haptic simulation do not appear to significantly alter the ability of a subject to identify or discriminate between the types of simulated objects described herein.

1 Introduction

The primary purpose of a haptic interface is to present an effective simulated mechanical environment to a human user. Though several published studies exist that compare the quality or effectiveness of various haptic simulation techniques relative to other simulation techniques, relatively few experimental comparisons exist that directly compare human performance in a haptically simulated environment relative to their performance in a real environment.

There is a vast amount of literature on human haptic recognition (identification) and discrimination of size in real environments. Several of these studies focus specifically on perception via a handheld probe (Chan & Turvey, 1991; Barac-Cikoja & Turvey, 1995); however, they do not address perceptual performance in simulated environments. Another body of literature presents studies that compare performance of tasks in simulated and real environments, including those by Buttolo, Kung, and Hannaford (1995); Shimojo, Shinohara, and Fukui (1997); West and Cutkosky (1997); Richard, Coiffet, Kheddar, and England (1999), and Unger et al. (2001). The studies by Buttolo et al., Richard et al., and Unger et al. all compare task-completion times for both real and simulated tasks. The experiments of Buttolo et al. indicated that the completion times for the real and simulated tasks were nearly the same. For a different set of experimental conditions, however, Richard et al. and Unger et al. reported significant increases in task-completion times for the simulated versus real tasks. Unlike these prior works, West and Cutkosky compared the ability of human subjects to identify spatial frequency in real and simulated sinusoidally varying textured surfaces. Their experiments indicated that the ability of a human to identify spatial frequency was impaired in the simulated versus real cases. Finally, Shimojo et al. performed comparisons in shape recognition between simulated and real objects for a pin-matrix-type tactile display. Their experiments also indicated that the ability of human subjects to identify shapes was impaired by the tactile interface.

This paper characterizes the ability of humans to identify and discriminate shape primitives simulated

Presence, *Vol. 14, No. 3, June 2005,* 366–376

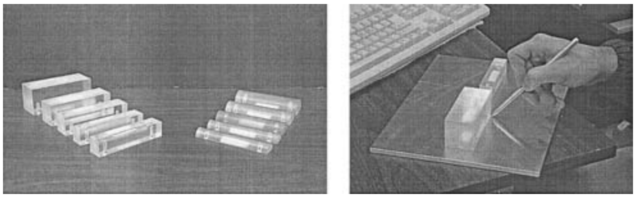

Figure 1. *Photograph of the real blocks and the environment for a square-object size-discrimination task.*

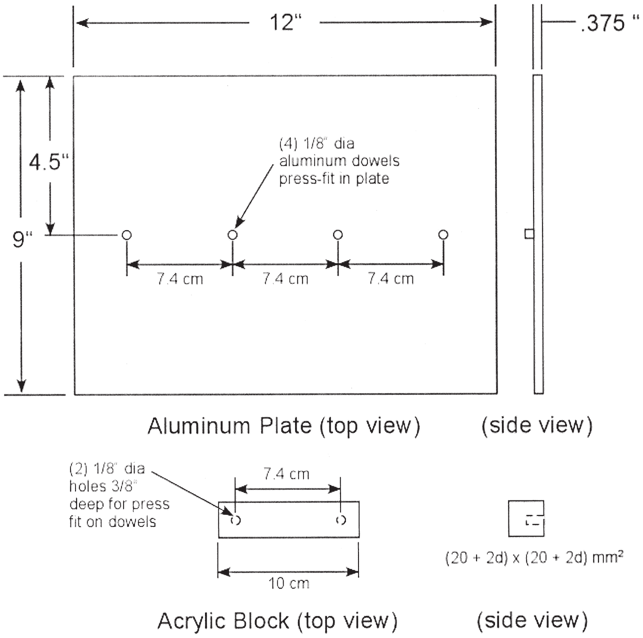

Figure 2. *Dimensions of aluminum plate and acrylic blocks used in real-environment experiments.*

with a haptic interface as compared to their performance of the same identification and discrimination tasks for real shape primitives. Shape primitives are defined as simple 3D shapes that can be combined to form more complex 3D objects. In this case, square cross-section and semicircular cross-section objects constitute the shape primitives. This study further investigates the performance of humans in identification and discrimination tasks in a virtual environment for the cases of limited force and stiffness output.

2 Methods

2.1 Testing Environments

The real environment used for the perception tests incorporated round and square cross-section shape primitives fabricated from acrylic, as shown in Figure 1. These objects were presented on an aluminum baseplate, which was fitted with four dowels that were used to secure the shape primitives to the baseplate. The dowels were arranged such that one block could be placed on the center of the baseplate (as was the procedure for size-identification experiments), or two blocks could be placed side by side for discrimination tasks. Figure 2 indicates the dimensions of the aluminum plate and the nominal dimensions of one of the acrylic shape primitives. A smooth-tipped aluminum stylus was fabricated to probe the shapes. The surfaces of all blocks were smooth to minimize friction (which was not modeled in the simulated environments). Contact paper on the surface of the blocks ensured that any machining irregularities could not be used as cues in the identification or discrimination tasks. Audio cues that arose during the experiments were masked by the sound of fans for the haptic device's motor amplifiers.

The simulated environment was constructed to emulate the experimental setup shown in Figure 1. As such, the simulation included a "baseplate," along with variable-sized round and square cross-section shape primitives of the same dimensions shown in Figure 1. The simulations were performed with a three degree-of-freedom point-contact force-reflecting haptic interface, shown in Figure 3, which was designed to exhibit minimal rotational inertia, minimal friction forces, zero backlash, and maximum link stiffness (Perry, 1997), physical characteristics that are generally known to facilitate high-fidelity haptic simulations (Ellis, Ismaeil, & Lipsett, 1993). The interface was used as an impedance operator for the haptic simulations, and as such, the 3D motion of the stylus endpoint was measured and the 3D force vector corresponding to the rendered haptic environment was displayed. The surfaces were simulated as a simple unilateral spring and damper at a sampling frequency of 3000 Hz. System bandwidth is approximately 100 Hz, limited by first-order low pass filters placed on each of the motor torque command signals. This partic-

Figure 3. *Test subject seated at testing station for simulated-environment experiments.*

ular apparatus is capable of displaying constant forces of over 10 N in the spatial region of the haptically displayed objects, and peak forces of roughly 40 N.

In order to test the real and simulated environments under what were considered like circumstances, the real-environment interactions were constrained via a probe, rather than allowing the subjects to use a more natural configuration of the hand. Recent findings of Lederman and Klatsky (2004) show that constrained manipulation involves a loss of information transfer to the subject by eliminating spatially distributed information and relaying it in a sequential manner. This ultimately results in significantly lower performance in tasks performed with a probe as compared to those with the real hand. Based on these findings, the subjects were constrained to interact with both the real and simulated environments via a probe.

It is expected that performance in the real environment will not differ much from that in the high- or low-fidelity simulated environments. This hypothesis is based on prior work that showed limited improvements in task performance in the virtual environments for forces and stiffnesses greater than 3 N and 470 N/m, respectively. Further, it is hypothesized that force has less of an effect on performance than stiffness, since relatively low values of force were sufficient whereas higher values of stiffness (relative to the achievable limit of the device) were necessary to reach the same level of performance.

2.2 Experimental Design and Procedures

Size-identification and size-discrimination experiments were performed in both real and simulated environments. For each experiment (identification or discrimination), the between-subjects factor was the size difference between objects, for which there were two levels (2.5 mm and 5 mm). The within-subjects factors included the object shape (two levels: square and round) and the environment type (four levels: real, high-fidelity simulated, low-fidelity simulated [force]; and low-fidelity simulated [stiffness]). Six subjects performed experiments in each of the simulated environments, and 13 subjects performed experiments in the real environment. The factors and levels are summarized in Table 1.

2.2.1 Size Identification. Size-identification tasks determine the ability of a test subject to classify similarly shaped objects, presented one at a time, by size alone. The identification of square cross-section shape primitives (Experiment 1A) and semicircular cross-section shape primitives (Experiment 1B) were conducted separately using three sizes for each size difference set, as described by Figure 4 and Table 2. The radius of the smallest shape primitive was always 1 cm. The medium and large sizes were generated by adding a constant *d*, the size difference (2.5 mm or 5 mm), to this radius. Note that for square shaped primitives, the "radius" corresponds to half the edge length.

2.2.2 Size Discrimination. Size-discrimination experiments test the ability of a human subject to notice size differences between objects placed side by side. The discrimination of square cross-section shape

Table 1. *Experiment Details—Factors and Levels for all Experiments*

		Experiment Factors		
Experiment	No.	Shape primitive	Size (mm)	Environment type (Number of subjects in parentheses)
Size identification	1A	Square	2.5 and 5	Real (13) High-fidelity simulated (6) Low-fidelity simulated (force) (6) Low-fidelity simulated (stiffness) (6)
Size identification	1B	Round	2.5 and 5	Real (13) High-fidelity simulated (6) Low-fidelity simulated (force) (6) Low-fidelity simulated (stiffness) (6)
Size discrimination	2A	Square	2.5 and 5	Real (13) High-fidelity simulated (6) Low-fidelity simulated (force) (6) Low-fidelity simulated (stiffness) (6)
Size discrimination	2B	Round	2.5 and 5	Real (13) High-fidelity simulated (6) Low-fidelity simulated (force) (6) Low-fidelity simulated (stiffness) (6)

Note: Object sizes correspond to half of object edge length for square cross-section shape primitives and to object radius for round cross-section shape primitives.

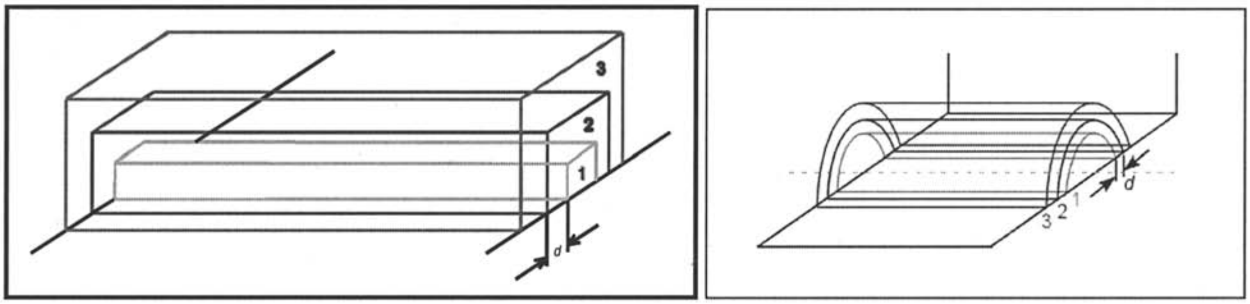

Figure 4. *Representation of square (left, Experiment 1A) and round (right, Experiment 1B) cross-section objects in three rendered sizes showing object size difference, d.*

primitives (Experiment 2A) and semicircular cross-section shape primitives (Experiment 2B) were conducted separately using two size differences (2.5 mm and 5 mm) for each shape primitive, as described by Figure 5 and Table 1.

2.3 Quality of Haptic Simulation

Experiments 1 and 2 were conducted for the real environment and for three different simulated environments corresponding to the maximum performance ca-

Table 2. *Object Sizes (mm) for each Test Session (Experiments 1A and 1B)*

Session number	Small (1)	Medium (2)	Large (3)	Difference in object size (mm)—d
1	10.00	12.50	15.00	2.50
2	10.00	15.00	20.00	5.00

pability of the haptic interface (i.e., high-fidelity simulation) and two variations of decreased performance (decreased-fidelity simulations). The high-fidelity performance capability of the haptic interface corresponds to a maximum continuous end-point force of 10 N, and a maximum simulated surface stiffness of 1000 N/m. The two lower fidelity performance conditions correspond to limiting the maximum continuous endpoint force to 3N and limiting the maximum simulated surface stiffness to 470 N/m, respectively. The damping coefficients used for the high- and low-fidelity stiffnesses were 100 Ns/m and 47 Ns/m, respectively.

The values of maximum endpoint force and maximum virtual surface stiffness were chosen based on previous findings of the authors. Prior work has characterized the effect of maximum force output and varying virtual surface stiffness on the ability of human subjects to perform perceptual tasks including size identification and size discrimination in a simulated environment (O'Malley & Goldfarb 2002; 2004). For the force-output experiments, results showed that 3 to 4 N of maximum force feedback to the user was sufficient to achieve good performance in the perception tasks, while the hardware was capable of up to 10 N of continuous force feedback. Higher levels of force feedback did not produce better human performance in the tasks. In the virtual surface stiffness experiments, test results indicated that performance, measured as a percent correct score in the perception experiments, improves in a nonlinear fashion as the maximum level of virtual surface stiffness in the simulation increases. Further, test subjects appeared to reach a limit in their perception capabilities at maximum stiffness levels of 470 N/m, while the hardware was capable of 1000 N/m of maximum virtual surface stiffness. These results indicate that haptic interface hardware may be able to convey sufficient perceptual information to the user with relatively low levels of maximum force output and virtual surface stiffness. However, performance of these perceptual tasks in simulated and real environments has not been directly compared, nor has performance in low-fidelity virtual environments been compared to performance in real environments. These performance comparisons are the focus of this paper.

2.4 Experimental Protocol

2.4.1 Experiment 1: Size Identification. Objects were placed, one at a time, in front of the subject. Each subject was asked to identify, via haptic interaction with the stylus-type probe, whether the object was the small, medium, or large object. The subject's hand and the environments, both simulated and real, were shielded from view by a curtain. A single test session consisted of one size difference, one shape primitive, and one of the four environments. A training session prior to each test session was used to familiarize the subjects with the three sizes of objects they would be classifying. Each subject completed two test sessions. During the testing, subjects were presented each of the three sizes 15 times in a random order, for a total of 45 stimuli. Six subjects were tested in each of the simulated environments, and 13 subjects were tested in the real environment. The order of testing was randomized, such that some subjects identified 2.5-mm size differences followed by 5-mm size differences, and some identified 5-mm differences followed by 2.5-mm differences.

2.4.2 Experiment 2: Size Discrimination. For each shape primitive (square or round cross-section), two objects were placed side by side in front of each subject. Each subject was asked to determine, via haptic interaction with the stylus-type probe, which of the two objects was larger. The subject's hand and the environ-

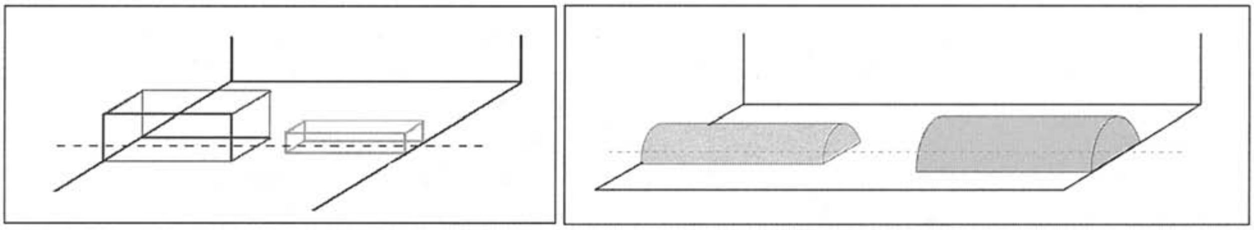

Figure 5. *Model of the simulated environment for the square and round object size-discrimination task.*

ments, both simulated and real, were shielded from view by a curtain. A single test session consisted of all size differences, one shape primitive, and one of the four environments. A training session was allowed prior to each test session that mimicked the actual experiment, yet gave feedback after each user response. Training sessions occurred in the same environment type that was to be tested, and subjects were allowed to determine the amount of training they underwent prior to each test session. Fourteen trials of each stimulus pair were presented for each test, for a total of 28 trials per test session. Six subjects were tested in each of the simulated environments, and 13 subjects were tested in the real environment.

3 Results

Table 1 reviews the factors and levels that were implemented in the experiments. The dependent variable for all test sessions was the mean accuracy achieved by a subject for a given experiment-shape-size-environment combination. Due to the varying number of subjects that participated in the experiments, a 3-factor mixed ANOVA using unequal n's was used for statistical analysis of the results, with size-identification and size-discrimination results analyzed separately. For each experiment, size difference was the within-subjects factor, while shape primitive and environment type were the between-subjects factors.

Results for the two experiments are presented in sections 3.1 and 3.2. For Experiments 1 and 2, pairwise comparisons were analyzed to determine significant differences between pairs of conditions. Tukey (1952; 1953) proposes a test designed specifically for pairwise comparisons based on the studentized range, sometimes called the "honestly significant difference test," that controls the MEER (Maximum Experimentwise Error Rate) when the sample sizes are equal. Tukey (1953) and Kramer (1956) independently propose a modification for unequal cell sizes. This method has fared extremely well in Monte Carlo studies (Dunnett, 1980). In addition, Hayter (1984) gives a proof that the Tukey-Kramer procedure controls the MEER for means comparisons, and Hayter (1989) describes the extent to which the Tukey-Kramer procedure has been proven to control the MEER for LS means comparisons. The Tukey-Kramer method is more powerful than the Bonferroni, Sidak, or Scheffe methods for pairwise comparisons.

Post hoc analyses, including higher level interactions and a comparison to chance, are presented in section 3.3.

3.1 Experiment 1: Size Identification

Figure 6 shows averaged results across subjects for Experiment 1A for the four environment cases (real, simulated high fidelity, simulated limited force, and simulated limited stiffness). As seen in the figure, the size identification for the maximum quality simulation falls within 10% of the real environment, followed closely by the limited force, then stiffness cases. Figure 7 shows averaged results across subjects for Experiment 1B for the four environment cases.

Shape, size difference, and environment type were all significant factors in the size-identification experiment [Shape: $F(1, 189) = 12.64, f = 0.0005$; Size: $F(1, 189) = 91.07, p < .0001$; Environment type: $F(3, 189) = 5.04, p = 0.0007$]. Pairwise comparisons for

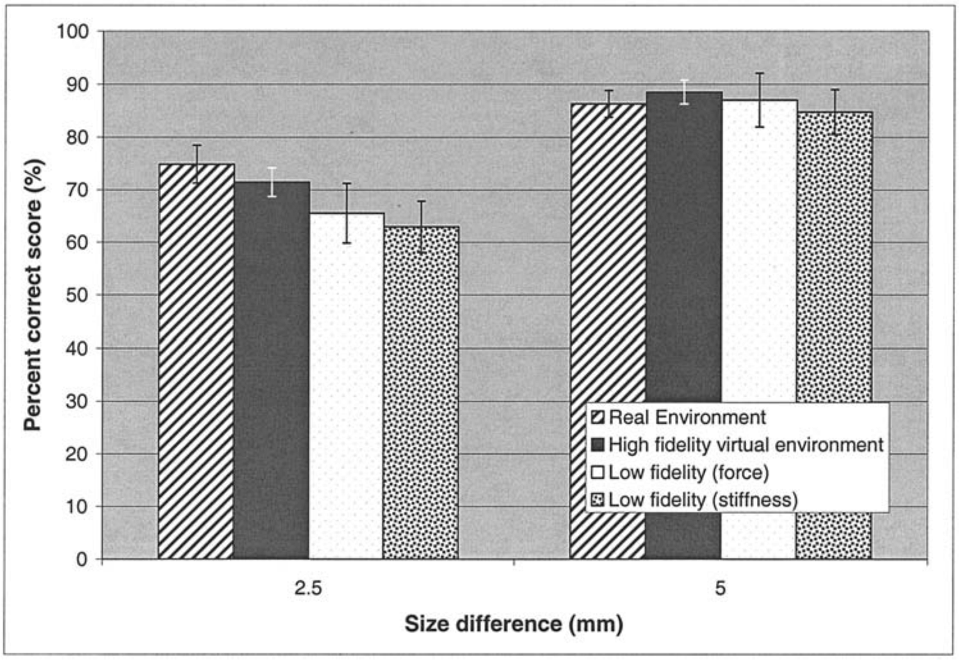

Figure 6. *Performance comparison for Experiment 1A (size identification of square shape primitives) for the four environment cases.*

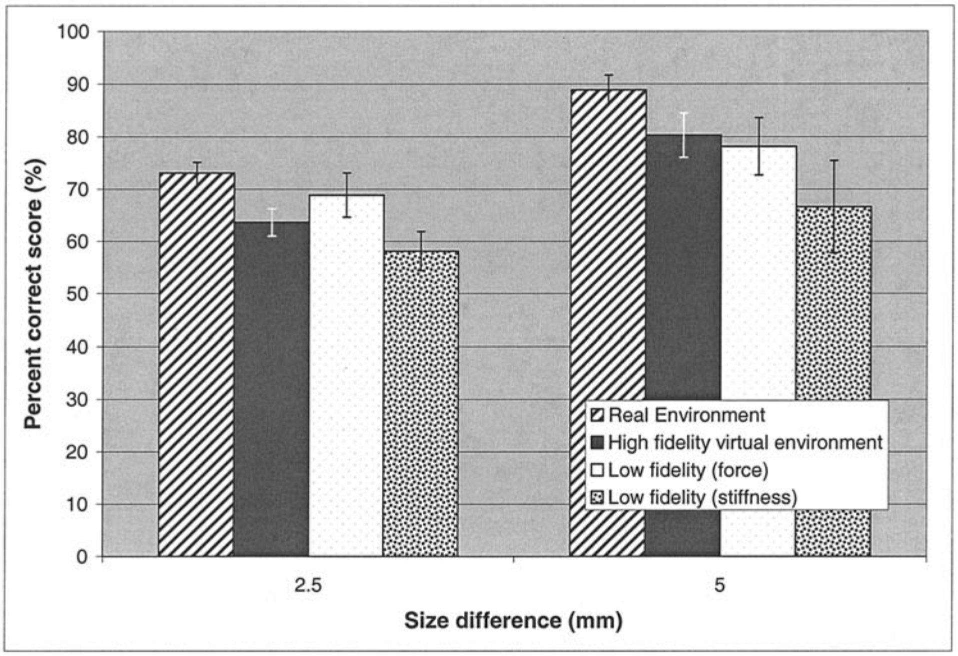

Figure 7. *Performance comparison for Experiment 1B (size identification of round shape primitives) for the four environment cases.*

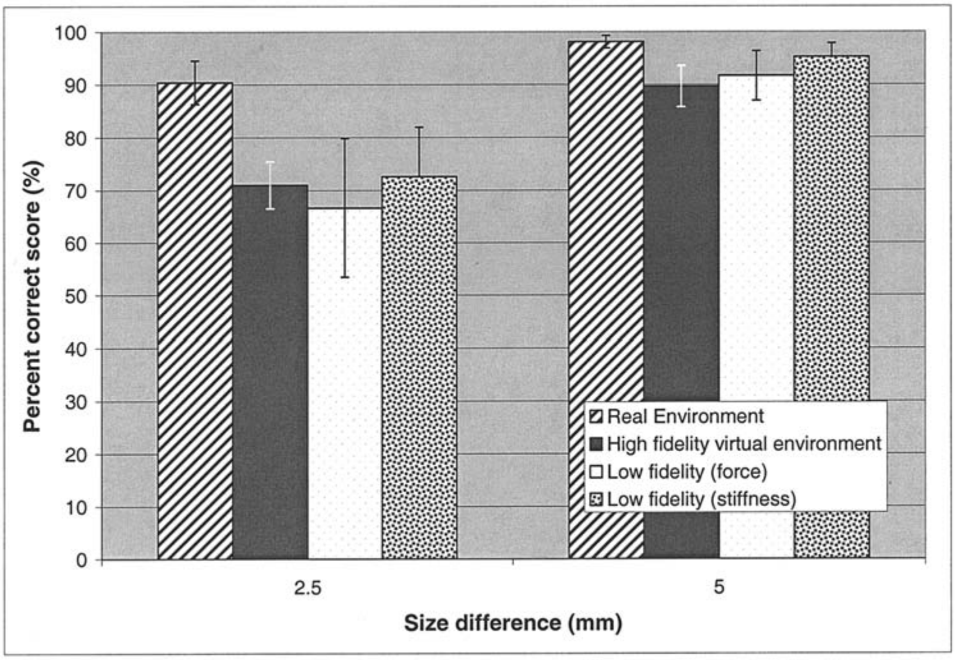

Figure 8. *Performance comparison for Experiment 2A (size discrimination of square shape primitives) for the four environment cases.*

environment type indicate that mean accuracies in the real environment and in the high-fidelity simulated environment are significantly different from those in the low-fidelity simulated environment where stiffness is reduced from the maximum achievable value for the haptic device, but not significantly different from performance in the low-fidelity simulated environment where maximum endpoint force is limited. Here we can conclude that for the size-identification task, performance in a low-fidelity simulated environment where maximum output force is limited to 3 N is comparable to performance in a high-fidelity simulated environment, and to performance in a real environment. For limitations in maximum virtual surface stiffness, performance is degraded significantly as compared to performance in a real or high-fidelity simulated environment. Therefore, to ensure good performance of size-identification tasks, designers should first aim to create simulated environments with high virtual surface stiffness, and should treat maximum force output of the haptic device as a secondary design goal.

3.2 Experiment 2: Size Discrimination

Figure 8 shows averaged results across subjects for Experiment 2A for the four environment cases (real, simulated high fidelity, simulated limited force, and simulated limited stiffness). Figure 9 shows averaged results across subjects for Experiment 2B for the four environment cases. As seen in the figures, the size-discrimination performance for the simulated environments approaches that of the real environment, but in general falls short.

Size difference and environment type were significant factors in the size-discrimination experiment, while shape was not significant [Shape: $F(1, 165) = 1.14$, $p = .2871$; Size difference: $F(1, 165) = 78.77$, $p < .0001$; Environment type: $F(3, 9.68)$, $p < .0001$]. Pairwise comparisons for environment type indicate that mean accuracies in the real environment are significantly different from those in any of the simulated environments. Performance in the high-fidelity simulated environment was not significantly different from that in either of the low-fidelity simulated environments, which supports prior work by the authors that showed that performance

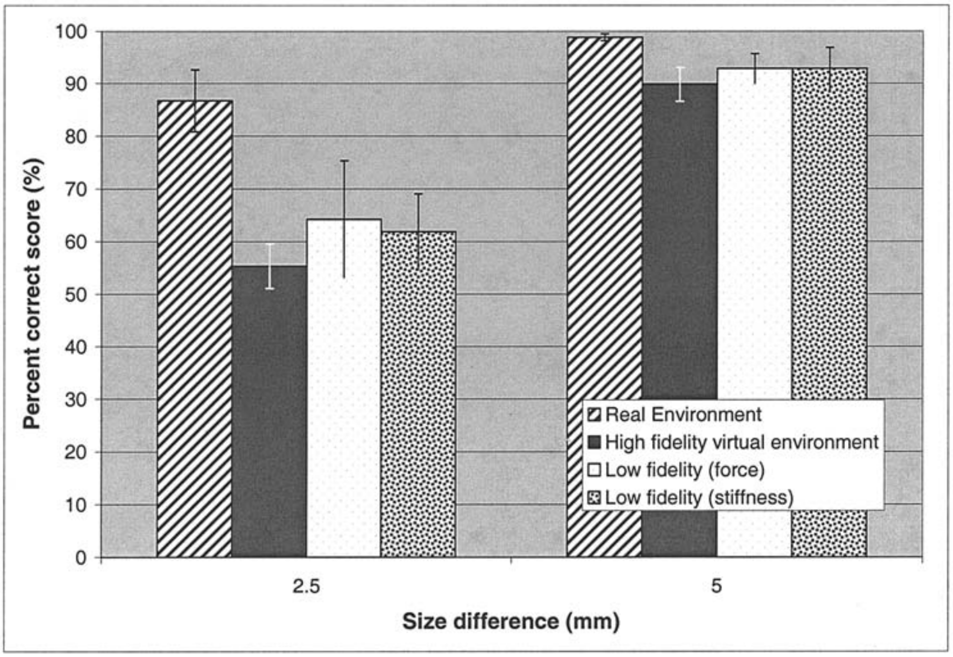

Figure 9. *Performance comparison for Experiment 2B (size discrimination of round shape primitives) for the four environment cases.*

in simulated environments with limited maximum force output or limited virtual surface stiffness did not significantly affect performance in size-identification and size-discrimination tasks (O'Malley & Goldfarb 2002; 2004). From these findings, designers of haptic devices should aim to reach recommended minimum levels of maximum force output and virtual surface stiffness (3 N and 470 N/m, respectively) to ensure acceptable performance of size-discrimination tasks, but should note that this performance will never reach the level that can be attained in a comparable real environment.

In recent related experiments with a commercial haptic interface, the PHANToM, an environment with both limited stiffness and limited force produced the same results in terms of object-discrimination performance as an environment with only limited stiffness, or an environment with only limited force (Upperman, Suzuki, & O'Malley, 2004). These results were obtained only for the discrimination task on a commercial haptic interface, not the device used for the work in this paper. Additionally, the results were not compared to performance of the discrimination task in a real environment. However, these findings suggest that limited force output coupled with limited virtual surface stiffness should not have a compounding detrimental effect on performance of size-discrimination tasks in simulated environments.

3.3 Additional Statistical Results

In addition to the results presented in sections 3.1 and 3.2, performance in all environments was compared to chance, to determine if the results were statistically different. For the size-identification experiments, chance would result in mean accuracy scores of 33%, since there were three possible responses that could be given. For the size-discrimination experiments, chance would result in mean accuracy scores of 50%, since there were two possible responses that could be given. Results of the binomial analysis showed that results for all experiment-type, shape-primitive, size-difference, and simulated-environment combinations were significantly better than chance.

Finally, two-way interactions were analyzed, using combined results from the size-identification and size-discrimination experiments. For this analysis, experiment type was treated as an additional factor, with two levels (identification and discrimination). There was only one significant interaction, which occurred between size and environment type [$F(3, 345) = 2.41$, $p = 0.0487$]. In other words, particular sizes may be easier to identify or discriminate with particular values of maximum endpoint force or virtual surface stiffness. This is reasonable based on the algorithm that is used to generate the simulated-environment surfaces. The force commanded to the haptic interface is proportional to distance of penetration of the probe tip into the virtual object and is also proportional to the velocity at which the probe is moving in the environment. As a result, the size of the virtual object constrains the maximum force that can be felt, since larger objects allow deeper penetration of the probe tip and therefore can give rise to larger forces. Also, the virtual surface stiffness will control the proportionality between the amount of penetration of the object and the force that is commanded to the motors. Because of these relationships, it is reasonable to expect a higher level interaction between size and environment type.

4 Conclusions

The experiments presented in this paper compare human perceptual performance in a real environment to performance in a simulated environment for two perception tasks, size identification and size discrimination. Findings indicate that performance of size-identification tasks with haptic interface hardware with reasonable maximum force output can approach performance in real environments, but falls short when virtual surface stiffness is limited. For size-discrimination tasks, performance in simulated environments was consistently lower than performance in a comparable real environment. Interestingly, significant variations in the fidelity of the haptic simulation do not appear to significantly alter the ability of a subject to identify or discriminate between the types of simulated objects described herein.

To ensure good performance of size-identification tasks, designers of haptic interfaces should first aim to create simulated environments with high virtual surface stiffness, and should treat maximum force output of the haptic device as a secondary design goal, since limited force output had an insignificant effect on performance when compared to performance in a real environment. For size-discrimination tasks, designers of haptic devices should aim to reach recommended minimum levels of maximum force output and virtual surface stiffness (3 N and 470 N/m, respectively) to insure acceptable performance, but should note that this performance will never reach the level that can be attained in a comparable real environment.

Acknowledgments

The work presented herein was supported by NASA Grant NGT-8-52859 and by the Micron Scholars program at Rice University. Many thanks go to Scott Baggett of the Statistical Consulting Lab at Rice for his assistance with the statistical analysis of results. The authors wish to thank all of the perception-test volunteers for their participation. Finally, thanks to the anonymous reviewers for their insightful comments, and special thanks to Susan Lederman for her suggestions and comments for improving this paper.

References

Barac-Cikoja, D., & Turvey, M. T. (1995). Does perceived size depend on perceived distance? An argument from extended haptic perception. *Perception Psychophysics, 57*(2), 216–224.

Buttolo, P., Kung, D., & Hannaford, B. (1995). Manipulation in real, simulated, and remote environments. *Intelligent Systems for the 21st Century. Systems, Man, and Cybernetics, 5,* 4656–4661.

Chan, T. C., & Turvey, M. T. (1991). Perceiving the vertical distances of surfaces by means of a hand-held probe. *Journal of Experimental Psychology—Human Perception and Performance, 17*(2), 347–358.

Dunnett, C. W. (1980). Pairwise multiple comparisons in the homogeneous variance, unequal sample size case. *Journal of the American Statistical Association, 75,* 789–795.

Ellis, R., Ismaeil, O., & Lipsett, M. (1993). Design and evaluation of a high-performance prototype planar haptic interface. *Advances in Robotics, Mechatronics, and Haptic Interfaces, DSC, 49,* 55–64.

Hayter, A. J. (1984). A proof of the conjecture that the Tukey-Kramer method is conservative. *The Annals of Statistics, 12,* 61–75.

Hayter, A. J. (1989). Pairwise comparisons of generally correlated means. *Journal of the American Statistical Association, 84,* 208–213.

Kramer, C. Y. (1956). Extension of multiple range tests to group means with unequal numbers of replications. *Biometrics, 12,* 307–310.

Lederman, S. J., & Klatzky, R. L. (2004). Haptic identification of common objects: Effects of constraining the manual exploration process. *Perception and Psychophysics* 66(4), 618–628.

O'Malley, M., & Goldfarb, M. (2002). The effect of force saturation on the haptic perception of detail. *IEEE/ASME Transactions on Mechatronics, 7*(3), 280–288.

O'Malley, M., & Goldfarb, M. (2004). The effect of virtual surface stiffness on the haptic perception of detail. *IEEE/ASME Transactions on Mechatronics, 9*(2), 448–454.

Perry, C. M. (1997). Design of a high performance robot for use in haptic interface and force feedback research. Unpublished master's thesis, Vanderbilt University, Nashville, TN.

Richard, P., Coiffet, P., Kheddar, A., & England, R. (1999). Human performance evaluation of two handle haptic devices in a dextrous virtual telemanipulation task. *Proceedings of the IEEE International Conference on Intelligent Robots and Systems,* 1543–1548.

Shimojo, M., Shinohara, M., & Fukui, Y. (1997). Human shape recognition performance for 3D tactile display. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, 4,* 3192–3197.

Tukey, J. W. (1952). Allowances for various types of error rates. Unpublished IMS address, Chicago, IL.

Tukey, J. W. (1953). *The problem of multiple comparisons.* Unpublished manuscript.

Unger, B. J., Nicolaidis, A., Berkelman, P. J., Thompson, A., Klatzky, R. L., & Hollis, R. L. (2001). Comparison of 3-D haptic peg-in-hole tasks in real and virtual environments. *IEEE International Conference on Intelligent Robots and Systems, 3,* 1751–1756.

Upperman, G., Suzuki, A. & O'Malley, M. (2004). Comparison of human haptic size discrimination performance in simulated environments with varying levels of force and stiffness. *Proceedings of the Twelfth Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems,* Chicago, IL.

West, A. M., & Cutkosky, M. R. (1997). Detection of real and virtual fine surface features with a haptic interface and stylus. *Proceedings of the ASME IMECE, DSC, 61,* 159–166.