A Study of Distance of Manipulation on the Responsive Workbench

Alexis Paljic, Sabine Coquillart, Jean-Marie Burkhardt, Paul Richard

A Study of Distance of Manipulation on the Responsive WorkbenchTM

A Study of Distance of Manipulation on the Responsive Workbench™

Alexis Paljic* — *INRIA* — *alexis.paljic@inria.fr*

Sabine Coquillart — *INRIA* — *sabine.coquillart@inria.fr*

Jean-Marie Burkhardt — *INRIA, Université Paris V* — *jean-marie.burkhardt@inria.fr*

Paul Richard — *ISTIA* — *paul.richard@istia.univ-angers.fr*

Abstract

Virtual environments such as the Responsive Workbench™[1] allow the close manipulation of virtual objects. We define close manipulation as either direct manipulation or manipulation at a relatively small distance. This paper reports on two studies on close manipulation using the Responsive Workbench. The first study investigates the influence of manipulation distance on performance in a 3D location task. The results indicate that direct manipulation and 20 cm distance manipulation are more efficient than for 40 and 55 cm distances. The second study investigates the effect of two factors: the presence or absence of a visual clue, and the scale value, which is a variation of the scale (1 or 1.5) used to map the user's movements to the pointer. Task performance is significantly lower when using the visual clue, and when using the 1.5 scale.

1. Introduction

The design of efficient interfaces for 3D applications is a concern for application developers and researchers. Virtual environment systems like the CAVE™[2] , the Responsive Workbench™(or Workbench), the FishTank, or Head-Mounted Displays (HMD's) are powerful technologies for interacting with 3D worlds. With such systems, the user is typically presented with stereoscopic views of the virtual world. By means of head tracking, the stereoscopic images can be presented to the user according to his viewpoint. In this case, the virtual and real spaces are superimposed and can be mixed together. By superimposing real and virtual spaces, systems utilizing head tracking allow close manipulation.

Close manipulation is often intended to provide the users with an intuitive way of interacting with 3D computer-generated worlds. In this paper, we define close manipulation as either direct manipulation, which is manipulation with the hand being as close as possible to the virtual objects, or manipulation at a relatively small distance (distances below one meter), with a gap separating the hand and the object being manipulated.

The workbench is one of the most attractive systems for close manipulation. Its table-like shape is well adapted to the close manipulation of virtual mock-ups or virtual models presented to the user. Unlike HMD, the workbench offers the advantage of not losing contact with the physical environment: the user's co-workers or his own body and hands. Furthermore, the workbench provides a large display and allows manipulation of models with a great freedom of movement.

As for most virtual environment systems, the design of user interfaces for the workbench is quite different from what we are used to with desktop configurations. Currently we have only a weak knowledge of user-interface principles with which to develop efficient interaction techniques suitable for the workbench. There is also clearly a need to provide both theoretically-grounded and empirically-validated rationales for the design of 3D interfaces. This paper presents two user studies addressing performance in different situations of close manipulation.

This study is also motivated by an observation frequently made during demonstrations on the workbench. Novice users often manipulate objects from a distance instead of bringing the hand as close as possible to the object. Holding the manipulated object at a distance is obviously a disadvantage at least for rotations because the center of rotation is at the hand and thus far from the object. However, does distance manipulation offer any compensating benefits? Is it more natural, more convenient, or more comfortable for the user? Should we advise, or even force the user to reduce as much as possible the distance of manipulation or not? More generally, close manipulation raises several questions: what is the influence of the distance of manipulation on task performance? Does the use of a virtual ray to fill the gap between the hand and the object in distance manipulation help the user? How does scaling user movements into the virtual

*Corresponding author: Alexis Paljic, INRIA Rocquencourt, Domaine de Voluceau 78153 Le Chesnay Cedex France.

[1] Responsive Workbench is a trademark of GMD.

[2] CAVE is a trademark of University of Illinois.

world affect user performance?

Some of these questions have been addressed previously, but never using a workbench. Extrapolating previous evaluation is hazardous because results differ depending on the configuration.

The objective of this study is to investigate the influence of the distance of manipulation on performance in a location task. An experiment using four values for distance was conducted so that the results could provide a view of performance variation for a large interval. A second experiment which studies the effect on user performance of two other conditions of manipulation is also presented. The first condition is the use of a virtual ray between the hand and the object being manipulated, the second consists of mapping user movements to the virtual objects with a 1.5 scale factor.

The next section describes the major previously-published studies related to close manipulation. In Section 3 we present the experiments and procedure used for the study. The results of the experiments are presented in Section 4, and a discussion is proposed in Section 5.

2. Related work

Two studies that compare direct manipulation and manipulation at a distance have been proposed by Mine [5] and Djajadiningrat et al. [2][3].

Mine [5] investigates the effect of proprioception on various interaction techniques using an HMD. In particular, the author carried out a study of direct manipulation to explore the differences between manipulating virtual objects co-located with one's hand and manipulating objects at a distance. The subjects performed a docking task using cubes. Three conditions were tested: manipulation of objects held in one's hand, objects held at a fixed offset and objects held at an offset varying according to the subject's arm extension. Statistical analysis shows that the manipulation of objects co-located with one's hand is significantly faster than the manipulation of objects at a fixed offset, as well as manipulation at a variable offset.

The study by Djajadiningrat et al. [2][3] compares direct manipulation and manipulation at a distance of 20 cm using two systems: a single screen fishtank-like setup and a small-size, three-screen, head tracked, non-stereoscopic virtual reality system called Cubby. The study was performed using a puzzle reconstruction task. No significant difference in performance between the two conditions for manipulation was found. However, subjects reported a subjective preference for direct over distant manipulation within each set-up.

A third study conducted by Poupyrev et al. [8] indirectly addresses the problem of close manipulation. An evaluation of different egocentric object manipulation techniques was carried out, among them the classical virtual hand and the go-go technique [7]. This study can be related to comparing direct and distant manipulation: the classical virtual hand method is direct manipulation since the object can be exactly co-located with the hand, and the go-go technique involves distance manipulation because the coupling between the input device and the cursor is non-linear. The study was done using a positioning task with an HMD system, so the user was totally immersed, without real-world visual feedback. For both techniques, the manipulated objects were located within arm's reach. When the task required moving the object closer or further away, statistical analysis did not reveal a significant difference in task performance due to technique. Conversely, when the initial and final positions of the object were at the same distance from the user, the direct manipulation technique was significantly 22% faster than the other techniques.

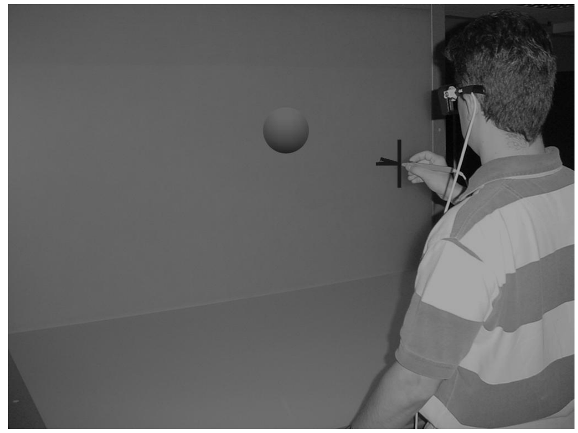

Figure 1. User performing the task.

3. Method

The objective of this study is to investigate the effect of manipulation distance on user performance in an environment that allows close manipulation. In Mine's study, distance of manipulation is randomly set between 10 and 60 cm, making it difficult to analyze the effect of variations of the distance of manipulation. We chose to conduct our study on fixed distance values. Four values were chosen in order to observe the effect of distance of manipulation within a large range of distances related to close manipulation. Another reason why we chose fixed values of distance is that it allows the user to stabilize his behaviour for a given distance, which is much harder when the distance is randomly changed for each trial, as in Mine's study.

Twenty-four participants were asked to perform a set of location tasks, in which they were presented with four manipulation distances. The users performed the task with a

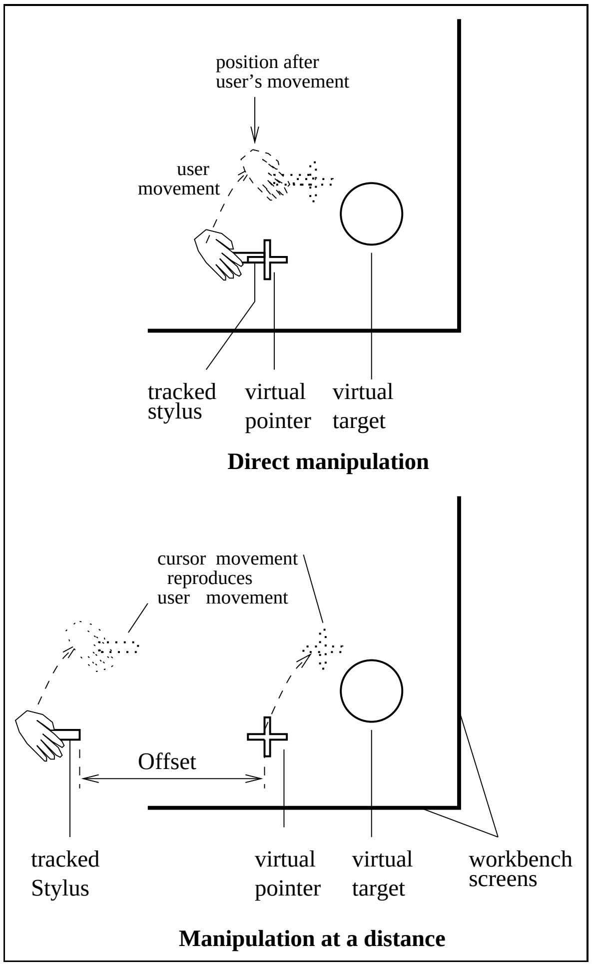

Figure 2. Side view of direct and distance manipulation on the workbench.

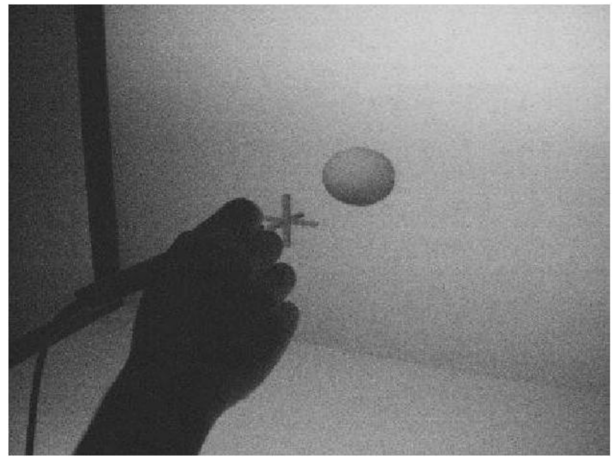

Figure 3. Direct manipulation: the cursor used to interact with the objects (cross) is as close as possible to the user's hand.

virtual pointer controlled by a tracker held in the hand (see Figure 1).

Distance of manipulation was set by changing the offset between the hand and the virtual pointer or cursor (see Figure 2). For direct manipulation, the cursor was displayed as close as possible to the hand of the user. This closest position is the tip of the tracked stylus. Figures 3 and 4 show two examples of close manipulation (direct and at a 55 cm distance).

As a continuation of the first study, and in order to help interpretation of its results, we carried out a second study to see how two common manipulation techniques could improve the user's performance in a 55 cm distance manipulation. The first manipulation technique consists of adding a virtual ray to bridge the gap between the user's hand and the cursor. The second maps the user's movements to the cursor with different scale values (1 and 1.5). The subjects, procedure, and equipment were the same for the two experiments, but the studied factors differ.

3.1. Equipment

The virtual environment system used for the experiment was a two-screen Responsive Workbench^{TM} using a 1280×1024 resolution for each screen, a 96Hz vertical refresh rate, and shutter glasses (Crystal Eyes^{TM}) for stereo vision. A 6 degrees of freedom tracker was used for head tracking, and a Fastrak Stylus tracker with one button for hand tracking and selection. The application was developed with the SGI Performer API, and runs on an SGI Onyx2.

3.2. Procedure

The location task consists of clicking on a start sphere and then clicking on a target sphere that appears at another location. Nine target sphere locations were chosen, each at an equal distance from the start sphere (see Figure 5). Let us assume that the x-axis is pointing towards the user and the z-axis is pointing upwards. The start sphere always appears 40 cm above the center of the horizontal screen at (0,0,0) coordinates. Two spheres were on the y-axis: p1(0,20,0) and p2(0,-20,0), two on the top-bottom z-axis: p3(0,0,20) and p4(0,0,-20). Five spheres were off-axis and located at p5(10,10,14.14), p6(10,10,-14.14), p7(10,-10,14.14), p8(10,-10,-14.14), and p9(-10,-10,14.14). No visual indications regarding the target positions were displayed, and in particular, the axes on Figure 5 were not displayed. During the test, the start sphere and the cursor

Figure 4. Distance manipulation: the cursor is 55 cm distant from the user's hand.

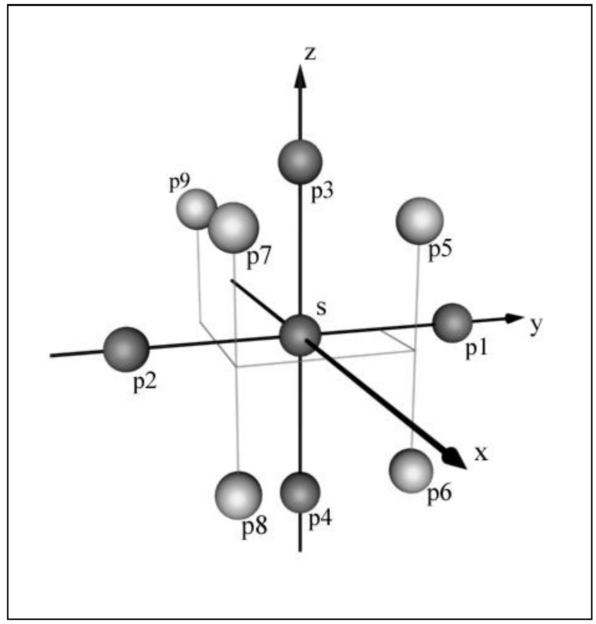

Figure 5. Target locations: s is the start sphere, p1 to p9 are the target spheres, the x-axis is pointing towards the user, the z-axis is pointing upwards.

were always displayed, and the current target sphere was displayed only during the task.

The subjects were instructed to perform the task as fast as possible. To perform the task, the user stood in front of the workbench and held the tracked stylus in his/her dominant hand to control the virtual pointer. The pointer was a cross-shaped cursor, used for clicking on the start and target spheres. The distance of manipulation was determined by the offset between the hand and the cursor (see Figure 2).

A click on a sphere triggers the next sphere to appear if the center of the cursor is inside the volume of the sphere (2.5 cm radius sphere) and the stylus button is pressed.

Only the translations of the tracked stylus are mapped onto the cursor, not rotations. The reason for this choice was to ensure that all users completed the task using the same displacement strategy, whatever the distance to the cursor was. Thus, to control the cursor the user had to translate his hand, not turn his wrist.

Each of the subjects was presented with one of the 24 permutations of 4 manipulation conditions studied in the experiment. For each condition, the subject accomplished a session composed of 3 trial tasks and 18 experimental tasks. For the experimental tasks, targets appeared exactly two times on each of the nine pre-defined locations in a random order, yielding 18 targets. At the end of each experiment, the user was asked to sort the conditions in terms of easiness.

3.3. Studied factors

For the first experiment, the design was composed of two factors: *target position* (the 9 pre-defined targets) and *offset* which is the distance between the tip of the stylus tracker and the cursor (4 values of the offset: 0, 20, 40 and 55 cm). The direction of this offset is in all cases parallel to the x-axis (see Figure 5), so the cursor location in space is the position of the hand, plus a translation (0, 20, 40 or 55 cm) on the x-axis. In this experiment the movement mapping between the hand of the user and the cursor was 1 to 1, meaning that the user's hand translations were exactly copied to the cursor without any multiplying constant.

In the second experiment, we compared four situations of manipulation yielded by the combination of two factors. The first is *visual clue*, which is a virtual ray linking the hand of the user to the cursor (two values: displayed, not displayed). When the ray is displayed, it is always directed on the x-axis, because no rotations are applied. The second factor is *scale*, for which two values are tested: 1 and 1.5. For a 1 scale, the offset between the hand and the cursor is 55 cm. Using a 1.5 scale results in a variable offset between the hand and the cursor (attached to the ray) during manipulation. For a 1.5 scale, the origin of the scale is chosen in order to have a 55 cm distance between the hand and the cursor when the user positions the cursor on the start sphere. The last factor studied in this second experiment is the *target position*, with the same 9 pre-defined targets as in the first experiment. The only difference in target position is for the 1.5 scale manipulation: targets were 1.5 times further from the start sphere than in the other situations. This is

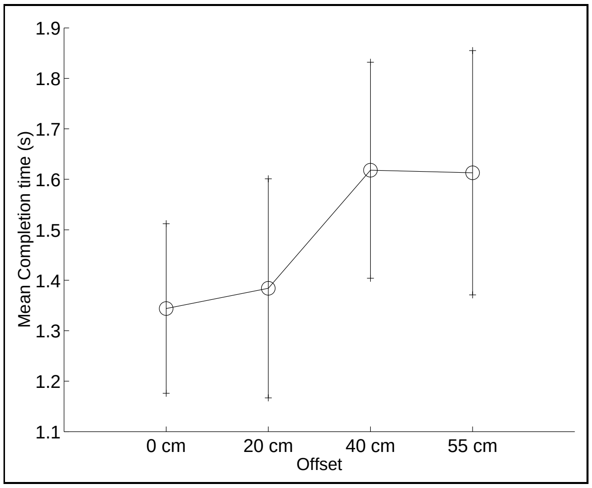

Figure 6. Mean task completion time for each of the offset conditions with standard deviation bars (first experiment).

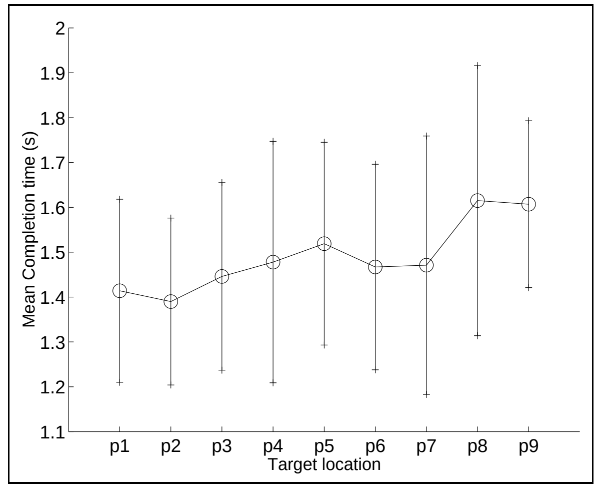

Figure 7. Mean task completion time for each target position with standard deviation bars (first experiment).

to ensure that the user's physical hand displacement remains the same as with the 1 scale.

3.4. Participants

Seven female and seventeen male subjects participated in the experiment. They all had normal or corrected vision. Eighteen were right-handed, five left-handed. One of the left-handed persons preferred to hold the tracked Stylus in his right hand.

4. Results

Task completion time is the dependent variable gathered for each task: it is the time elapsed between the click on the start sphere and the click on the target sphere. Subject ratings for easiness were also gathered to examine the subjective perception of the various manipulation conditions.

4.1. First experiment

This study is a 4 (*offset value*)×9 (*target position*) design. Task completion times for the two repetitions of each *target position* are averaged together to yield a single value. Figure 6 shows mean completion times. Analysis of variance (ANOVA) for mean task completion time revealed a significant effect of the *offset value* ($F(3,828)=131.191$ $p<0.0001$). Post-hoc analysis for offset value using the Scheffé test shows that subjects perform the task faster for 0 cm and 20 cm offsets than for 40 cm and 55cm offsets. There are no significant differences between 0 and 20 cm, nor for 40 and 55 cm *Target position* ($F(8,828)=16.55$ $p<0.0001$) has a significant effect on task performance. Post-hoc analysis using the Scheffé test for target positions shows that the off-axis targets p8 and p9 required significantly more time than the others. Target p5 takes significantly more time than p2. On-axis targets were always completed faster than off-axis targets, except for p4, which required more time than p6 or p7 to reach, but these results are not significant. p2 is the easiest target overall. Significant interaction effects were found for *offset values* × *target position* ($F(24,828)=5.574$ $p<0.0001$). Figure 7 shows mean completion time for each target.

4.2. Second experiment

This study is a 2 (*visual clue*) × 2 (*scale*) × 9 (*target position*) design. Anova for mean task completion time indicates that there is a significant effect of the *visual clue* ($F(1,828)=20.249$ $p<0.0001$). Globally, the users take more time to complete the task when the visual clue (virtual ray) is displayed than without this added information.

There is also a significant effect of the *scale factor* ($F(1,828)=253.330$ $p<0.0001$). Subjects performed the task significantly slower when the scale was 1.5. Figure 8 shows mean completion time for each condition. A significant effect was found for *target position* ($F(8,828)=11,16$ $p<0.0001$). Figure 9 shows mean completion time for each position. As in the first experiment, post-hoc analysis using the Scheffé test shows that the off-axis targets p8 and p9 required significantly more time than the others. Significant effects were found for both two-way interaction *scale* ×

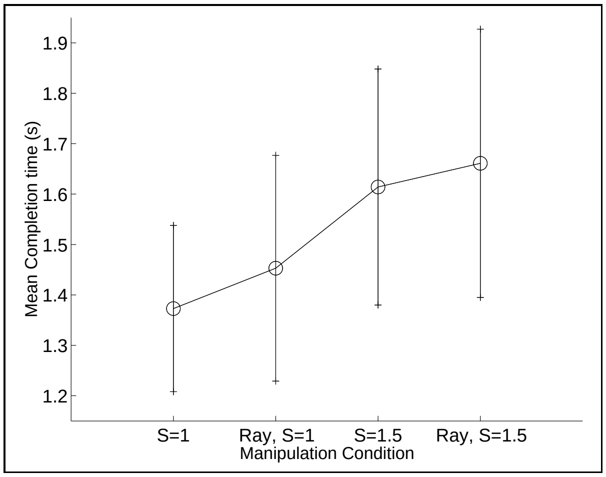

Figure 8. Mean completion time (seconds) for the second experiment conditions: Ray means that the virtual ray between the hand and the cursor is visible, S is the scale value.

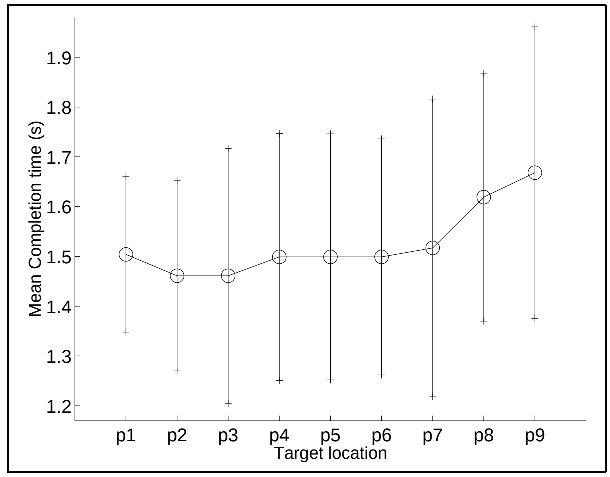

Figure 9. Mean completion time (seconds) for each target position in the second experiment

target positions (F(8,828)=4.44 p<0.0001) and *visual clue* × *target positions* (F(8,828)=2.27 p=0.0213). The three-way interaction for *scale* × *visual clue* × *target position* (F(8,828)=4.85 p<0.0001) is also significant.

4.3. Subjects comments

Subjects were asked to sort, in terms of easiness, the 4 manipulation conditions for each experiment. Table 1 summarizes rankings for the first experiment. Each column, corresponding to one condition, contains the number of persons who rated that condition respectively 1 (easiest), 2, 3 and 4 (hardest). For the first experiment, 0 and 20 cm distance manipulations are significantly better rated (first and second) than 40 and 55cm ($\chi^2(3)$=14.33, p<0.005). Conversely, the users show no significant preference between the manipulation conditions (χ^2 non significant) in the second experiment.

	0 cm	20 cm	40 cm	55 cm
1	10	5	2	7
2	6	12	4	2
3	3	4	12	5
4	5	3	6	10

Table 1. Subject rating for conditions in the first experiment, 1 is easiest, 4 is hardest.

5. Discussion

For the first experiment, a simple model of user behaviour could have been that performance in completion time decreases linearly with manipulation distance. The results show that this is not the case and that user behaviour is more complex.

The results for task completion time for the first experiment suggest the existence of two sets of manipulation distances between which user performance significantly differs. Distances from 0 to 20 cm can be interpreted as a direct manipulation area, since subject performance appears to be equivalent. This result replicates Djajadiningrat's result of no significant difference between 0 and 20 cm distance manipulations. For distances from 40 to 55 cm the task completion time is significantly higher than for 0 and 20 cm offsets. Furthermore, these higher distances seem to correspond to an area where the subjects adopt a different behaviour, or at least seem to be engaged in a more demanding task. It is possible that these results replicate those of Mine showing that distant manipulation results in higher completion times. Indeed, Mine et al. use a random distance from 10 to 60 cm for each trial, which could be considered as a manipulation distance of 35 cm on average, thus being much closer to our 40 cm condition, than the 0 to 20 cm area. However, the fact that the distance is not fixed in Mine's experiment is a important shift in settings between the two experiments.

We propose to interpet these results in terms of consistency among visual and propioceptive information, and their respective informativeness relative to the task. For the first set of distances, we can consider that the frame of reference for the user is his hand, meaning that the task is completed by bringing the hand close to the desired location. In such manipulation, visual and proprioceptive information are strongly related and the user can easily use them jointly. For the second set of distances, we can consider that the frame of reference is the cursor, meaning that the task is done by bringing the cursor, not the hand, to the desired location. In this case, joint use of visual and proprioceptive information requires an additional treatment, because the two types of information are not as closely related as when the hand is the reference. The control of hand movement may be more complex because the user has to take into account the difference between the frame of reference (the cursor) and the frame associated with his hand. This additional treatment could explain the decrease in performance for the 40 to 55 cm set of distances.

Another explanation for the decrease in performance for the 40 cm to 55 cm interval could be a change in depth perception. Work by Yoshida et al. [9] addressed the effect of the distance between the eyes and the screen on the perceived depth of a virtual object in a stereoscopic display. Using four values of distance between 60 and 100 cm, they found that the greater the distance, the more users underestimate the distance to the object. The error is always inferior to 10 mm. In our experiment, the users slightly moved back (maximum 40 cm) for the 40 and 55 cm offsets in order to have their arm in a comfortable position. Since head tracking was used, the user is likely to present the same depth perception errors indicated by Yoshida et al. However, these errors are small compared to the size of the spheres used for the task (5 cm diameter), suggesting that the biases due to depth perception are not likely to affect the user performance in any way.

For the second experiment, our hypothesis was that adding a ray (visual clue) that bridges the gap between hand and cursor would help the user. But on the contrary, results show that user performance is significantly lower when the ray is displayed[3]. This experiment was done with a 55 cm distance of manipulation where, according to the first hypothesis, the user would choose the cursor as the frame of reference. One hypothesis that can explain the results for *visual clue* is that visually linking the cursor and the hand forces the user to use his hand as the frame of reference, whereas in a normal distance manipulation (without visual clue) he would have chosen the cursor as the frame of reference. This might result in a cognitively confusing situation. However, further investigation is needed to address this cognitive issue.

The results for *target position* in the first experiment show that p8 and p9 take significantly longer to be clicked. We can raise different hypotheses. Target p9 is the furthest target from the user's eyes, so it is visually smaller and can be harder to see than the other targets. Target p8 is located in the lower-left manipulation area, and can be hidden by the arm of left-handed users when they click on the start sphere: this can lead to a greater time to find the target. The mean completion time for most of the off-axis targets (all except p6 and p7, which are faster to reach than p4) is higher than for on-axis targets, but these results are not statistically significant and further work is required to address this problem. Comparison between on-axis target completion times shows a result that is consistent with the results of a study of 3D point location done by Boritz and Booth [1]: up and down (p1 and p2) targets take more time to reach than left and right targets (respectively p3 and p4). In our setting, these differences are not statisically significant. Boritz and Booth found one significant difference between left and down targets (p2 and p4).

Significant interaction effects between *offset* $\times$ *target position* show that the difference for p8 and p9 compared to other targets is higher for 40 and 55 cm offsets. One explanation could be that the treatment for manipulation at 40 and 55 cm is more important as the offset grows, particularly for these two off-axis targets.

The results for the scale factor in the second experiment show that users manipulating with a 1 to 1 translation mapping (scale=1) perform better than with a variable offset (scale=1.5). This result is consistent with previous work done by Mine et al.[6]. The effect of the scale can be interpreted in light of the hypothesis on consistency between proprioceptive and visual information: as a higher offset induces more information treatment to control the movement, scale may also be a limit because it adds a supplementary difference between visual and proprioceptive information.

Subject comments on manipulation distance are consistent with the performance variation that we measured. Most users rate the closer manipulation to be easier. As distance of manipulation grows, subjects rate it to be harder. It is surprising to see that users prefer direct manipulation, whereas the reaction of users that are new to immersive displays is to manipulate at a distance from the virtual objects. In this experiment, for direct manipulation the users had to get close to the objects, otherwise they couldn't complete the task. This obligation may make them realize that direct manipulation is more comfortable. This tends to support the decision to give the users instructions to manipulate at the closest possible distance. These results are consistent with user comments retrieved by Mine and Djajadiningrat, where users prefer direct manipulation. In Poupyrev and al.'s experiment, all but three users do not prefer the direct manipulation technique, which could be due to the fact that the classical virtual hand technique is perceived as less innovative than the other techniques, particularly the go-go technique.

[3] We discarded the hypothesis that a bias could be due to a change in image refresh rate by checking that it is the same for both conditions.

6. Conclusion and Future Work

An experiment was conducted to study the influence of distance of manipulation on a location task. Results show that manipulation at a closer distance (direct manipulation to 20 cm) is significantly more efficient than for 40 cm to 50 cm in our setting. The second experiment studied the influence of adding a visual clue or a scale on performance in the same location task. It was found that adding a visual clue (virtual ray) to fill the gap between the hand and the cursor lowers user performance, and also that using a 1.5 scale factor to map user movement to the cursor significantly lowers task performance.

To be able to use these results as rationales for 3D user interface design, it is important to replicate them on similar systems in order to check their stability.

In future work, we want to focus on the effect of training on performance in 3D location tasks within the context of direct and distance manipulation. The objective is to see how the performance between direct and distant manipulation can change with training, and particularly if distant manipulation can become as efficient as direct manipulation. Secondly, we plan to study the variation in performance in the transition interval (between 20 cm and 40 cm distance of manipulation) to see if, for example, there is a gap value of the offset for which performance significantly changes, or if performance decreases continuously.

Acknowledgements

This study was partially funded by the Regional Council of Ile-de-France and the French Ministry of Research within the RNTL Perf-rv project. The Responsive Workbench used in the study has been partially funded by the Regional Council of Ile-de-France. The authors want to thank M. Stephen Kreisler and all the subjects who participated in the experiment.

References

[1] J. Boritz and K. Booth. A study of interactive 3d point location in a computer simulated virtual environment. In *ACM VRST '97*, 1997.

[2] J. Djajadiningrat. *Cubby : What You See is Where You Act. Interlacing the display and manipulation spaces*. Ph.D. thesis, Delft University of Technology, 1998.

[3] J. Djajadiningrat, C. Overbeeke, and P. Stappers. Cubby: A unified interaction space for precision manipulation. In *Proc. of ITEC*, 2001.

[4] W. Kruger, C.-A. Bohn, B. Frohlich, H. Schuth, W. Strauss, and G. Wesche. The responsive workbench: A virtual work environment. *IEEE Computer*, 28(7):42–48, 1995.

[5] M. Mine. Exploiting proprioception in virtual-environment interaction. Technical report, TR97-014 UNC Chapel Hill Computer Science, 1997.

[6] M. R. Mine, F. P. Brooks, Jr., and C. H. Sequin. Moving objects in space: Exploiting proprioception in virtual-environment interaction. *Computer Graphics*, 31(Annual Conference Series):19–26, 1997.

[7] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The go-go interaction technique: Non-linear mapping for direct manipulation in VR. In *ACM Symposium on User Interface Software and Technology*, pages 79–80, 1996.

[8] I. Poupyrev, S. Weghorst, M. Billinghurst, and T. Ichikawa. Egocentric object manipulation in virtual environments : Empirical evaluation of interaction techniques. *In Proc. of Eurographics*, 1998.

[9] S. Yoshida, T. Hoshino, T. Oseki, S. Miyazaki, J. Hasegawa, and T. Fukumura. An improvement method for depth perception in stereoscopic display. In *International Archives of Photogrammetry and Remote Sensing.*, pages 42–48. 1998.