

The Role of Contextual Haptic and Visual Constraints on Object Manipulation in Virtual Environments

Yanqing Wang and Christine L. MacKenzie

School of Kinesiology
Simon Fraser University
Burnaby, BC V5A 1S6
Canada

+1 604 291 5794

{wangy, cmackenz}@move.kines.sfu.ca

ABSTRACT

An experiment was conducted to investigate the role of surrounding haptic and visual information on object manipulation in a virtual environment. The contextual haptic constraints were implemented with a physical table and the contextual visual constraints included a checkerboard background ("virtual table"). It was found that the contextual haptic constraints (the physical table surface) dramatically increased object manipulation speed, but slightly reduced spatial accuracy, compared to free space. The contextual visual constraints (presence of the checkerboard) actually showed detrimental effects on both object manipulation speed and accuracy. Implications of these findings for human-computer interaction design are discussed.

KEYWORDS

Human performance, virtual reality, visual information, haptic information, 3D, docking, controls and displays, task context, force feedback, graphic interface, degrees of freedom, augmented environment.

INTRODUCTION

Virtual environments provide domain constraints and contextual constraints for human object manipulation. Domain constraints include intrinsic properties of the object being manipulated such as controller and cursor size and shape [14]. Contextual constraints are the surrounding information and environment for object manipulation. The goal of this experiment is to explore the role of contextual haptic and visual constraints on multidimensional object manipulation in virtual environments.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI '2000 The Hague, Amsterdam
Copyright ACM 2000 1-58113-216-6/00/04...\$5.00

Contextual haptic constraints

There are two kinds of contextual haptic constraints in virtual environments: active and passive haptic constraints. Active haptic constraints are provided with force feedback devices. Passive haptic constraints are implemented with the real object in augmented environments, for example, where graphic cues are augmented with surrounding physical cues. This study deals with the effects of passive haptic constraints on object manipulation. Recent research shows that such passive haptic feedback not only provides realism in virtual environments [7], but also enhances human performance [9]. Lindeman, Sibert and Hahn compared human performance on docking a graphic object to a "floating" graphic panel with docking when the panel was augmented with a physical paddle [9]. They found that the passive haptic feedback provided by the augmented paddle resulted in a 44% decrease in the movement time and a 38% increase in accuracy. Zhai reported a study on a six degrees of freedom elastic controller for object manipulation tasks [19]. The elastic constraint can be considered as a kind of haptic constraint on the controller. They found that human performance was better when the elastic device was used, compared with isometric devices, and suggested that the elastic property of the controller provided more sensitivity for position control [19].

Contextual haptic constraints on object manipulation are essentially the problem of degrees of freedom (DOF) for movement control. Previous research has not explicitly examined the effect of contextual haptic constraints in terms of degrees of freedom of movements. Furthermore, in an augmented environment, contextual haptic constraints are usually spatially related to contextual visual constraints. We are unaware of research that addresses contextual haptic constraints in relation to contextual visual constraints. These aspects of contextual haptic constraints are investigated in this study.

Contextual visual constraints

Contextual visual constraints such as graphic checkerboard or groundplane background have been a standard technique to provide depth cues in graphic interfaces [2]. The depth cues provided by contextual visual constraints can facilitate human performance in virtual environments [1] [12]. Robertson, Czerwinski and Larson suggested that the spatial graphic background improves the user's spatial memory for information visualization [11]. The role of contextual visual constraints on object manipulation in virtual environments needs further investigation.

Research hypotheses

In this experiment, we used a physical table to provide the contextual haptic constraints for object manipulation. We compared human performance on object manipulation on the table surface with that in free space. The movement on the table surface had fewer degrees of freedom than in free space. A graphic checkerboard, a "virtual table", served as the contextual visual constraint. The physical table surface was overlaid with the checkerboard. We tested two hypotheses:

1. Contextual haptic constraints enhance human performance on object transportation and orientation;
2. Contextual visual constraints facilitate human performance on object transportation and orientation.

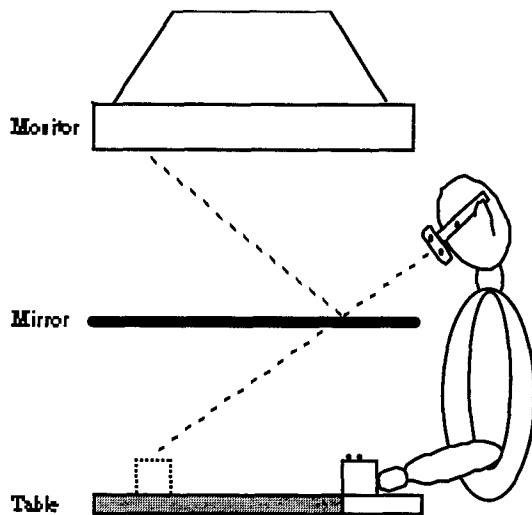


Figure 1. The Virtual Hand Laboratory setup for this experiment. The stippled part of the table surface is removable. When this part is removed, the subject manipulates the wooden cube in free space. A graphic checkerboard, or "virtual table" appears as if on the table surface. The graphic target cube (dashed line) is drawn to appear on the table surface. The wooden cube (solid line) is the controller. Markers on the goggles and wooden cube drive the stereoscopic, head-coupled graphics display.

METHOD

Subjects

Eight university student volunteers were each paid \$20 for participating in one, two-hour experimental session. All subjects were right-handed, and had normal or corrected-to-normal vision. All subjects had experience using a computer. Informed consent was provided before the experimental session.

Experimental setup

This experiment was conducted in the Virtual Hand Laboratory (VHL) [14], shown in Figure 1. A SGI monitor was placed screen down on a specially constructed cart. A half-silvered mirror was placed parallel to and between the computer screen and the table surface. The image on the screen was reflected by the mirror, and was perceived by the subject as if it were in a workspace on the table surface. The images for the left and right eye were alternatively displayed and were synchronized with goggles to provide the subject with a stereoscopic view. Three infrared (IRED) markers were placed on the side of goggles. An OPTOTRAK 3D motion system monitored 3D position information from these markers to provide a head-coupled view.

The physical object, a wooden cube, was the controller, serving as the input device. Three IREDs were placed on top of the controller cube. The 3D position information from these three IREDs was used to drive a red six-degree-of-freedom wireframe graphic object, the cursor cube (not shown in Figure 1). The information from these IREDs on the top of the wooden cube was also recorded for data analysis. The transportation data were collected with the marker on the top center of the wooden cube. The orientation data were derived with two markers on the top of the wooden cube.

The target was a red wireframe graphic cube generated on the monitor, appearing on the table surface for the subject looking into the mirror, as shown in Figure 1. The graphic target was placed along the horizontal center axis on display, which was aligned with the subject's body midline. The controller, cursor and target cubes had the same size of 30 mm. The graphic target was located 70, 140 or 210 mm away from the starting position, with an angle rotated 22.5 or 45 degrees clockwise about a vertical axis.

As shown in Figure 1, one part of the table surface was removable. The other part of the table with the same height was used to support the controller at its start position. The graphic target was perceived by the subject through the mirror, as if on the table surface. When the table surface was present, the subject could slide the controller on the table surface; when the table surface was removed, the subject had to move the controller to the target in the air, without the table as a supporting surface. The part of the table surface at the start location of the controller was

always present so that the controller was supported at the beginning of the movement.

This setup provided three contextual haptic conditions: table-slide, table-lift and no-table. In the table-slide condition, the physical table was present and the subject was instructed to slide the controller on the table surface. The table-lift condition was when the physical table was present, but the subject was instructed to slightly lift the controller from the table surface and land the controller on the table surface at the final position. In the no-table condition, the table was removed and the subject had to move the controller in the air to its final position. In all cases, the controller started in a constant, supported location and the subject's task was to align the cursor cube with the target cube.

In the table-slide condition, the wooden cube (controller) and the cursor cube were constrained to three degrees of freedom, two for translation and one for rotation. In the table-lift condition, the wooden cube and cursor cube were constrained to three degrees of freedom only at the start and the end of the movement; the controller and cursor had six degrees of freedom for free motion between the start and the end. In the no-table condition, the wooden cube and the cursor cube had six degrees of freedom after it left the start position.

A black and white checkerboard was displayed with a block size of 13.5 by 13.5 mm. The checkerboard was superimposed on the planar table surface. When the checkerboard was not present, object manipulation was performed on a black background.

Procedure

System calibration was first performed [13]. The workspace on the table surface, including the checkerboard, was calibrated so that the checkerboard was aligned to the tabletop. The cursor cube was registered to superimpose with the wooden cube. The individual subject eye positions were also calibrated to obtain a customized, stereoscopic, head-coupled view.

During the experiment, subjects saw only the graphic cursor and target, with no vision of the hand and the wooden cube. The task was to match the location and angle of the cursor cube to those of the graphic target as fast and accurately as possible. Trials were blocked on contextual haptic constraint and visual constraint conditions. At the beginning of each block, subjects were given 20 trials for practice. The order of target distances and angles were randomly generated over trials within each block. Ten trials were repeated in each experimental condition.

Data analysis

Independent variables for this experiment were contextual haptic constraints, contextual visual constraints, target distances and target angles. Object transportation and orientation data collected from the IRED markers on the wooden cube were analyzed separately. Temporal dependent measures were: total task completion time (CT), object transportation time (TT) and object orientation time (OT). CT was the time between the start of the cube movement to the end of the cube movement, either cube translation or orientation. TT was the cube translation time and OT was the cube orientation time. Spatial error measures were: constant errors of distance (CED), variable errors of distance (VED), constant errors of angle (CEA), and variable errors of angle (VEA). The constant error was the average distance or angle off the target for all trials under an experiment condition, and the variable error was the standard deviation of trials for that condition. ANOVAs were performed on the balanced design of $3 \text{ haptic constraints} \times 2 \text{ visual constraints} \times 3 \text{ target distances} \times 2 \text{ target angles}$ with repeated measures within subjects.

RESULTS

Time Measures

Object transportation and orientation showed a parallel structure, that is, the transportation process completely contained the orientation process [17]. Noting that the results of transportation time (TT) were almost the same as task completion time (CT), only TT and OT are reported in detail here.

Transportation time (TT)

The average transportation time (TT) was 906 ms, taking up 97% of the task completion time (CT) of 933 ms. Haptic constraints had effects on TT, $F(2, 14) = 78.62, p < .001$. TT was 749 ms in the table-slide condition, 778 ms in the table-lift condition, and increased to 1192 ms in the no-table condition.

Haptic conditions interacted with target distance to affect TT, $F(4, 28) = 16.98, p < .001$, as shown in Figure 2. Post hoc analysis was performed on haptic constraint conditions for each target distance separately. Post hoc tests revealed that for each target distance, the no-table condition had a significantly longer TT than the table-slide and table lift conditions ($p < .05$). TTs between the table-slide and table lift conditions did not significantly differ from each other. The longer the target distance, the larger the no-table effect on TT.

Visual constraints significantly interacted with the target angle, $F(1, 7) = 6.20, p < .05$. The checkerboard actually increased TT, but it appeared that at 22.5 degrees, the presence of the checkerboard resulted in a larger increase in TT than at 45 degrees, as shown in Figure 3.

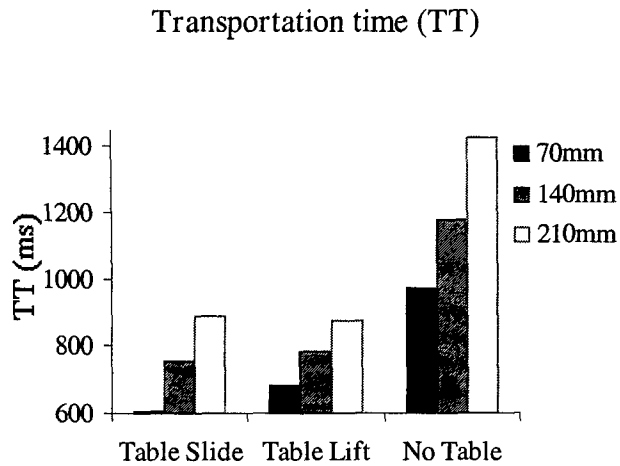


Figure 2. Transportation time with haptic constraints and target distances.

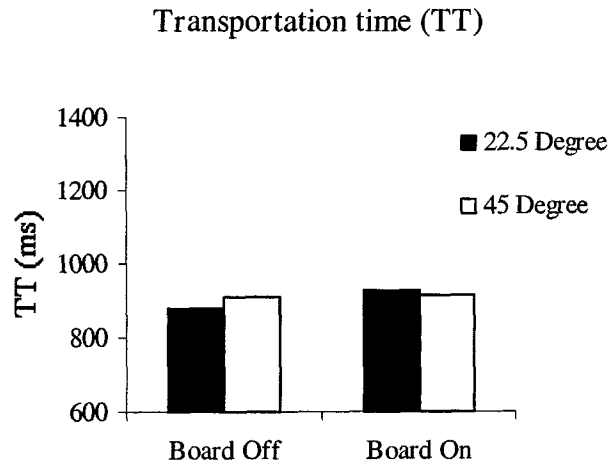


Figure 3. Transportation time with visual constraints and target angles. Board Off = Black background; Board On = Checkerboard background.

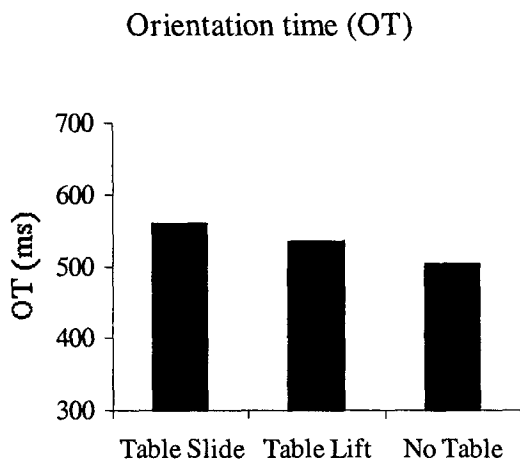


Figure 4. Orientation time with haptic constraints.

Orientation time (OT)

The orientation time (OT) was 533 ms on average, 57% of the task completion time (CT). The effect of haptic constraints on OT was significant, $F(2, 14) = 4.38, p < .05$, and showed a trend opposite to TT, as demonstrated in Figure 4. The longest OT occurred in the table-slide condition, 560 ms, and reduced to 536 ms in the table-lift condition, and then further decreased to 504 ms in the no table condition. Post hoc tests revealed that OT in the no-table condition was significantly different from that in the table-slide condition ($p < .05$). No significant difference in OT was found between the no-table and table-lift conditions nor between the table-lift and table-slide conditions.

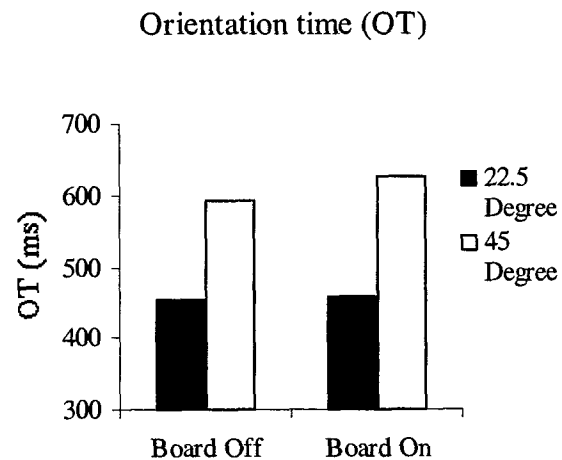


Figure 5. Orientation time with visual constraints and target angles.

There was a significant interaction on OT between visual constraints and target angles, $F(1, 7) = 10.18, p < .05$, as shown in Figure 5. It appeared that the checkerboard had more impact on OT at 45 degrees than at 22.5 degrees of target angle.

Spatial Errors

Overall, the spatial errors were very small. The average value of constant errors of distance (CED) was 0.04 mm, not significantly different from the target location. The average value of constant errors of angle (CEA) was 0.8 degree, not significantly off the target angle. Constant errors reflect system errors and subjects' individual bias while

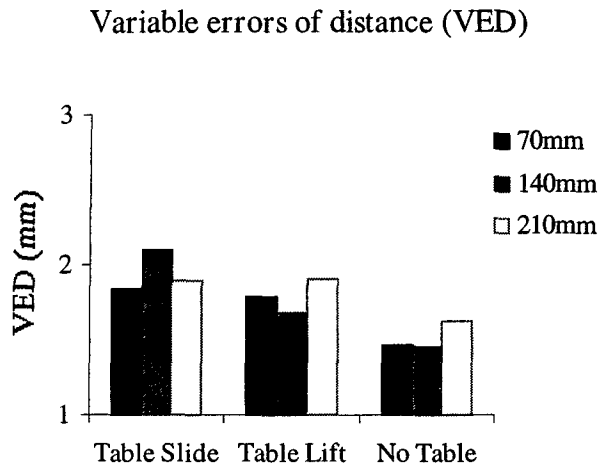


Figure 6. Variable errors of distance with haptic constraints and target distances.

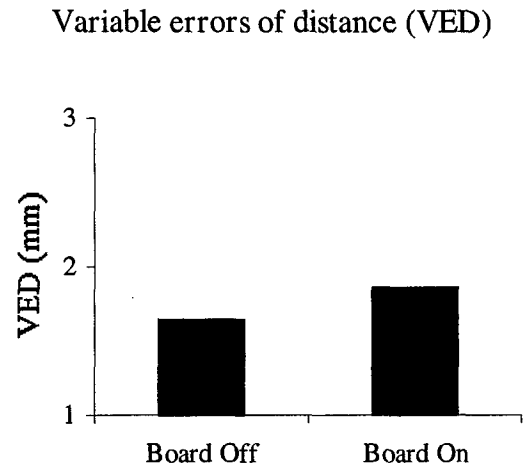


Figure 7. Variable errors of distance with visual constraints

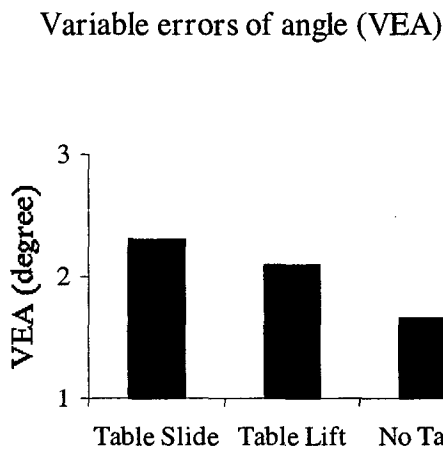


Figure 8. Variable errors of angle with haptic constraints.

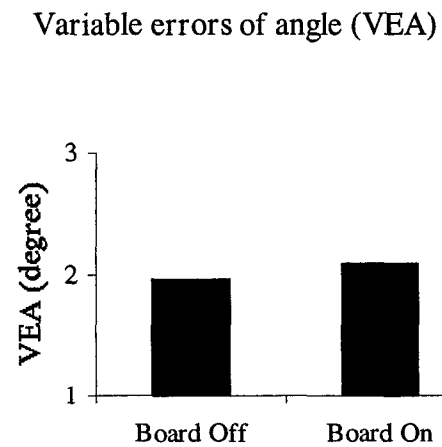


Figure 9. Variable errors of angle with visual constraints.

variable errors indicate human performance consistency under a certain interface system. Only the results of variable errors are presented as follows.

Variable errors of distance (VED)

Haptic constraints, visual constraints, and target angles had main effects on the variable errors of distance (VED). VED increased a small amount with changing haptic constraints (e.g., 1.5 mm in the no-table condition, 1.8 mm in the table-lift condition, and 1.9 mm in the table-slide condition), $F(2, 14) = 5.12, p < .05$. Post hoc analysis revealed a significant difference in VED between the table-slide and no-table conditions ($p < .05$), but there were no significant

differences between the table-slide and table-lift conditions nor between the no-table and table-lift conditions.

There was an interaction between the haptic constraint and the target distance, $F(4, 28) = 3.67, p < .05$, as shown in Figure 6. Post hoc analysis revealed that at the target distance of 140 mm, VED in the table-slide condition was significantly larger than that seen in the table-lift condition or in the no-table condition, but VEDs between the table-lift condition and the no-table condition did not differ from each other. The differences in VED among haptic constraints were not significant at target distances of 70 mm and 210 mm.

The presence of the checkerboard background increased VED from 1.6 mm to 1.9 mm ($F(1, 7) = 16.54, p < .01$), as shown in Figure 7. This was a small and unexpected result.

There was a three-way interaction on VED among the haptic constraint, visual constraint and target distance, $F(4, 28) = 4.33, p < .01$. VEDs consistently increased in the presence of the checkerboard at the target distance 70 mm across haptic constraints. The checkerboard appeared not to make a difference in VED in the table-lift at 210 mm or in the no-table at 140 mm.

Variable errors of angle (VEA)

The variable errors of angle (VEA) had an average value of 2.0 degrees. The haptic constraint had significant effects on VEA, $F(2, 14) = 12.27, p < .001$. VEA was 1.7 degrees in no-table, 2.1 degrees in table-lift, and 2.3 degrees in table-slide, as shown in Figure 8. Post hoc revealed that VEA in the no-table condition was significantly smaller than that both in the table-lift and table-slide conditions ($p < .05$). There was no difference in VEA between the table-lift and table-slide conditions.

The checkerboard led to a small increase in VEA, $F(1, 7) = 6.56, p < .05$, from 2.0 to 2.1 degrees (Figure 9). It was unexpected that the presence of the checkerboard would have detrimental effects on both VED and VEA.

There was a three-way interaction on VEA among the haptic constraint, visual constraint and target distance, $F(4, 28) = 4.35, p < .01$. VEAs were similar or generally increased with the presence of the checkerboard with the exception in the table-lift condition at the target distance of 70 mm. In this particular condition, it appeared that the presence of the checkerboard reduced VEA.

DISCUSSION

The role of contextual haptic constraints

Haptic constraints had profound effects on human performance in object transportation and orientation. The task completion time (transportation time) was reduced dramatically with the tabletop, compared to when no supporting surface was present. This was consistent with the speed findings by Lindeman et al. [9]. This result also supports our suggestions in previous research that haptic information has more impact on object manipulation speed than visual information [15] [16]. The fact that the contextual haptic constraint was imposed in the control space indicates the importance of human motor control systems in object manipulation.

In this experiment, the task required three degrees of freedom for control. When the controller was moved in free space, it had six degrees of freedom. The degrees of freedom were reduced to three when the controller slid on the tabletop, and the controller actually became a three-

degree-of-freedom input device. Jacob, Sibert, McFarlane, and Mullen found that object manipulation speed increased when the structure (dimensions) of tasks was matched with the structure of input devices [8]. Evidence from this experiment supports their theory. Hinckley, Tullio, Pausch, Proffitt and Kassell conducted an experiment to compare two DOF input with three DOF input for three DOF rotation tasks [6]. They found in an orientation matching task, that users completed the task up to 36% faster when using three DOF input than two DOF input devices, without significant loss of accuracy.

No significant difference was found in task completion time between table slide and table lift conditions. At the end of movements, the controller was constrained to three degrees of freedom for both table slide and table lift. This suggests that the contextual haptic constraint on the end of the movement or the target is more critical than during the course of the movement. Similar results were found by MacKenzie for a 3D pointing task where subjects pointed to a solid target faster than to a hole [10]. MacKenzie suggested that the solid target (haptic constraint) helps to stop the movement, compared to pointing to a hole where subjects have to take extra time to decelerate the pointing device.

At the same time, however, the haptic constraint, the table surface, consistently increased the spatial variable errors in task performance, although the increase was quite small. This finding is counter-intuitive and contrary to Lindeman et al.'s results [9]. Hinckley et al. did not find that the constraints on degrees of freedom for object orientation had effects on spatial errors [6]. It is not clear what factors in these experiments cause such inconsistency. This needs further investigation.

The role of contextual visual constraints

It was originally predicted that the contextual visual constraint would be used as guidance for object manipulation. It is surprising that the visual constraint, the checkerboard background, generally deteriorated human performance in both times and spatial errors. Even though the effect of visual constraint was small, it is theoretically and practically important.

It appeared that the visual constraint, the checkerboard, interfered with the manipulation task rather than provided extra visual cues to enhance object manipulation performance. One interpretation is that the checkerboard background distracted the subject's attention from the target. Other factors such as the pattern and color of the checkerboard may also cause the interference.

If the same effect is replicated with various depth cues such as groundplanes or stereoscopic view, it poses an important question. Can depth cues of graphics, supposed to help in

visualization, benefit interaction in general? The results from previous research are not conclusive [1] [3] [4] [11] [18]. Recent research by Boritz shows that the depth cues provided by head-coupled view are generally detrimental for object docking in virtual environments [3]. Cao, MacKenzie and Payandeh found that the depth cues from stereoscopic view help for certain tasks, but not for others [4]. Our results show that the cues provided by the checkerboard actually hindered object manipulation, yet the checkerboard-like background has been widely used in human-computer interaction design [2].

There is strong evidence showing that humans have two visual systems, one for perception and another for prehension [5]. The theory of two visual pathways suggests that the depth cues that facilitate perception may not necessarily benefit action (object manipulation). We hypothesize that contextual visual constraints play different roles in object visualization than object manipulation. Future work is definitely needed to test this hypothesis.

Implications for human-computer interaction design

Virtual environment design should take advantage of passive haptic constraints. The significant benefit gained in the task completion time from the haptic constraint can be weighed against the relatively small reduction in accuracy control. Passive haptic constraints such as table surfaces, walls or paddles are cheap and reliable, and could be easily implemented in virtual environments. For example, a graphic 2D command menu in virtual environments could be augmented with a physical plate whenever possible.

Recently, force feedback devices have been implemented in virtual environments to enhance the realism of interaction. It is generally believed that force feedback devices improve human performance in accuracy control, but lack of empirical evidence. We found that the significant benefit of passive haptic constraints is for speed control rather than accuracy control. Thus, more attention should be paid to utilization of passive haptic feedback for applications where speed is a major concern. Force feedback devices should be designed to simulate passive haptic feedback as well as active haptic feedback.

We issue a caveat about the role of contextual visual constraints in virtual environment design. This experiment suggests that the background depth cues may benefit object perception, but degrade object manipulation. For example, in medical virtual reality applications, a checkerboard background may be used for diagnosis, but not for surgery. We cannot assume that graphic cues beneficial for perception will always be beneficial for interaction. The effect of contextual visual constraints should be carefully evaluated before implementation.

Contextual haptic and visual constraints have significant effects on human object manipulation in virtual environments. Research in this area is very limited, compared to research on domain constraints of object manipulation. Contextual constraints are ubiquitous in virtual environments, and human-computer interaction design should be guided by understanding contextual constraints as well as domain constraints.

CONCLUSIONS

It is concluded from this experiment:

1. Contextual haptic constraints, such as a table surface, improve human performance in the task completion time, but slightly reduce the spatial accuracy.
2. Contextual visual constraints, such as a checkerboard background, degrade human performance on object manipulation for both speed and accuracy.

ACKNOWLEDGEMENTS

This research was supported by the National Science and Engineering Research Council of Canada (NSERC) Strategic Project program. We would like to thank Evan D. Graham, Valerie A. Summers and Kellogg S. Booth for their help in software design for the Virtual Hand Laboratory.

REFERENCES

1. Arthur, K.W., Booth, K.S. and Ware, C. (1993). Evaluating 3D task performance for fish tank virtual world. *ACM Transactions On Information Systems*, 11 (3), 239-265.
2. Balakrishnan, R. and Kurtenbach, G. (1999). Exploring bimanual camera control and object manipulation in 3D graphics interfaces. *Proceedings of the Conference on Human Factors in Computing Systems CHI '99 / ACM*, 56-63.
3. Boritz, J. (1998). *The Effectiveness of Three-dimensional Interaction*. Ph.D. Thesis, Dept. of Computer Science, University of British Columbia, Vancouver, B.C., Canada.
4. Cao, C. G. L., MacKenzie, C. L. and Payandeh, S. (1996). Task and motion analyses in endoscopic surgery. *Proceedings of the ASME Dynamic Systems and Control Division*, 58, 583-590.
5. Goodale, M.A., Jakobson, L.S and Servos, P. (1996). The visual pathways mediating perception and prehension. In Wing, A. M., Haggard, P. and Flanagan, J. R. (Eds.), *Hand and Brain*, 15-31, New York: Academic Press.

6. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D. and Kassell, N. (1997). Usability analysis of 3D rotation techniques. *Proceedings of UIST '97*, 1-10.
7. Hoffman, H.G. (1998). Physical touching virtual objects using tactile augmentation enhances the realism of virtual environments. *Proceedings of IEEE VRAIS*, 59-63.
8. Jacob, R. J. K., Sibert, L. E., McFarlane, D. C. and Mullen, M. P. Jr. (1994). Integrality and separability of input devices. *ACM Transactions on Computer-Human Interaction*, 1 (1), 1-26.
9. Lindeman, R.W., Sibert, J.L. and Hahn, J.K. (1999). Towards usable VR: An empirical study of user interfaces for immersive virtual environments. *Proceedings of the Conference on Human Factors in Computing Systems CHI '99/ACM*, 64-71, Pittsburgh, PA.
10. MacKenzie, C.L. (1992). Making contact: Target surfaces and pointing implements for 3D kinematics of humans performing a Fitts' task. *Society for Neuroscience Abstracts*, 18, 515.
11. Robertson, G., Czerwinski, M. and Larson, K. (1998). Data mountain: Using spatial memory for document management. *Proceedings of UIST '98/ACM*, 153-162.
12. Servos, P., Goodale, M.A., and Jakobsen, S. C. (1992). The role of binocular vision in prehension: A kinematic analysis. *Vision Research*, 32 (80), 1513-1521.
13. Summers, V.A., Booth, K.S., Calvert, T., Graham, E. and MacKenzie, C.L. (1999). Calibration for augmented reality experimental testbeds. *ACM Symposium on Interactive 3D Graphics*, 155-162.
14. Wang, Y. (1999). *Object Transportation and Orientation in Virtual Environments*. Ph.D. Thesis, School of Kinesiology, Simon Fraser University, Burnaby, BC, Canada.
15. Wang, Y. and MacKenzie, C.L. (1999a). Object manipulation in virtual environments: Relative size matters. *Proceedings of the Conference on Human Factors in Computing Systems CHI '99 /ACM*, 48-55.
16. Wang, Y. and MacKenzie, C.L. (1999b). Effects of orientation disparity between haptic and graphic displays of objects in virtual environments. *INTERACT '99*, 391-398.
17. Wang, Y., MacKenzie, C.L., Summers, V. and Booth, K.S. (1998). The structure of object transportation and orientation in human-computer interaction. *Proceedings of the Conference on Human Factors in Computing Systems CHI '98/ACM*, 312-319.
18. Wickens, C.D. (1992). *Engineering Psychology and Human Performance*. Columbus, OH: Harper Collins.
19. Zhai, S. (1995). *Human Performance in Six Degree of Freedom Input Control*. Ph.D. Thesis, Dept. of Computer Science, University of Toronto, Toronto, Ontario, Canada.